Identification of Pyroptosis-related Gene Prognostic Signature in Patients with Hepatocellular Carcinoma

Ruipeng Chen

Department of pathology, The Second Xiangya Hospital, Central South University

Pansong Li

Geneplus-Beijing Research Institute

Mengmeng Song

Geneplus-Beijing Research Institute

Liyan Ji

Geneplus-Beijing Research Institute

Yi Sun (✉ yi_sun@csu.edu.cn)

Department of pathology, The Second Xiangya Hospital, Central South University

Xuan Gao

Geneplus-Beijing Research Institute

Xuefeng Xia

Geneplus-Beijing Research Institute

Research Article

Keywords: Hepatocellular carcinoma, Pyroptosis, Signature, Prognosis, Nomogram, Bioinformatics

Posted Date: February 3rd, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1270464/v1>

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

[Read Full License](#)

Abstract

Background:

Pyroptosis has been recently identified as a hallmark of cancer biology; however, the potential of pyroptosis-related genes (PRGs) as prognostic markers has not been fully elucidated in hepatocellular carcinoma (HCC).

Methods:

35 PRGs were obtained from the published literature, and pyroptosis subtypes were identified by bioinformatics methods. The risk score model was established applying the LASSO Cox regression method in the TCGA cohort and validated in the ICGC datasets. Additionally, immune infiltration, enriched pathways and genomic alterations were analyzed between high and low risk-score subgroups. Finally, a nomogram containing the pyroptosis risk score and other prognostic related clinical factors was created for predicting the overall survival of HCC.

Results:

Based on the expression profile of PRGs, we determined two pyroptosis-related subtypes (cluster A and cluster B) of HCC with different immune characteristics and significantly different prognoses. The risk score model showed that up-regulation of *GPX4*, *CASP8*, *NOD2*, and *GSDME* was associated with poor prognosis, while *NLRP6* was the opposite. Compared with patients with lower risk scores, the high-risk score group had a worse prognosis ($P < 0.0001$) and a higher level of immune cell infiltration. Functional analysis suggested that the highly expressed genes in the high-risk group were mainly concentrated in various signaling pathways, while the low-expressed genes in the high-risk group were mainly enriched in different biochemical metabolic translations. Genomic alterations in high-risk and low-risk populations implied that mutations in the TP53 gene are highly associated with pyroptosis in HCC patients. A nomogram including risk score and TNM stage for 1-year, 3-year, and 5-year overall survival indicated a good prognostic prediction ability.

Conclusions:

We developed and verified a prognostic risk model based on PRGs for HCC sufferers, which may provide a robust tool for forecasting prognosis in patients with HCC.

Background

Hepatocellular carcinoma (HCC), which accounts for the most significant proportion of liver cancers, is one of the most common malignancies in humans and the fourth leading cause of cancer-related deaths worldwide¹. Studies have shown that pathogenic factors for HCC include hepatitis B and C viruses, alcohol abuse, dietary toxins such as aflatoxin and aristolochic acid, metabolic liver disease (mainly non-alcoholic fatty liver), and exposure to carcinogenic substances such as hexachlorobenzene². In the past

few decades, due to the popularization of health knowledge and advanced screening procedures, including alpha-fetoprotein (AFP) detection and non-invasive imaging technology, significant progress has been made in the treatment of HCC²⁻⁶. However, the survival rate of advanced HCC is still poor⁷. Nowadays, the indicator of tumor lymph node metastasis (TNM) staging is widely applied in clinical neoplasm management for predictive risk assessment and treatment decisions. Nevertheless, due to the high degree of molecular heterogeneity, even among patients with the same clinicopathological characteristics, the risk of recurrence and death may vary greatly^{8,9}. Therefore, there is an urgent need for new prognostic signatures to identify the risk of HCC patients more accurately.

Pyroptosis, also referred to as inflammatory cell death, is defined as gasdermin-mediated programmed necrosis^{10,11}. The characteristic of cell pyroptosis is that gasdermin is activated to form cell membrane pores, which is different from the mechanism of cell apoptosis. Subsequently, the intact cell membrane is destroyed under the dual action of pro-inflammatory cytokines and alarm factors, eventually leading to cell death¹⁰. The initial research on pyroptosis was mainly focused on its function in fighting infections. Nowadays, a large number of studies have shown that pyroptosis plays an indispensable role in the evolution of human malignancies. The interactions between pyroptosis and cancers are incredibly complicated¹². On the one hand, the growth of tumors can be inhibited by mediating the pyroptosis of cancer cells. Still, on the other hand, pyroptosis provides nutritional and environmental support for the occurrence and development of neoplasm¹³. Increasing evidence has highlighted the role of pyroptosis in cancer^{13,14}.

Research on pyroptosis-related genes (PRGs) and ovarian cancer, lung adenocarcinoma, and gastric cancer have been published recently¹⁵⁻¹⁷. The molecular characteristics, clinical significance, and immune interaction of PRGs in the above-mentioned cancer species have been elucidated by applying bioinformatics methods. These studies revealed that using PRGs as a new gene signature to forecast the survival prognosis of cancer suffers is extremely valuable.

However, the expression levels, clinical prognostic characteristics, and immune infiltration status of the PRGs in HCC have not yet been clarified, making our research significant. In this study, gene expression data from The Cancer Genome Atlas (TCGA) were employed to identify molecular subtypes of HCC based on genes related to pyroptosis. To develop a prognostic risk model, five genes were filtered from 35 PRGs by LASSO-Cox regression. The risk model can assess the survival status of HCC suffers and be verified by the International Cancer Genome Consortium (ICGC) external validation cohorts. This risk score model can be utilized as an independent prognostic evaluation mark for suffers from HCC.

