A new nomogram model for prognosis of hepatocellular carcinoma based on novel gene signature that regulates cross-talk between immune and tumor cells

Youpeng Wang¹, Yeni Yang¹, Ziyin Zhao², Hongfa Sun¹, Dingan Luo¹, Lakshmi Huttad³, Bingyuan Zhang¹* and Bing Han¹*

Abstract

Background: The combined application of immune cells and specific biomarkers related to the tumor immune microenvironment has a better predictive value for the prognosis of HCC. The purpose of this study is to construct a new prognostic model based on immune-related genes that regulate cross-talk between immune and tumor cells to assess the prognosis and explore possible mechanisms.

Method: The immune cell abundance ratio of 424 cases in the TCGA-LIHC database is obtained through the CIBERSORT algorithm. The differential gene analysis and cox regression analysis is used to screen IRGs. In addition, the function of IRGs was preliminarily explored through the co-culture of M2 macrophages and HCC cell lines. The clinical validation, nomogram establishment and performing tumor microenvironment score were validated.

Results: We identified 4 immune cells and 9 hub genes related to the prognosis. Further, we identified S100A9, CD79B, TNFRSF11B as an IRGs signature, which is verified in the ICGC and GSE76427 database. Importantly, IRGs signature is closely related to the prognosis, tumor microenvironment score, clinical characteristics and immunotherapy, and nomogram combined with clinical characteristics is more conducive to clinical promotion. In addition, after co-culture with M2 macrophages, the migration capacity and cell pseudopod of MHCC97H increased significantly. And CD79B and TNFRSF11B were significantly down-regulated in MHCC97H, Huh7 and LM3, while S100A9 was up-regulated.

Conclusion: We constructed an IRGs signature and discussed possible mechanisms. The nomogram established based on IRGs can accurately predict the prognosis of HCC patients. These findings may provide a suitable therapeutic target for HCC.

Keywords: Hepatocellular carcinoma, Immune-related genes (IRGs), Tumor immune microenvironment(TIME), Cancer immunotherapy, Prognostic model

Introduction

Hepatocellular carcinoma (HCC) ranks third in the global cancer-related mortality rate [1], usually caused by chronic hepatitis and liver fibrosis [2, 3]. Surgical resection and liver transplantation are often used as the two main treatments for HCC. However, due to
the shortage of liver donors and the high recurrence rate of patients, the overall prognosis is not satisfactory [4]. In China, HCC is usually diagnosed at a late stage, which leads to the existing treatment methods with greater limitations and poor results. Less than 14.1% of patients live for up to 5 years [5]. Therefore, there is a need for early prediction of the survival status of patients, exploring new treatment methods to provide patients with personalized treatment and improving the clinical prognosis of patients.

Some studies have shown that immunotherapy has shown broad application prospects in treating many advanced cancers, especially for virus-induced cancers [6]. In China, most HCC patients are associated with HBV and suffer from chronic hepatitis. Meanwhile, the liver is considered to be an immune-tolerant organ. It can limit hypersensitivity to antigens and bacteria through the portal vein and can effectively receive allogeneic liver transplantation, creating an immunosuppressive microenvironment for the liver [7]. This shows that HCC patients may be more appropriate for immunotherapy.

The development of immunotherapy focuses on the tumor immune microenvironment (TIME) [8]. In addition to tumor cells, various immune cells, mesenchymal cells, secreted cytokines, chemokines and other non-tumor components that are also infiltrated in TIME have shaped different tumor heterogeneities [9, 10]. It has been reported that the HCC TIME has varieties of cytokines and is closely relevant to the prognosis of patients in many research, such as IL-6, IL-10, etc. [11, 12]. Therefore, we believe that the different cytokines and cell components in TIME have important guiding significance for the prognosis of patients. However, there is still a lack of immune-related genes that regulate the immune and tumor cells to effectively assess the heterogeneity of TIME and the prognosis of patients. Therefore, looking for key immune genes as HCC markers, clinicians can better understand the immunological characteristics of HCC and provide directions for patient prognosis and immunotherapy [13].

In this study, we downloaded the clinical survival information and RNA expression data of 424 cases in the Tumor Genome Atlas (TCGA-LIHC) database, and analyzed the content of 22 immune cells in the patients based on the CIBERSORT algorithm. Four immune cells related to survival in HCC were identified. And we also screened three IRGs that regulate the level of immune cell immersion. In the validation study, we chose M2 macrophages and three HCC cell lines as our cell models. We reported three essential IRGs that regulate the “cross-talk” between immune cells and tumor cells in TIME. Subsequently, we constructed a prognostic nomogram combining IRGs signature and clinical factors, which guides forecasting the prognosis of patients, and it may be a proper therapeutic target for HCC patients.

Materials and methods

Data source
From the Cancer Genome Atlas (TCGA) data portal, we downloaded the RNA-Seq gene expression profiles (FPKM and COUNT format) of 374 HCC and 50 adjacent normal HCC tissues, as well as clinical data on patient age, survival time, tumor staging, etc. In addition, RNA-Seq gene expression profiles and clinical information on 243 HCC specimens were validated from the ICGC database and 115 HCC specimens from the GEO database (GSE76427), respectively.