Results

Defining of the expression of PRGs between normal and tumor tissues

Normal tissues include normal human liver tissue samples in GTEx and adjacent normal tissues in TCGA, and tumor tissues derived from TCGA. We have determined 34 differentially expressed genes (DEGs)

related to pyroptosis in 160 normal tissues and 343 tumor tissues (all P -value <0.01) (Fig 1A). Among them, 28 genes (*PLCG1*, *CASP1*, *CASP4*, *CASP5*, *CASP6*, *CASP8*, *CASP9*, *PJKK*, *ELANE*, *PRKACA*, *GPX4*, *GSDMB*, *TIRAP*, *GSDMD*, *GSDME*, *SCAF11*, *GZMB*, *IL18*, *IL1B*, *IL6*, *NOD2*, *NLRC4*, *NLRP1*, *TNF*, *NLRP2*, *NLRP3*, *NLRP6*, and *NOD1*) were downregulated in the tumor group while six other genes (*CASP3*, *GSDMA*, *GSDMC*, *GZMA*, *PYCARD*, and *AIM2*) were enriched in the tumor group.

In order to further understand the interaction of these pyroptosis-related genes at the protein level, we constructed a PPI network. The minimum interaction score required for PPI analysis was set to 0.4 (the default recommended value). Through MCC, DMNC, and MNC algorithms in Cytoscape, we determined that *NLRP1*, *PYCARD*, *AIM2*, *IL18*, *CASP1*, *NLRC4*, *CASP4*, *CASP8*, *NLRP3*, and *TNF* were hub genes (Fig 1B). These hub genes are also DEGs between the normal specimens and HCC.

Molecular subtypes based on pyroptosis-related genes

To investigate the relationship between the expression profiles of 35 PRGs and HCC subtypes, we conducted a consistent cluster analysis on all 343 HCC samples in the TCGA training cohort. By adjusting the clustering variable (k), we observed that when $k=2$, the inter-group correlation was lowest and the intra-group correlation was highest, suggesting that based on 35 PRGs, 343 HCC patients could be well divided into two robust clusters (cluster A contains 59 patients and cluster B contains 284 patients) (Fig 1C and 1D).

Prognostic differences and immune landscape of the subtypes

The prognosis signature among molecular subtypes was further analyzed. Survival analysis in OS showed that cluster A had a worse prognosis than cluster B (Figure 2A). We applied ssGSEA to study the immune cell infiltration of the two molecular subtypes of the pyroptosis gene and observed that compared with cluster-B, cluster-A not only had more infiltration of innate immune cells (macrophages, mast cells, and immature dendrites cells and MDSC) but also adaptive immune cells, such as activated B cells and activated CD4 T cells, effector memory CD4 T cells and effector memory CD8 T cells (Fig 2B). Interestingly, in the 17 immune-related pathways, cluster-A was also more abundant than cluster-B (Figure 2C). In addition, through the ESTIMATE package, we found that cluster A had higher stromal scores, immune scores, and ESTIMATE scores than cluster B, but the tumor purity was the opposite (Figure 2D), which was consistent with the above findings (Figure 2B; Figure 2C). We noticed that patients with poor prognoses among the subtypes of pyroptosis molecules were more abundant in immune infiltration, revealing the complex relationship between the pyroptosis pathway and the immune microenvironment in HCC.

Development of a Prognostic Risk Model in the Training Dataset

The risk prognostic model of PRGs was generated based on the TCGA cohort. First, we used LASSO-Cox regression analysis to reduce the amount of PRGs. The minimum lambda (0.06544434) was obtained by

ten-fold cross-validation to construct our prognostic model containing five genes (Fig 3A; Fig 3B). The formula of the pyroptosis risk model was as follows:

Risk score=

$$0.004182175*GPX4+0.091401425*CASP8+0.018589062*NOD2+0.091845249*GSDME-0.031080493*NLRP6$$

According to the above formula, the up-regulation of *GPX4*, *CASP8*, *NOD2*, and *GSDME* was associated with poor prognosis, while *NLRP6* was the opposite. With the best cutoff value of risk score obtained by the "survminer" package (cutoff = -0.01471409), sufferers in the TCGA cohort were well divided into high-risk groups and low-risk groups. In Figure 3C, an obvious difference in the KM survival curve between the two subgroups was observed ($P<0.0001$). Furthermore, in Figure 3D, we followed that the basic conditions of the high-risk and low-risk groups include survival and death status, population distribution trends, and the expression heat map of the five genes that made up the risk score. The AUC in the 1-year, 3-year, and 5-year ROC curves were 0.63, 0.65, and 0.60, respectively (Figure 3E), indicating that the risk score had high sensitivity and accuracy in forecasting the prognosis of hepatocellular carcinoma.

External validation of the prognostic risk model

The ICGC cohort was regarded as an external validation dataset to verify the robustness of the risk model established by the TCGA training cohort. Using the same method as described above, the ICGC cohort was divided into high-risk and low-risk groups based on its best cutoff value (cutoff = -0.00717451). The KM survival curve illustrated that the prognosis of the high-risk group is much worse than that of the low-risk group (Figure 4A). Figure 4B displayed the overall trend of the ICGC cohort, including survival and death status, population distribution, and five prognostic PRGs expression levels, which is consistent with the TCGA cohort. In addition, the ROC curve generated based on the risk score was applied to evaluate the 1-year, 3-year, and 4-year prognosis of the validation cohort, as shown in Figure 4C.

Comparison of the immune activity between subgroups

We further investigated the distinction in the abundance of immune cell infiltration between the high-risk score group and the low-risk score group. Through the "GSVA" package, we observed that CD56dim natural killer cells and eosinophil infiltration in the low-risk score group were relatively higher than in the high-risk group. In contrast, the cells with higher infiltration in the high-risk score group included activated CD4 T cells, central memory CD4 T cells, effector memory CD4 T cells, $\gamma\delta$ T cells, immature dendritic cells, MDSC, memory B cells, and natural killer T cells, plasmacytoid dendritic cells, follicular helper T cells, and type 2 helper T cells (Figure 4D). Subsequently, by the TIMER database, compared with the low-risk group, we found that the high-risk score group had a higher degree of infiltration of B cells, CD4 T cells, CD8 T cells, neutrophils, macrophages, and dendritic cells (Figure 4E).

We next explored relationships between molecular subtypes of pyroptosis and the risk score. The higher risk score was found in patients in cluster-A (Fig 4F), previously confirmed to have a poor prognosis.

Analysis of Enrichment Pathways and Genome Changes between Risk Groups

We employed the Wilcoxon rank-sum test (implemented in the R software) to evaluate differentially expressed genes between the high-risk and low-risk scoring groups. The threshold was set as $FDR < 0.05$ and fold change (expressed in $\log_2$ (average FPKM ratio between the high-risk group and low-risk group) ≥ 1). Using the above settings, we found that in the high-risk group of the TCGA cohort, there were 2629 significantly up-regulation genes and 76 significantly down-regulation genes (Additional file: Figure S1). Refer to the attached file for detailed illustrations of highly expressed genes and low expressed genes (Table S3; Table S4).

GO analysis showed that high-expressed genes in the high-risk group were mainly enriched in various synaptic signaling pathways, embryonic organ morphogenesis, and G protein-coupled receptor signaling pathway (Figure 5A). However, low-expressed genes in the high-risk group were mainly enriched in various biological metabolic processes, such as the xenobiotic metabolic process, lipid hydroxylation, and cellular response to the xenobiotic stimulus. (Figure 5B). KEGG analysis revealed that upregulated genes were mainly enriched in neuroactive ligand-receptor interaction, nicotine addiction, and Rheumatoid arthritis (Figure 5C), while downregulated genes were enriched primarily on pathways such as chemical carcinogenesis, retinol metabolism, and drug metabolism (Figure 5D).

As shown in Figure 5E and Figure 5F, the 20 most frequently mutated genes in the high-risk and low-risk cohorts were not wholly consistent, which might reveal the differences in the genomic alteration of the risk score group. Based on the frequency of mutation, the mutated genes in the high-risk cohort included *TP53*, *TTN*, *CTNNB1*, *MUC16*, *PCLO*, *AXIN1*, *XIRP2*, *APOB*, *RYR1*, *ABCA13*, *CSMD3*, *FAT3*, *LRP1B*, *ARID1A*, *CACNA1E*, *CCDC168*, *RYR2*, *ABCB5*, *COL6A6*, and *FRAS1* (Figure 5E). The mutant genes in the low-risk cohort were *CTNNB1*, *TTN*, *TP53*, *MUC16*, *ALB*, *MUC4*, *RYR2*, *ABCA13*, *APOB*, *OBSCN*, *PCLO*, *BAP1*, *USH2A*, *ARID2*, *FLG*, *HMCN1*, *LRP1B*, *SPTA1*, *CUBN*, and *PRKDC* (Figure 5F). We used Fisher's exact test (P threshold was set to < 0.05) to explore different mutant genes in the high and low-risk groups and found that *TP53* occupies the first position (Figure 5G). This result implied that *TP53* has a high correlation with pyroptosis in patients with HCC. The lollipop chart revealed the different mutation sites of *TP53* between the two cohorts (Figure 5H), and the difference in survival prognosis of *TP53* mutation in HCC was observed (Figure 5I). In addition, as shown in Figure 5J and Figure 5K, in the co-occurrence and mutually exclusive mutations of the high and low-risk groups, we observed a particular *TP53-CSDM1* co-occurrence mutation, suggesting that their respective mutations in HCC might cause a combined effect.

Establishment and Assessment of a Nomogram based on Risk score

Univariate and multivariate Cox analysis suggested that the pyroptosis risk score could be identified as an independent prognostic signature to forecast the survival status of HCC (Table 1). Additionally, we integrated the risk score and prognostic clinical characteristics to develop a nomogram to forecast the 1-year, 3-year, and 5-year overall survival rates of sufferers with hepatocellular carcinoma (Figure 6A). The AUC of the time-dependent ROC curve illustrated that the performance of the nomogram prediction was better than that of the risk score alone (Figure 6B). The result of the KM curve of the OS of the nomogram

was significant (Figure 6C). Furthermore, compared with the ideal model, the calibration chart showed that the nomogram with the risk score and TNM stage performed well (Figure 6D).

Discussion

HCC is highly malignant and progresses rapidly, and the prognosis of patients with HCC is usually poor. The symptoms of early HCC are insidious, and most patients with HCC are at the advanced stage when they are diagnosed, which means that most of them have already lost the opportunity to undergo surgical liver resection. Although radiotherapy, chemotherapy, liver transplantation, and other potential treatment methods have produced various effects and prolonged the survival time of patients, the prognosis of HCC is still unsatisfactory owing to the high risk of recurrence and intrahepatic spread. Pyroptosis is a unique sort of programmed cell death characterized by the swelling and lysis of cells and the release of many pro-inflammatory factors. The relationship between malignancies and pyroptosis is complex, and pyroptosis of different human tissues and genetic landscapes has different effects on cancer. With the continuous deepening of tumor molecular biology research, prognostic gene signatures reflecting malignancies progression have received more attention in predicting survival in HCC, which may help achieve more accurate individualized treatment management. This study is significant in exploring the relationship between pyroptosis genes and hepatocellular carcinoma.