Identifying survival-related immune cells
The RNA-Seq (FPKM format) of 424 specimens were analyzed using the CIBERSORT algorithm and obtained a ratio matrix of 22 immune cells (perm = 100) [14, 15]. Owing to those samples with CIBERSORT P-value > 0.05 may represent samples with low immune cell infiltrate, they cannot be ignored. Therefore, we select 127 samples with CIBERSORT P-value < 0.1 and follow-up days ≥ 30 days for follow-up analysis [16]. Then, we analyzed the correlation in 22 immune cells in 127 patients. Finally, the Kaplan–Meier analysis for overall survival was used to identify survival-related immune cells, whose cut-off level was set at the median value according to the abundance ratio of 22 immune cells. Through using independent sample t-test and one-way ANOVA test, we analyzed the relationship between the abundance ratio of immune cells and tumor grade, clinical stage, and stage T.

Identifying Differentially expressed Immune-Related Genes (DEIRGs)
Cox proportional hazards regression was established based on the four survival-related immune cells identified in the previous steps. The 127 samples were sorted into high-risk (n = 64) and low-risk (n = 63) groups based on risk scores. We got 2498 unique immune-related genes from Immport database (https://www.immport.org/home), and we established the expression matrix of immune-related genes in 127 samples (count format) [17]. Through the edgeR R package for analysis of DEIRGs with the conditions: |logFC| > 1 and P < 0.05 [18].

Protein–Protein Interaction Network Construction and Hub Genes Screening
The 412 differential IRGs were analyzed in the STRING database (https://stringdb.org/), with the condition: combined-score ≥ 0.4 [19]. This network was visualized using Cytoscape 3.8.2 and analyzed by the MCODE plugin. Ultimately, we obtained 11 modules and 10 seed genes.
At the same time, Cytohubba was used to screen the top 20 nodes ranked by degree. We selected 30 genes as immune microenvironment-related hub genes.

**Relationship between clinical characteristics and hub genes**

For our study, 127 patients were grouped and using Kaplan–Meier survival analysis, and the overall survival rate was analyzed according to the expression level of the 30 hub genes. Here, we identified 9 survival-related hub genes. We analyzed and visualized the hub genes’ connection with clinical characteristics by the "WGCNA" R package.

**Construction of the IRGs signatures**

To develop a prognostic model, 9 survival-related genes in Kaplan–Meier survival analysis were included in multivariate proportional hazards regression analysis. The 127 patients were sorted according to their risk score, which was derived from gene expression multiplied by a linear combination of regression coefficients obtained from the multivariate Cox regression. The 63 patients with the low-risk score were defined as the low-risk group, and the remaining 64 patients were in the high-risk group. Using the Kaplan–Meier analysis to compare OS between the two groups of patients and the "survival ROC" package to plot receiver operating characteristic (ROC) curve.

**External validation of the IRGs**

243 HCC specimens in the ICGC database and 115 HCC specimens in the GSE76427 were used as a verification cohort to verify the prognostic accuracy of the IRGs signature risk score model. The samples were divided into high-risk and low-risk groups by calculating risk scores based on the same formula, and their Kaplan–Meier and ROC curve were analyzed, respectively.

**Enrichment analysis of differentially expressed genes (DEGs) between low-risk and high-risk groups**

Through the edgeR R package for analysis of DEGs between low risk ($n = 63$) and high risk ($n = 64$) groups in TCGA with the conditions: $|\log FC| > 1$ and $P < 0.05$. Analyzing DEIRGs in the GO (Gene Ontology) and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways via the DAVID 6.8 (https://david.ncifcrf.gov/) [20, 21]. The GO terms and KEGG signaling pathways are then visualized via R Package "ggplot2" with the conditions: FDR < 0.05 and counts $\geq 4$.

**Analysis of the degree of immune infiltration between low-risk and high-risk groups**

The ssGSEA was executed to probe into the different infiltration degrees of immune cell types, immune-related functions and pathways in the expression profile of low-risk and high-risk groups using the R package "GSVA" based 29 immune-related gene sets [22]. To prove the effectiveness of IRGs risk scores and to picture clustering heatmap, we made use of R package "ESTIMATE" to study the expression level of RNA-seq to count the tumor purity, estimate score (ES), immune score (IS), and stromal score (SS). Using the R package "ggpubr", we obtained the vioplots of ES, IS, and SS in low-risk and high-risk groups. The correlation of immune cells with IRGs signature risk score was analyzed and visualized by the "corrplot" package in R.

**Construction of prognostic nomogram**

To provide a quantitative analysis tool to predict the survival risk of HCC patients, we were further constructed the nomogram on the basis of IRGs as well as clinical parameters in 127 patients in TCGA. In order to evaluate the accuracy of the nomogram, we used 115 patients in GSE76427 for external validation, and the calibration curve and DCA curves are drawn with the R-pack "rms" and "ggDCA".

**Cell line culture**

MHCC-97H, Huh7, LM3 HCC cell lines and THP-1 cell line were purchased from the Shanghai cell bank (Chinese Academy of Sciences, Shanghai, China). HCC cell lines were cultured in a medium with 10% FBS and 1% P/S (Gibco, Thermo Fisher Scientific, Waltham, USA), and THP-1 cells were cultured in RPMI 1640 medium (Hyclone, Thermo Fisher Scientific, Waltham, USA) with 10% FBS. All cells were maintained in a humidified atmosphere with 5% CO2 at 37 °C.