Herein, we clarified the mRNA levels of 35 PRGs from the published literature in HCC and normal tissues and determined that most of them are differentially expressed. Compared with healthy liver tissue, the expression of 28 PRGs decreased in HCC, and 6 PRGs increased in HCC. Then, we determined that *PYCARD*, *AIM2*, *IL18*, *CASP1*, *CASP4*, *CASP8*, *NLRC4*, *NLRP1*, *NLRP3*, and *TNF* were hub genes in the pyroptosis pathway.

We also applied the unsupervised clustering algorithm to distinguish pyroptosis-related subtypes (cluster-A and cluster-B). Cluster-A was overexpressed in immune cells and immune-related pathways, and the prognosis was worse than cluster-B. Due to the prognostic effects of the PRGs mentioned above, we employed LASSO Cox regression analysis to develop a predictive risk model based on 5 PRGs (*GPX4*, *CASP8*, *NOD2*, *GSDME*, and *NLRP6*). This model had high accuracy and practicability in forecasting the 1-year, 3-year, and 5-year overall survival of HCC sufferers. Subsequently, we used the ICGC validation set to validate the dependability of the model. The results indicated that *GPX4*, *CASP8*, *NOD2*, and *GSDME* were risk factors, and *NLRP6* was a protective factor for HCC prognoses.

In recent years, glutathione peroxidase 4 (*GPX4*) has been considered a vital regulatory factor of ferroptosis tumor cell death. Unlike other *GPXs*, *GPX4* has the function of catalyzing the reduction of lipid peroxides in a complex cell membrane environment, indicating that *GPX4* has a distinctive role in physiology¹⁸. Inhibition of *GPX4* can induce the death of cancer cells that are resistant to conventional chemotherapy or radiotherapy^{19,20}. Herein, we revealed that up-regulation of *GPX4* is associated with poor prognosis and a high risk of hepatocellular carcinoma in patients. In this way, inhibiting the high

expression of *GPX4* may offer a novel therapeutic strategy for reducing the mortality rate in patients with HCC.

Caspase-8 (*CASP8*) plays an essential enzyme in the pyroptosis pathway²¹. The significance of *CASP8* in initiating necrosis receptor-induced pyroptosis and maintaining immune homeostasis and surveillance is well established. It has long been held that the expression levels of *CASP8* in most tumors are usually down-regulated, leading to pyroptosis escape and enhancing resistance to anticancer therapies²². Multiple studies have shown that, compared with healthy liver tissues, the expression levels of *CASP8* are down-regulated in HCC, which is consistent with our results^{23–25}. Nevertheless, the relationship between the levels of *CASP8* expression and the prognosis of HCC patients has not yet been confirmed. Our analysis revealed that up-regulation of *CASP8* could increase the mortality risk of in HCC which may be related to the RIPK1/FADD/Caspase-8 signal complex causing liver parenchymal cell damage, thereby impairing the normal functioning of the liver²⁶.

Nucleotide-binding oligomerization domain 2 (*NOD2*), a recognized innate immune sensor, has a significant effect on carcinogenesis²⁷. It was reported that *NOD2* dysregulation is involved in the pathogenesis of Crohn's disease (CD) and colitis-related colon cancer. *NOD2* gene polymorphism was associated with colorectal cancer, lymphoma, lung cancer, gastric cancer, ovarian cancer, laryngeal cancer, and breast cancer²⁸. In their research on HCC, Ma et al. showed that *NOD2* restrains tumorigenesis and enhances tumor chemotherapeutic sensitivity through targeted regulation of the AMPK pathway, which may reveal a novel strategy for the treatment of HCC based on *NOD2* adjustment²⁹. A recent study on HCC demonstrated that *NOD2* initiates the nuclear autophagy pathway, leading to DNA damage and increased genomic instability, and proved that *NOD2* as a poor prognostic factor is closely related to the increased risk of death in HCC patients³⁰. Our conclusion that overexpression of *NOD2* had a bad prognosis and high-risk score is consistent with the above research.

Gasdermin E (*GSDME*), also known as *DFNA5*, was first identified as a mutated gene inherited in an autosomal dominant manner, causing the loss of cochlear hair cells and eventually developing into progressive, sensorineural and non-syndromic hearing impairment. In recent years, multiple studies have shown that *GSDME*, a member of the Gasdermin family, is closely related to cancer³¹. Compared with healthy human tissues, due to epigenetic suppression of methylation, the expression of *GSDME* is down-regulated in gastric cancer, colorectal cancer, breast cancer, and most human cancer cell lines¹⁰. This is consistent with our research results. *GSDME* is a molecular drug target for treating human malignancies, and clinicians can choose appropriate chemotherapeutic drugs according to its expression levels to improve the sensitivity of chemotherapeutic drugs and reduce drug resistance¹³.

NLRP6 (originally called *PYPAF5*) is a novel node-like receptor (NLR) family member that produces inflammasomes. *NLRP6* consists of an N-terminal pyrin domain, a central NACHT domain, and a terminal leucine-rich repeat (LRR) domain³². Nowadays, the majority of research on *NLRP6* has concentrated on the intestine in the mouse model, and only a few studies have involved human patients to date^{32,33}. The

survey by Domblides et al. determined that down-regulation of *NLRP6* is related to the poor prognosis of human colorectal cancer³⁴. Our research revealed that *NLRP6* is a protective factor for the prognosis of hepatocellular carcinoma, which provides new clues for the further study of *NLRP6* in HCC.

It is well known that the tumor microenvironment plays a vital role in anti-tumor molecular therapy. Therefore, we also explored the abundance of immune and stromal cells in the high-risk and low-risk groups. Interestingly, unlike the phenomenon observed in ovarian cancer and gastric cancer^{16,17}, in our research, the degree of immune cells and stromal cells infiltration of patients in the high-risk group was higher than that in the low-risk group. This may partly indicate that the mechanism of pyroptosis genes in the microenvironment of hepatocellular tumors is more complicated. Unlike the low-risk group related to various biological metabolic pathways, the high-risk group mainly participated in signal transduction pathways. It is worth noting that the high-risk group is also significantly associated with nicotine addiction and autoimmune diseases (rheumatoid arthritis). By comparing the genomic changes in different risk groups, we observed that the high-risk group was significantly associated with more aggressive molecular alteration (such as TP53 mutations). More importantly, the combination of univariate and multivariate Cox analysis illustrated that the nomogram using the pyroptosis risk score and TNM stage has high prediction capacity and may serve as a meaningful indicator in personalized medicine.

Our research had some limitations. First, the pyroptosis risk score based on five genes was constructed and validated based on retrospective cohorts. Moreover, further in vivo and in vitro experiments could be applied to further verify our results.

Conclusions

In summary, we determined two pyroptosis molecules subtypes in hepatocellular carcinoma patients with different prognoses. A risk model with good prognostic performance containing five genes (*GPX4*, *CASP8*, *NOD2*, *GSDME*, and *NLRP6*) was constructed. This risk prognosis model could be applied as a prognostic predictor for the survival status of sufferers with hepatocellular carcinoma.

Materials And Methods

Datasets and preprocessing

The TCGA-liver hepatocellular carcinoma (LIHC) data set was composed of the gene expression RNA sequencing (RNA-seq) data of HCC and adjacent normal samples and corresponding clinical prognostic traits and was obtained from the TCGA database (<https://portal.gdc.cancer.gov/>). The validation group data of ICGC contained the gene expression data of HCC specimens with clinical characteristics and was received from the ICGC data portal(<https://icgc.org/>). After removing data with no clinical information and OS < 30 days, this study included 343 HCC and 50 adjacent normal samples in the TCGA cohort and 228 HCC samples in the ICGC cohort. The gene expression RNA-seq data of 110 normal human liver

tissue samples were downloaded from the GTEx data set (<https://xenabrowser.net/datapages/>). The somatic mutation profile was gained from the TCGA data portal (<https://portal.gdc.cancer.gov/>). Subsequently, we used the R package "maftools" to process somatic mutation profiles sorted in the form of mutation annotation format (MAF).

Identification of differentially expressed pyroptosis-related genes

We collected 35 PRGs from the published literature^{13,15-17,35}, including *GSDMD*, *CASP5*, *GSDME*, *GSDMB*, *PRKACA*, *SCAF11*, *NOD2*, *NLRP2*, *CASP3*, *CASP9*, *IL18*, *CASP4*, *ELANE*, *GPX4*, *GSDMC*, *PLCG1*, *PYCARD*, *AIM2*, *CASP8*, *TIRAP*, *NLRP6*, *NLRP7*, *NLRP3*, *IL6*, *CASP6*, *TNF*, *GSDMA*, *NOD1*, *NLRP1*, *PJVK*, *IL1B*, *NLRC4*, *CASP1*, *GZMA*, and *GZMB*. In the gene differential expression analysis, there are differences in the magnitude of the HCC samples and the normal tissue samples adjacent to cancer in the TCGA database, which may lead to statistical errors in the results. Therefore, we also included normal liver tissue samples from the GTEx database in this study. This website (<https://xenabrowser.net/datapages/>) has already processed the data to remove the batch effect, so that we can use it directly. Then, the mRNA expression of PRGs was compared between HCC and normal samples using the Wilcoxon test.

Protein-protein interaction networks and hub genes

The online database STRING³⁶ (v11.0, <http://www.string-db.org/>) was used to visualize the PPIs (Protein-Protein Interaction Networks) between 35 pyroptosis-related genes. Cytoscape^{37,38} (<https://cytoscape.org/>) was used to visualize the PPI network of PRGs and calculate hub genes.

Identification of pyroptosis-related subtypes

The RNA-seq expression profiles of 35 PRGs were used to perform k-means unsupervised clustering through the "ConsensusClusterPlus" package, which was repeated 1000 times^{39,40}.

Estimation of immune infiltration and immune-related pathway activity

In order to clarify the immune infiltration status of each specimen in the TCGA data set corresponding to the pyroptosis subtype, we used single-sample gene set enrichment analysis (ssGSEA) through the "GSVA" package based on immune cell gene sets and immune-related pathways obtained from published literature^{41,42} (Additional file: Table S1; Table S2). In addition, we applied the ESTIMATE algorithm as a tool assessing each tumor specimen to compute the immune score, stromal score, ESTIMATE score, and tumor purity⁴³.

Construction of the risk model based on PRGs

We directly used the least absolute shrinkage and selection operator (LASSO) penalty method to further narrow down the prognostic-related pyroptosis genes to construct a prognostic risk model. Then, we determined the best parameter λ through ten-fold cross-validation. According to the optimal cut-off value calculated by the "survminer" package, we divided the TCGA cohort into the high-risk and low-risk groups

based on the prognostic risk score obtained from the screened genes. In addition, we plotted time-dependent receiver operating characteristic (ROC) curves by the "timeROC" package, to assess the sensitivity and specificity of the risk model in the TCGA and ICGC cohorts.