**THP-1-derived M2 macrophages and Establishment of co-culture system**

THP-1 cells were treated with phorbol 12-myristate 13-acetate (PMA) (Sigma, Saint-Quentin Fallavier, France, 100 ng/mL) for 24 h to polarize THP-1 cells into macrophages. The IL-4 and IL-13 were then polarized into M2 macrophages (Sino Biological Als, China, 5 μg). MHCC-97H cells ($1 \times 10^6$ cells) were placed in the lower chamber of a 6-well transwell plate. After 24 h, M2 macrophages ($1 \times 10^6$ cells) derived from THP-1 were placed on the 0.4-μm porous membrane in the upper chamber to establish a co-culture system in a serum-free DMEM medium [23]. Then 48 h later,
MHCC-97H cells were collected for RNA extraction and other experiments.

**Cell migration assay**
Using transwell compartments (8 um pore) to assess the cell migration capacity (Corning, 353,097). Normal MHCC-97H cells and M2 macrophages co-culture treated MHCC-97H cells (5 × 10⁴ cells) were suspended in serum-free medium in the upper compartment of a 24-well transwell plate, while medium with 30% FBS is placed in the lower chamber. After 24 h at 37 °C, the translocated cells were stained with 0.5% crystal violet for 20 min.

**Quantitative real-time PCR (qPCR)**
Total RNA was extracted with RNA-easy isolation reagent (Vazyme, Nanjing, China). PrimeScript™ RT Kit (TaKaRa, RR047A) and SYBR Premix EX Taq™ (TaKaRa, RR820A) were used to cDNA synthesis and qPCR on the FTC-3000P real-time PCR system (Funglyn Biotech, Shanghai, China). Supplementary Table 1 shows the PCR primers used.

**Statistical analysis**
Using IBM SPSS Statistics performed all analyses (version 23). A $P<0.05$ indicated statistical significance.

**Results**

**Identifying survival-related immune cells**
Using the CIBERSORT algorithm to analyze the abundance ratio of 22 immune cells in 127 samples, revealing the relative content of 22 immune cells in normal and tumor samples (Fig. 1A, 1B). As shown in Fig. 1A, M0/M1/M2 Macrophages, CD8 + T cells, and dendritic cells occupy a large proportion in the sample. In adjacent normal HCC tissues, the content of active mast cells, M0 macrophages and Tregs were significantly higher than that of tumor samples ($P<0.05$), while the content of M2 Macrophages, plasma cells and monocytes were significantly lower than that of tumor samples ($P<0.05$) (Fig. 1B). Correlation analysis further suggests that there are connections between 22 immune cells (Fig. 1C). CD8 + T cells are positively correlated with the content of T cells follicular helper, active CD4 + T memory cells, and Plasma cells, but negatively correlated with the content of resting CD4 + T memory cells, M0 macrophages, and M2 macrophages. In addition, Fig. 1D-G show that the abundance ratios of the four types of immune cells are related to survival rates by Kaplan–Meier analysis, among CD8 + T cells ($P=0.006$), Plasma cells ($P=0.01$), and CD4 + memory resting T cells ($P=0.05$) are indicators of favorable prognosis, while M2 Macrophages ($P=0.05$) are indicators of unfavorable prognosis. The correlation between abundance ratios of the four immune cells and clinical characteristics reveals that CD8 + T cells, Plasma cells, and resting CD4 + T memory cells decreased with the increase of stage T, clinical stage, and tumor grade, while M2 Macrophages is the opposite (Supplementary Fig. 1).

**Identifying immune-related genes and enrichment analysis**
Cox proportional hazards regression was established based on the four survival-related immune cells. Risk scores = Plasma cells*( -7.76) + CD8 + T cells *( -3.26) + resting CD4 + T memory cells *( -4.42) + M2 Macrophages * 1.08. According to the risk score, the samples were divided into high-risk and low-risk groups. We analyzed the immune-related genes related to the risk score level and obtained 412 immune-related differential genes (Supplementary Fig. 2A). Using the DAVID website, GO/KEGG enrichment performed the analysis of 412 immune-related differential genes. The supplementary Fig. 2B-E shows the top 12 enrichment results. The results showed that the differential genes were mainly located in T cell receptor complex and extracellular exosome, significantly involved in complement activation, inflammatory response, antigen binding, transmembrane signaling receptor activity, and were mainly enriched in the chemokine signaling pathway, natural killer cell-mediated cytotoxicity, Jak-STAT signaling pathway. In conclusion, 412 immune-related gene proteins are mainly involved in various signaling pathways and the regulation of immune responses, cell proliferation and apoptosis, closely connecting various immune cells, stromal cells and tumor cells in the tumor microenvironment.

**Protein–protein interaction network construction and hub genes screening**
To probe into the interrelationship between immune-related genes and get hub genes, we performed PPI and module analysis to obtain 30 hub genes. Supplement Table 2 shows the full names and primary functions of 30 hub genes, meanwhile Cytoscape analysis was used to get the first two most important modules (Fig. 2A-B). The functional analysis of genes involved in this module was analyzed using DAVID. Module 1 is mainly related to HIV and lung cancer. It is primarily concentrated in immune cell activation and chemotaxis, and cell proliferation (Table 1). Module 2 is mainly related to HCC, HBV infection, lung cancer and is primarily enriched in the proliferation and differentiation of immune cells and apoptosis (Table 1). Both are closely related to cancer and immune signaling pathways, such as the chemokine signaling pathway, Jak-STAT signaling pathway, TNF signaling pathway.
Fig. 1 The relationship between the abundance ratios of immune cells and overall survival. A Differences in the expression of 22 immune cells in adjacent normal HCC and HCC tissue. B The abundance ratio of immune cells in the 127 samples. C The relationship between the abundance ratios of various immune cells. D-G The survival analysis for the abundance ratios of the four immune cells.
Fig. 2 The top two modules and survival analyses of the hub genes. A Two modules in MCODE. Redder indicates that the higher the number of interactions with other proteins, the smaller the greener the number. B-J The nine genes are significantly related to survival. Red represents high gene expression, and blue represents low gene expression.
Relationship between clinical characteristics and hub genes

Through Kaplan–Meier survival, we analyzed 30 hub genes, and obtained 9 immune-related genes with prognostic significance \((P<0.05)\). Including CCL5, CCR7, CD79B, CD247, CXCL1, CXCL5, CXCR3, LTBP1, S100A9, TNFRSF11B (Fig. 2B–J). Table 2 shows the correlation analysis between these nine prognostic-related hub genes and their clinical characteristics. CCR7 and CD79B have a significant positive correlation with stage (I/II/III/IV) and stage T, and LTBP1 has a significant negative correlation with stage.

Establishment and verification of IRGs signature

9 hub genes were tested for their prognostic significance to perform univariate COX analysis and included hub genes with \(P<0.1\) into the multivariate COX analysis (Table 3). To get the best model, these 9 genes were analyzed using the Cox proportional hazards model method of R package "survival". Finally, 3 immune-related genes were used to construct Cox proportional hazards model as follows: Riskscore = CD79B*(0.000946) + TNFRSF11B*0.000158 + S100A9*0.00025.
Further, the risk scores of 127 HCC patients are sorted and divided into high-risk (n = 64) and low-risk (n = 63) groups according to the median risk score. Through K-M analysis, the higher the risk score of patients, the worse the prognosis (P < 0.001, Fig. 3A). We plotted not only the ROC curves of our IRGs signature, but also the ROC curves of other previously published IRGs signatures (Liu's signature [13] and Dai's signature [24]), which have been published previously (Fig. 3B-D). Interestingly, the 1-year, 3-year and 5-year AUC of our IRGs signature was significantly higher than Liu's signature and Dai's signature in the training cohort. These prove the excellent value of the IRGs we constructed in predicting the prognosis of HCC patients.

**External validation of the IRGs signature**

External verification was performed on 243 HCC patients in the ICGC database and 115 HCC patients in the GSE76427. Sorted based on risk scores, patients were divided into high-risk (n = 121 in ICGC, n = 57 in GSE76427) and low-risk groups (n = 122 in ICGC, n = 58 in GSE76427) with the median risk score as the cut-off value. K-M analysis showed that patients in high-risk groups had a worse prognosis. (Fig. 3E, G). The ROC curve analysis was performed, and the results showed that the AUC values of 1, 3, and 5 years were 0.663, 0.608, and 0.635 respectively, and 0.609, 0.703 and 0.575 in the GSE76427 respectively (Fig. 3F, H). This proves that the IRGs signature has a strong predictive ability.

**Identifying DEGs between high-risk and low-risk groups and Enrichment Analysis**

Through differential expression analysis between low-risk and high-risk groups, we obtained 2371 DEGs (Fig. 4A). Interestingly, GO analysis results showed that the changes in the biological process (Fig. 4B) were significantly enriched in immune response, particularly in the B cell receptor signaling pathway, T cell co-stimulation and granulocyte–macrophage colony-stimulating factor production. The cellular component which has changes (Fig. 4C) was mainly enriched in the extracellular region, plasma membrane, extracellular exosome. KEGG pathway (Fig. 4D) was mainly enriched in natural killer cell-mediated cytotoxicity, T cell receptor signaling pathway, Jak-STAT signaling pathway. These results suggest that our IRGs signature regulates the immune response primarily by participating in stimulation or activation of T cell and B cell, as well as the polarization of macrophages.

**The degree of immune infiltration between low-risk and high-risk groups**

The RNA sequencing data of HCC samples were analyzed by the ssGSEA method and got the abundance levels of 29 immune-related cells and types in 127 HCC samples. Patients in the low-risk group had a higher degree of immune cell infiltration (Fig. 5A). Meanwhile, we calculate the stromal score (SS), immune score (IS) and estimate score (ES) by using the ESTIMATE algorithm. The result has shown that the SS, IS, and ES of the low-risk group was significantly higher than that of the high-risk group (P < 0.05) (Fig. 5A-B). Immune cell correlation analysis showed that as the risk score increased, plasma cells and CD8+ T cells gradually decreased, and conversely M2 cells and M0 cells gradually increased (p < 0.05) (Fig. 5C-F).

**Establishment of a risk-nomogram for predicting survival in HCC patients**

Combined with other univariate and multivariate COX analyses of significant and important clinical features (P < 0.1), such as AFP level, hepatitis, tumor status and other factors that help in the disease detection, we have established a convenient and clinically adaptable risk nomogram to predict the survival probability of HCC patients in the training cohort of TCGA and the validation cohort of GSE76427. By calculating the sum of the scores corresponding to the corresponding levels of each factor in the nomogram, the higher the total score, the worse the patient's 1-year, 3-year, and 5-year OS rate (Fig. 6A, C).