Correlations of the risk score with immune infiltration

We applied the above-mentioned ssGSEA method to estimate the difference in infiltration levels of 28 immune cells between the high-risk score group and the low-risk score group in the TCGA cohort. However, using a single algorithm and marker gene set alone to evaluate tumor-infiltrating immune cells (TIIC) may cause calculation errors. To avoid these problems, we also used the TIMER database (<https://cistrome.shinyapps.io/timer/>) to quantify the immune infiltration per specimen to make the results more convincing⁴⁴.

Gene function and pathway enrichment analysis

WebGestalt (<https://www.webgestalt.org/option.php>) was utilized to explore the function and pathway enrichment of target genes⁴⁵. First, we utilized Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) database overexpression enrichment analysis method to input the genes of interest to the WebGestalt server. Then, we applied the "ggplot2" package to plot the results.

Construction and validation of the prognostic nomogram

Univariate and multivariate Cox regression analyses were utilized to evaluate the survival prognosis of pyroptosis risk score and clinical characteristics. Subsequently, we integrated the risk score and clinical characteristics that have prognostic value to construct a prognostic nomogram. Calibration curves for 1-year, 3-year, and 5-year survival rates were applied to assess the deviation of the nomogram from the ideal model. The ROC curve was constructed and the area under the curve (AUC) was calculated to evaluate the survival prognosis prediction ability of the nomogram containing the pyroptosis gene risk score and TNM staging.

Statistical analysis

Data were analyzed with the R (version 4.0.5) and R Bioconductor packages. Comparisons between two groups were evaluated with the student's-test or Wilcoxon test. ROC curves were generated with the "timeROC" package. The nomogram was developed with the "rms" package. Overall survival (OS) was compared in pyroptosis molecular subgroups and risk score groups with Kaplan-Meier curves and log-rank tests. Unless otherwise stated, all statistical tests were two-way, and a *P*-value <0.05 was considered significant.

Abbreviations

PRG: pyroptosis-related gene

HCC: hepatocellular carcinoma

AFP: alpha-fetoprotein

TNM: tumor lymph node metastasis

TCGA: The Cancer Genome Atlas

ICGC: International Cancer Genome Consortium

MAF: mutation annotation format

PPI: Protein-Protein Interaction

ssGSEA: single-sample gene set enrichment analysis

LASSO: least absolute shrinkage and selection operator

ROC: receiver operating characteristic

TIIC: tumor-infiltrating immune cell

GO: Gene Ontology

KEGG: Kyoto Encyclopedia of Genes and Genomes

AUC: area under the curve

OS: Overall survival

DEG: differentially expressed gene

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

The data that support the findings of this study are available from the corresponding author upon reasonable request. The raw data were obtained from the TCGA database (<https://portal.gdc.cancer.gov/>), ICGC database (<https://icgc.org/>), and GTEx database

(<https://xenabrowser.net/datapages/>), respectively. Details of the data processing steps can be found in the Materials and methods section of the article.

Competing interests

The authors have no conflict of interest related to this publication.

Funding

This study was supported by a grant from the Natural Science Foundation of Hunan Province, China (No. 2019JJ40422).

Author Contributions:

Designed the research, analyzed the data and wrote the manuscript (RPC, PSL, MMS, LYJ, XG, XFX, YS), participated in data preparation (RPC, PSL, MMS), analysis of data (RPC, LYJ), and figure preparation (RPC). All authors read and approved the manuscript for publication.

Acknowledgements

Not applicable

References

1. Llovet JM, Kelley RK, Villanueva A, et al. Hepatocellular carcinoma. *Nat Rev Dis Primers* 2021; **7**(1): 6.
2. Yang JD, Hainaut P, Gores GJ, Amadou A, Plymoth A, Roberts LR. A global view of hepatocellular carcinoma: trends, risk, prevention and management. *Nat Rev Gastroenterol Hepatol* 2019; **16**(10): 589–604.
3. Llovet JM, Montal R, Sia D, Finn RS. Molecular therapies and precision medicine for hepatocellular carcinoma. *Nat Rev Clin Oncol* 2018; **15**(10): 599–616.
4. European Association for the Study of the Liver. Electronic address eee. Corrigendum to "EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma" [*J Hepatol* 69 (2018) 182–236]. *J Hepatol* 2019; **70**(4): 817.
5. Marrero JA, Kulik LM, Sirlin CB, et al. Diagnosis, Staging, and Management of Hepatocellular Carcinoma: 2018 Practice Guidance by the American Association for the Study of Liver Diseases. *Hepatology* 2018; **68**(2): 723–50.
6. Llovet JM, De Baere T, Kulik L, et al. Locoregional therapies in the era of molecular and immune treatments for hepatocellular carcinoma. *Nat Rev Gastroenterol Hepatol* 2021; **18**(5): 293–313.
7. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018; **68**(6): 394–424.