To verify the validity of this nomogram, we calculated the C-statistic discriminatory index and the calibration plot of the prediction models in the training cohort and the validation cohort. Figure 6D shows the calibration plot for predicting the 1-year, 3-year and 5-year OS rates, with an accompanying C-statistic discriminatory index value of 0.739 and 0.668, demonstrating the good predictive ability and effectiveness of our nomograms. This proves that the IRGs signature combines clinical characteristics such as AFP levels and hepatitis to further enhance clinical value and predictive power (Table 4). In addition, according to the DCA curve, the nomogram showed exceptional performance in the training cohort (Fig. 6B).
Fig. 3 (See legend on previous page.)
Effects of M2 macrophage levels on HCC cell migration
We used PMA, IL-4 and IL-13 to induce M2 macrophage formation (Fig. 7A). First, the M0 macrophage markers CD11b and CD14 were detected by qPCR. Then, the expression of M2 macrophage marker CD206 and CD163 was significantly up-regulated compared to M0 macrophages by qPCR ($P < 0.05$) (Fig. 7D). Figure 7B shows morphological changes in THP-1, M0 and M2 cells. We then probe into whether M2 macrophages affected HCC cell migration levels. After co-culture with M2 macrophages, we observed a significant increase in the ability of MHCC-97H to migrate, and the pseudopod of MHCC-97H cells has increased significantly in morphology (Fig. 7C). It is proved that M2 macrophages increase the malignancy of HCC.

IRGs identified as key factors of tumor microenvironment in HCC
To determine the essential value of CD79B, S100A9, TNARSF11B in the immune infiltration of HCC. Compared with three normal HCC cell lines, in HCC cell lines co-cultured with M2 macrophages, the expression of CD79B and TNFRSF11B was significantly down-regulated, while the expression of S100A9 was up-regulated considerably (Fig. 7E-G).

Discussion
The liver receives blood from the hepatic artery and portal vein and has a rich blood supply [25]. Therefore, the liver will be exposed to intestinal-derived microorganisms and food-derived harmless antigens at a high level for a long time, giving the liver a unique immune privilege.
Fig. 5  The degree of immune infiltration between low-risk and high-risk groups. A Differences in 29 immune-related gene sets between the high-risk and low-risk groups. Using ESTIMATE algorithm to calculate the SS, IS, ES. B The violin plots of SS, IS, ES. C-F Spearman’s correlation analysis of immune cells and IRGs signature risk scores.
Despite a major histocompatibility complex (MHC) mismatch, the liver can accept allogeneic transplantation by inducing immune tolerance [26, 27]. However, the survival rate and activity of hepatitis viruses, bacteria and tumors in the immunosuppressive microenvironment are higher, leading to persistent infections and rapid cancer progression. Various components of the immune microenvironment regulate the occurrence and progress
of HCC [28–30]. In recent years, immunotherapy has made great breakthroughs in tumor treatment, including HCC [31, 32]. Considering the close connection between the immune microenvironment in tumorigenesis and immunotherapy [33]. This article hopes to provide guidance for the prediction of prognosis of HCC patients and
probe into potential immunotherapy targets for HCC by screening immune cell and gene targets closely related to immune infiltration and clinical characteristics.

The interaction of various components in TIME leads to complex functions of the body. Consequently, understanding the relationship between immune infiltration and patient prognosis is an urgent goal. This study first explored the relationship between M2 macrophages, plasma cells, resting CD4+ T memory cells, CD8+ T cells immune cells and the survival of HCC patients. CD4+ T cells and CD8+ T cells can activate the immune system to kill tumor cells, a research hotspot in tumor immunotherapy [34, 35]. CD8+ T cells can differentiate into effector cytotoxic T lymphocytes (CTLs), which have two main ways to kill tumor cells: granular exocytosis and Fas ligand (FasL)-mediated apoptosis induction [36] (Fig. 8). Most effector cells can return to a resting state and form memory cells [37]. CD4+ T memory cells are essential for adaptive immunity [38] and can be divided into central memory T cell (TCM) and effector memory T cell (TEM) [39]. When re-infected, TEM can release IFN-γ, IL-4 and other cytokines and chemokine receptors in a short period to quickly perform immune functions. TCM can maintain immune memory while expressing IL-2 and dividing rapidly. In the peripheral blood of patients with advanced cancer, the proportion of TCM cells decreases, showing a state of T cell exhaustion [40]. Our research suggests that patients with low levels of resting CD4+ T memory cells and CD8+ T cells have a shorter overall survival period, consistent with the theory of T cell exhaustion. Following antigen exposure and T cell licensing, B cells can secrete IL-2, IL-4, IFN-γ, TNF-α and other cytokines to enhance T cell toxicity [41]. It usually differentiate into potent antibody-secreting plasma cells, which are essential for humoral immunity [42] (Fig. 8). A large amount of evidence indicates that the high expression of plasma cell-related genes is related to the excellent prognosis of various tumors [43, 44]. Tumor-associated macrophages (TAMs) are macrophages that infiltrate around tumor cells. They are

Fig. 8 schematic overview. Schematic overview of functional interactions between B cells, PC, T cells and macrophages in the tumor microenvironment. B cells can enhance T cell responses by producing IL-2, IL-4 and other cytokines. Meanwhile, it can be differentiated into FC, which produces antibodies against tumor-associated antigens and triggers antibody-dependent cellular cytotoxicity (ADCC) responses. It also supports tumor-associated macrophages (TAMs) to take up tumor antigens and polarize M1 type to exert phagocytosis. Tumor induces the formation of M2 macrophages, which can increase the expression of tumor cells 5100A9, further enhance M2 polarization and recruit M2 macrophages. It can also reduce the expression of TNFRSF11B and enhance the migration ability of tumors. Meanwhile, it can reduce the expression of CD79B, a component of the BCR complex, and weaken the immune function of B cells
mainly divided into classically activated macrophages (M1) and alternately activated macrophages (M2). The overall appearance of the M2 macrophage phenotype promotes immune evasion of tumor cells [45]. The mutual transformation of M1 and M2 macrophages regulates tumorigenesis [46]. Evidence shows that TAMs play an essential role in the development, invasion, drug resistance, immune escape and angiogenesis of HCC [47]. Consistent with previous studies, our study shows that the enrichment of M2 macrophages in HCC indicates a poor prognosis, and the enrichment of plasma cells is an indicator of a favorable prognosis. Overall, the four survival-related immune cells identified in this study are significant in immune infiltration and immunotherapy of HCC, confirming the significance of related gene analysis based on immune cells.

We have identified 30 hub genes, nine of which are related to patient prognosis. Finally, with univariate and multivariate COX regression, we screened 3 Hub genes as an IRGs signature for HCC patients, including S100A9, CD79B, and TNFRSF11B. Previous studies have linked these three genes to tumor development, but the regulatory effect of the HCC immune microenvironment is still unclear. CD79B, as one of the main components of the B-cell antigen receptor complex (BCR), is mainly responsible for transducing antigen-recognition signals to the inside of the cell, expressed on almost all B cell surfaces [48, 49]. BCR signaling primarily affects key functions such as immune synapse formation, antigen affinity, cell migration, etc., and is critical for the maturation and maintenance of B cells [50]. The downregulation of CD79B in the tumor immune microenvironment may lead to abnormal BCR signaling, affecting cell activity and leading to immunosuppression. S100A9 is a secretory protein in the inflammatory microenvironment that is significantly up-regulated in TAMs and is mainly expressed in neutrophils and circulating monocytes [51, 52] and is related to poor differentiation of HCC, vascular infiltration, invasion and metastasis [53]. Interestingly, M2 macrophages have been reported to secrete more S100A9, enhancing the stem cell-like properties of HCC cells through the AGE/NF-κB axis signaling pathway [54]. TNFRSF11B is able to bind and inhibit TRAIL (TNF-related apoptosis-inducing ligand) to exert anti-apoptotic effects, suggesting that it may provide a survival advantage for cells [55]. It is also controversial in the study of HCC [56, 57]. Studies have shown that in highly aggressive liver cancer cells, the expression of TNFRSF11B is often lower than that of low aggressive liver cancer cells [58]. The same point of view is that the down-regulation of TNFRSF11B expression can promote HCC bone metastasis in vivo [59]. However, there is also evidence that the high expression of TNFRSF11B is an important reason for the high metastatic potential of HCC cells [60]. Hence, we need to study the mechanisms of the three genes further to help develop new HCC treatment strategies.

We have established a new IRGs signature with S100A9, CD79B, and TNFRSF11B. Compared to other previously published IRGs signatures (Liu’s signature [13] and Dai’s signature [24]), our IRGs signature has higher accuracy, and our signature has been successfully verified in the GSE76427 and ICGC databases (Fig. 3). To explore the regulatory role of our IRGs signature, the 127 patients in TCGA were sorted according to the risk score and divided into high-risk(n = 64) and low-risk(n = 63) groups. Interestingly, GO analysis results the DEGs were significantly enriched in the B cell receptor signaling pathway and T cell co-stimulation, particularly in granulocyte–macrophage colony-stimulating factor production (Fig. 4). And our IRGs signature risk score and M0/M2 macrophages are significantly positively correlated (Fig. 5). This evidence suggests that our IRGs signature is closely related to macrophages.

In addition, immune cells and stromal cells are the two major non-tumor components in the tumor microenvironment. Identifying the right tumor immune subtypes can accurately guide the clinical treatment and prognosis of tumors. We got the abundance levels of 29 immune-related cells between low-risk and high-risk groups (Fig. 5), and found that patients in the low-risk group had a higher degree of immune infiltration. Using ESTIMATE algorithm, we generated immune scores, stromal scores and purity of tumor cells for 127 HCC patients, which have been used to evaluate the immunological characteristics of HCC [61]. Compared to the low-risk group, the immune scores and stromal scores were lower in high-risk group. This suggests that the proportion of immune cells in the tumor microenvironment and tumor purity have an essential impact on the prognosis of HCC. This implies that our IRGs signature has important guiding implications for immunoclassification in HCC patients. For patients with higher risk scores, different immunotherapies combined with conventional treatments should be given priority, to improve the prognosis. Meanwhile, we combined a few selected clinicopathological characteristics, such as BMI, AFP, tumor status, stage, etc., to establish a predictive nomogram model to evaluate the 1-year, 3-year, and 5-year prognosis of HCC patients to achieve accurate prediction of survival. The calibration plots show that the model has a high prediction accuracy. This indicates that the new scoring system established in this study is of great help to the patients who need additional treatment or enhanced follow-up, and can better promote accurate prevention and personalized health management of patients.
We found a significant positive correlation between IRGs signature risk scores and M0/M2 macrophages (Fig. 5E, F). To further explore the functions of the IRGs signature genes associated with M2 macrophages, we first co-cultured three HCC cell lines with M2 macrophages, respectively. The results showed that the migration capacity of MHCC-97H was significantly enhanced, and the pseudopod of MHCC-97H cells has increased significantly in morphology after co-culture with M2 macrophages (Fig. 7C). This proves that M2 macrophages cause changes in gene expression in HCC cells, leading to an increase in the malignancy of HCC. M2 macrophages may have released some macromolecular substances through exosomes that regulate the expression of IRGs signature genes in the microenvironment, such as miR-200c, miR-203 [62]. Considering that IRGs signature genes play a key regulatory role in co-culture systems, we detected changes in the expression of IRGs signature genes in MHCC97H, Huh7 and LM3 cell lines through qPCR. The results showed that in the tumor microenvironment created by M2 macrophages, the expression of CD79B in tumor cells was down-regulated, which may lead to abnormal BCR signaling. Consistent with previous studies [54], the expression of S100A9 in tumor cells was significantly increased after M2 macrophages were co-cultured, and the expression of TNFRSF11B was significantly reduced, which may increase the malignancy of tumor cells (Fig. 8). This implies that we constructed an IRGs signature that may regulate the "cross-talk" between immune cells and tumor cells.