8. Craig AJ, von Felden J, Garcia-Lezana T, Sarcognato S, Villanueva A. Tumour evolution in hepatocellular carcinoma. *Nat Rev Gastroenterol Hepatol* 2020; **17**(3): 139–52.
9. Zheng H, Pomyen Y, Hernandez MO, et al. Single-cell analysis reveals cancer stem cell heterogeneity in hepatocellular carcinoma. *Hepatology* 2018; **68**(1): 127–40.
10. Liu X, Xia S, Zhang Z, Wu H, Lieberman J. Channelling inflammation: gasdermins in physiology and disease. *Nat Rev Drug Discov* 2021; **20**(5): 384–405.
11. Shi J, Gao W, Shao F. Pyroptosis: Gasdermin-Mediated Programmed Necrotic Cell Death. *Trends Biochem Sci* 2017; **42**(4): 245–54.
12. Xue Y, Enosi Tuipulotu D, Tan WH, Kay C, Man SM. Emerging Activators and Regulators of Inflammasomes and Pyroptosis. *Trends Immunol* 2019; **40**(11): 1035–52.
13. Xia X, Wang X, Cheng Z, et al. The role of pyroptosis in cancer: pro-cancer or pro-"host"? *Cell Death Dis* 2019; **10**(9): 650.
14. Broz P, Pelegrin P, Shao F. The gasdermins, a protein family executing cell death and inflammation. *Nat Rev Immunol* 2020; **20**(3): 143–57.
15. Lin W, Chen Y, Wu B, Chen Y, Li Z. Identification of the pyroptosis-related prognostic gene signature and the associated regulation axis in lung adenocarcinoma. *Cell Death Discov* 2021; **7**(1): 161.
16. Shao W, Yang Z, Fu Y, et al. The Pyroptosis-Related Signature Predicts Prognosis and Indicates Immune Microenvironment Infiltration in Gastric Cancer. *Front Cell Dev Biol* 2021; **9**: 676485.
17. Ye Y, Dai Q, Qi H. A novel defined pyroptosis-related gene signature for predicting the prognosis of ovarian cancer. *Cell Death Discov* 2021; **7**(1): 71.
18. Brigelius-Flohe R, Maiorino M. Glutathione peroxidases. *Biochim Biophys Acta* 2013; **1830**(5): 3289–303.
19. Kim DH, Kim WD, Kim SK, Moon DH, Lee SJ. TGF-beta1-mediated repression of SLC7A11 drives vulnerability to GPX4 inhibition in hepatocellular carcinoma cells. *Cell Death Dis* 2020; **11**(5): 406.
20. Shin D, Kim EH, Lee J, Roh JL. Nrf2 inhibition reverses resistance to GPX4 inhibitor-induced ferroptosis in head and neck cancer. *Free Radic Biol Med* 2018; **129**: 454–62.
21. Ni C, Ye Y, Wang M, et al. A six-nucleotide insertion-deletion polymorphism in the CASP8 promoter is associated with risk of coal workers' pneumoconiosis. *J Toxicol Environ Health A* 2009; **72**(11-12): 712–6.
22. Mandal R, Barron JC, Kostova I, Becker S, Strebhardt K. Caspase-8: The double-edged sword. *Biochim Biophys Acta Rev Cancer* 2020; **1873**(2): 188357.
23. Cho S, Lee JH, Cho SB, et al. Epigenetic methylation and expression of caspase 8 and survivin in hepatocellular carcinoma. *Pathol Int* 2010; **60**(3): 203–11.
24. Yamaguchi Y, Shiraki K, Fuke H, et al. Adenovirus-mediated transfection of caspase-8 sensitizes hepatocellular carcinoma to TRAIL- and chemotherapeutic agent-induced cell death. *Biochim Biophys Acta* 2006; **1763**(8): 844–53.

25. Boege Y, Malehmir M, Healy ME, et al. A Dual Role of Caspase-8 in Triggering and Sensing Proliferation-Associated DNA Damage, a Key Determinant of Liver Cancer Development. *Cancer Cell* 2017; **32**(3): 342-59 e10.
26. Kondylis V, Polykratis A, Ehlken H, et al. NEMO Prevents Steatohepatitis and Hepatocellular Carcinoma by Inhibiting RIPK1 Kinase Activity-Mediated Hepatocyte Apoptosis. *Cancer Cell* 2015; **28**(5): 582–98.
27. Caruso R, Warner N, Inohara N, Nunez G. NOD1 and NOD2: signaling, host defense, and inflammatory disease. *Immunity* 2014; **41**(6): 898–908.
28. Kutikhin AG. Role of NOD1/CARD4 and NOD2/CARD15 gene polymorphisms in cancer etiology. *Hum Immunol* 2011; **72**(10): 955–68.
29. Ma X, Qiu Y, Sun Y, et al. NOD2 inhibits tumorigenesis and increases chemosensitivity of hepatocellular carcinoma by targeting AMPK pathway. *Cell Death Dis* 2020; **11**(3): 174.
30. Zhou Y, Hu L, Tang W, et al. Hepatic NOD2 promotes hepatocarcinogenesis via a RIP2-mediated proinflammatory response and a novel nuclear autophagy-mediated DNA damage mechanism. *J Hematol Oncol* 2021; **14**(1): 9.
31. De Schutter E, Croes L, Ibrahim J, et al. GSDME and its role in cancer: From behind the scenes to the front of the stage. *Int J Cancer* 2021; **148**(12): 2872–83.
32. Li R, Zhu S. NLRP6 inflammasome. *Mol Aspects Med* 2020; **76**: 100859.
33. Ghimire L, Paudel S, Jin L, Jeyaseelan S. The NLRP6 inflammasome in health and disease. *Mucosal Immunol* 2020; **13**(3): 388–98.
34. Domblides C, Soubeyran I, Lartigue L, et al. Prognostic Role of Inflammasome Components in Human Colorectal Cancer. *Cancers (Basel)* 2020; **12**(12).
35. Karki R, Kanneganti TD. Diverging inflammasome signals in tumorigenesis and potential targeting. *Nat Rev Cancer* 2019; **19**(4): 197–214.
36. Szklarczyk D, Franceschini A, Wyder S, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* 2015; **43**(Database issue): D447-52.
37. Doncheva NT, Morris JH, Gorodkin J, Jensen LJ. Cytoscape StringApp: Network Analysis and Visualization of Proteomics Data. *J Proteome Res* 2019; **18**(2): 623–32.
38. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003; **13**(11): 2498–504.
39. Bello-Orgaz G, Menendez HD, Camacho D. Adaptive k-means algorithm for overlapped graph clustering. *Int J Neural Syst* 2012; **22**(5): 1250018.
40. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* 2010; **26**(12): 1572–3.
41. Jia Q, Wu W, Wang Y, et al. Local mutational diversity drives intratumoral immune heterogeneity in non-small cell lung cancer. *Nat Commun* 2018; **9**(1): 5361.