Overall, we identified 4 immune cells related to survival and 9 immune-related genes related to the prognosis of HCC patients. These cells and genes can be considered as biomarkers for judging the prognosis of HCC, and may serve as important targets for HCC immunotherapy. Highlighting the outcomes of our study, firstly, we screened 3 hub genes to establish an IRGs signature, and verified and evaluated the signature in the multiple data sets, demonstrating the reliability of the signature. Secondly, we have conducted a comprehensive and in-depth study on the correlation between IRGs signature and immune cells and immune checkpoints, which provides a potential direction for the research of HCC immunotherapy. And finally, we established a quantitatively calculated nomogram, which is more conducive to clinical application. However, the limitation of this study is that the public data set lacks some key clinical pathology and neoadjuvant therapy related information of patients. Meanwhile, markers based on immune-related genes need to be applied in clinical patients to verify the predictability of their prognosis.

Abbreviations
HCC: Hepatocellular carcinoma; OS: Overall survival; DEGs: Differential expression genes; TCGA: The cancer genome atlas; ROC: Receiver operating curve; AUC: Area under the curve; IRGs: Immune-related genes; TIME: Tumor immune microenvironment; GO: Gene Ontology; AFP: Alpha-fetoprotein; MHC: Major histocompatibility complex; CTLs: Cytotoxic T lymphocytes; TCM: Memory T cell; TEM: Effector memory T cell; TAMs: Tumor-associated macrophages.

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s12885-022-09465-9.

Additional file 1: Figure 1. The relationship between the abundance ratios of the immune cells and clinical characteristics. (A–C) The relationship between the abundance ratios of each immune cell and stage T, stage and tumor grade. Figure 2. (A) Volcano plots of the hepatocellular carcinoma gene expression profiles grouping by the risk score. Red/blue symbols classify the upregulated/downregulated genes according to the criteria: |log2FC| > 1.5 and P-value < 0.05. (B–E) Represent the enrichment analysis results of genes involved in immune cell infiltration, namely biological processes, cellular components, molecular functions, and KEGG. The main 12 results of each term are shown.

Additional file 2: Table 1. qPCR primer sequence.

Additional file 3: Table 2. Function roles of the 30 hub genes.

Acknowledgements
Not applicable.

Authors’ contributions
All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

Funding
Natural Science Foundation of Shandong Province (CN) (ZR201911030198).

Availability of data and materials
The datasets analysed during the current study are available in TCGA, ICGC and GSE76427. All data and outcomes generated during this study are included in this published article and its supplementary information files.

Declarations
Ethics approval and consent to participate
The study was approved by the Affiliated Hospital of Qingdao University. All methods were carried out in accordance with relevant guidelines and regulations.

Consent for publication
This is a bioinformatics study based on data retrieved from online databases. Since all of the datasets were retrieved from published literature, it was confirmed that all written informed consent were obtained.

Conflict of Interest
The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author details
1 Department of Hepatobiliary and Pancreatic Surgery, The Affiliated Hospital of Qingdao University, 16 Jiangsu Street, Qingdao 266005, China. 2 Organ Transplantation Center, The Affiliated Hospital of Qingdao University, Qingdao, China.
18. Varet H, Brillet-Gueguen L, Coppee JY, Dillies MA. SARTools: a DESeq2- and EdgeR-Based R pipeline for comprehensive differential analysis of RNA-Seq Data. PLoS One. 2016;11(6):e0157022. https://doi.org/10.1371/journal.pone.0157022.

19. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P, Jensen LJ, von Mering C. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res. 2017;45(D1):D362–8. https://doi.org/10.1093/nar/gkw937.

20. Kaneshia M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res. 2016;44(D1):D457–62. https://doi.org/10.1093/nar/gkv1070.

21. Kaneshia M, Goto S. KEGG kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28(1):27–30. https://doi.org/10.1093/nar/28.1.27.

22. Shen Y, Peng X, Shen C. Identification and validation of immune-related IncRNA prognostic signature for breast cancer. Genomics. 2020;112(3):2640–6. https://doi.org/10.1016/j.ygeno.2020.02.015.

23. Ning Y, Cui Y, Li X, Cao X, Chen A, Xu C, Cao J, Luo X. Co-culture of ovarian cancer stem-like cells with macrophages induced SKOV3 cells stemness via IL-8/STAT3 signaling. Biomed Pharmacother. 2018;103:262–71. https://doi.org/10.1016/j.biopha.2018.04.022.