42. Li Y, Jiang T, Zhou W, et al. Pan-cancer characterization of immune-related lncRNAs identifies potential oncogenic biomarkers. *Nat Commun* 2020; **11**(1): 1000.
43. Yoshihara K, Shahmoradgoli M, Martinez E, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun* 2013; **4**: 2612.
44. Li T, Fu J, Zeng Z, et al. TIMER2.0 for analysis of tumor-infiltrating immune cells. *Nucleic Acids Res* 2020; **48**(W1): W509-W14.
45. Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res* 2019; **47**(W1): W199-W205.

Tables

Table1 Univariate and multivariate analyses of the TCGA

Univariable analysis			Multivariable analysis		
Characteristic	HR (95% CI)	P value	Characteristic	HR (95% CI)	P value
5-gene signature	19(4.5-81)	<0.001***	5-gene signature	19.94(2.71-146.5)	0.003**
Age	1(1-1)	0.204	Age	1.01(0.99-1.0)	0.173
Race			Race		
Asian	reference		Asian	reference	
Black	1.6(0.70-3.9)	0.256	Black	4.89(1.81-13.2)	0.002**
White	1.3(0.87-1.9)	0.205	White	1.32(0.74-2.4)	0.339
Gender			Gender		
Female	reference		Female	reference	
Male	0.81(0.56-1.2)	0.25	Male	0.81(0.49-1.3)	0.41
Vascular infiltration			Vascular infiltration		
NO	reference		NO	reference	
YES	1.4(0.94-2.2)	0.09	YES	1.19(0.67-2.1)	0.549
Stage			Stage		
I	reference		I	reference	
II	1.5(0.87-2.4)	0.151	II	1.02(0.50-2.1)	0.962
III	3.0(1.94-4.6)	<0.001***	III	2.07(1.13-3.8)	0.018*
IV	17.3(2.3-129.2)	0.006**	IV	28.09(3.31-238.4)	0.002**

* p<0.05; ** p<0.01; *** p<0.001

Figures

B PPI network showing the interactions of the pyroptosis-related genes, Hub gene: red.

C Consensus matrix heatmap to cluster patients into 2 to 5 clusters, showing the clustering stability after the 1000 times k-means cluster approach.

D Heatmap of the expression of PRGs in two molecular subtypes.

Figure 2

The different OS and immune profiles between the two molecular subtypes.

A Kaplan–Meier curves of overall survival between the two clusters.

B Discrepancies of immune cell infiltration between cluster-A and cluster-B

C Difference of expression of immune-related pathways between cluster-A and cluster-B

D Distributions of immune score, stromal score, EATIMATE score, and tumor purity between cluster-A and cluster-B. The distance of both ends of boxes represents the interquartile range of values and the thick line represents the median value.

Figure 3

Construction of a prognostic PRG model in HCC

A LASSO-Cox regression coefficient selection and variable screening.

B Cross-validation in the LASSO-Cox regression model to select the tuning parameter. Lambda. min = 0.06544434

C Kaplan–Meier curves of overall survival of patients in the high-risk score and low-risk score groups in the TCGA cohort.

D Distribution of risk score, survival status, and the expression of five prognostic PRGs in the TCGA cohort.

E ROC curves for predicting 1-, 3-, 5-year overall survival in the TCGA cohort.

- B Distribution of risk score, survival status, and the expression of five prognostic PRGs in ICGC cohort.
- C ROC curves for predicting 1-, 3-, 4-year overall survival in ICGC cohort.
- D Immune cell infiltration discrepancy between the high-risk score and low-risk score groups in ssGSEA.
- E Immune cell infiltration discrepancy between the high-risk score and low-risk score groups in TIMER.
- F Relationships between molecular subtypes of pyroptosis and risk score.

Figure 5

Comprehensive analyses of enriched pathways and genomic alterations between different risk groups.

A & B Gene Ontology enrichment analysis was performed with significantly upregulated and downregulated genes, respectively.

C & D Kyoto Encyclopedia of Genes and Genomes analysis was performed with significantly upregulated and downregulated genes, respectively.

E & F Top 20 most frequently mutated genes were illustrated in High-Risk Score Cohort and Low-Risk Score Cohort.

G TP53 occupies the top 1 position among differently mutated genes between High-Risk Score Cohort and Low-Risk Score Cohort.

H A lollipop plot showed the different mutation spots of TP53 between two cohorts.

I Kaplan-Meier analysis shows the independent relevance between overall survival and TP53 mutation.

J & K The heatmap illustrates the co-occurrence and mutually exclusive mutations of the top 25 frequently mutated genes in each cohort.

Figure 6

Development of the nomogram in the TCGA cohort.

A Nomogram for predicting the 1-year, 3-year, and 5-year OS of Hepatocellular carcinoma patients.

B ROC curves and AUC for 1-year, 3-year, and 5-year survival of the nomogram.

C KM curves of the OS of the nomogram.

D Calibration plot of the nomogram for predicting 1-year, 3-year, and 5-year OS.

Supplementary Files

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

- [FigureS1VolcanoMapofDifferentialGenes.pdf](#)
- [TableS1.pdf](#)
- [TableS2.pdf](#)
- [TableS3RiskHighdowngene.pdf](#)
- [TableS4RiskHighupgene.pdf](#)