24. Dai Y, Qiang W, Lin K, Gui Y, Lan X, Wang D. An immune-related gene signature for predicting survival and immunotherapy efficacy in hepato-cellular carcinoma. Cancer Immunol Immunother. 2021;70(4):967–79. https://doi.org/10.1007/s00262-020-02743-0.

25. Gang S, Kumar KH, Sahni D, Yadav TD, Aggarwal A, Gupta T. Anatomy of the hepatic arteries and their extrahepatic branches in the human liver: a cadaveric study. Ann Anat. 2020;227:151409. https://doi.org/10.1016/j.aanat.2019.07.010.

26. Jiang Y, Que W, Zhu P, Li XK. The role of diverse liver cells in liver transplantation tolerance. Front Immunol. 2020;11:1203. https://doi.org/10.3389/fimmu.2020.011203.

27. Du X, Chang S, Guo W, Zhang S, Chen ZX. Progress in liver transplant tolerance and tolerance-inducing cellular therapies. Front Immunol. 2020;11:1326. https://doi.org/10.3389/fimmu.2020.01326.

28. Wu Q, Zhou L, Lv D, Zhu X, Tang H. Exosomes-mediated communication in the tumor microenvironment contributes to hepatocellular carcinoma development and progression. J Hematol Oncol. 2019;12(1):153. https://doi.org/10.1186/s13045-019-0739-0.

29. Zhang Q, Lou Y, Bai XL, Liang TB. Immunometabolism: a novel perspective of liver cancer microenvironment and its influence on tumor progression. World J Gastroenterol. 2018;24(31):3500–12. https://doi.org/10.3748/wjg.v24.i31.3500.

30. Tahmasebi Birgani M, Carloni V. Tumor Microenvironment, a Paradigm in Hepatocellular Carcinoma Progression and Therapy. Int J Mol Sci. 2017;18(2). https://doi.org/10.3390/ijms18020405.

31. Banerjee K, Kumar S, Ross KA, Gautam S, Poelaret B, Nasser MW, Atibal A, Bhata A, Wannemuehler MJ, Narasimhan B, Solheim JC, Batra SK, Jain M. Emerging trends in the immunotherapy of pancreatic cancer. Cancer Lett. 2018;417:35–46. https://doi.org/10.1016/j.canlet.2017.12.012.

32. Jiang Y, Han QJ, Zhang J. Hepatocellular carcinoma: Mechanisms of progression and immunotherapy. World J Gastroenterol. 2019;25(25):3151–67. https://doi.org/10.3748/wjg.v25.i25.3151.

33. Wang H, Guo J, Jiang X, Wang Z. Less immune cell infiltration and worse prognosis after immunotherapy for patients with lung adenocarcinoma who harbored STK11 mutation. Int Immunopharmacol. 2020;94:106574. https://doi.org/10.1016/j.intimp.2020.106574.

34. Zander R, Schauder D, Xin G, Nguyen C, Wu X, Zajac A, Cui W. CD4(+) T cell help is required for the formation of a cytolytic CD8(+) T cell subset that protects against chronic infection and cancer. Immunity. 2019;51(6):1028-1042 e4. https://doi.org/10.1016/j.immuni.2019.10.009.

35. Jiang X, Xu J, Liu M, Xing H, Wang Z, Huang L, Mellor AL, Wang W, Wu S. Adoptive CD8(+) T cell therapy against cancer: challenges and opportunities. Cancer Lett. 2019;462:23–32. https://doi.org/10.1016/j.canlet.2019.07.017.

36. Farhood B, Najafi M, Morretazee K. CD8(+) cytotoxic T lymphocytes in cancer immunotherapy: a review. J Cell Physiol. 2019;234(6):8509–21. https://doi.org/10.1002/jcp.27782.

37. Gouverguil FB, Ellis JS, Wan X, Dhakal M, Hoeman CM, Cascio JA, Zaghouani H. Antigen-free adjuvant assists late effector CD4 T cells to transit to memory in lymphopenic hosts. J Immunol. 2013;191(3):1126–35. https://doi.org/10.4049/jimmunol.1202262.
38. Swain SL, Agrewnala JN, Brown DM, Jelley-Gibbs DM, Golech S, Huston G, Jones SC, Kamiperschoer C, Lee WH, Mckinstry KK, Roman E, Strutt T, Weng NP. CD4+ T-cell memory: generation and multi-faceted roles for CD4+ T cells in protective immunity to influenza. Immunol. Rev. 2006;211:18-22. https://doi.org/10.1111/j.0105-2896.2006.00388.x.

39. Schenkel JM, Masopust D. Tissue-resident memory T cells. Immunity. 2014;41(6):866-89. https://doi.org/10.1016/j.immuni.2014.12.007.

40. Klebanoff CA, Gattinoni L, Restifo NP. CD8+ T-cell memory in tumor immunology and immunotherapy. Immunol. Rev. 2006;211:214-24. https://doi.org/10.1111/j.0105-2896.2006.00389.x.

41. Sharonov GV, Serebrovskaya EO, Yuzhakova DV, Britanova OV, Chudakov DM. B cells, plasma cells and antibody repertoires in the tumour microenvironment. Nat Rev Immunol. 2020;20(5):294-307. https://doi.org/10.1038/s41577-019-0257-x.

42. Schwartz M, Zhang Y, Rosenblatt JD. B cell regulation of the anti-tumor response and role in carcinogenesis. J Immunother Cancer. 2016;4:40. https://doi.org/10.1186/s40425-016-0145-x.

43. Alistar A, Chou JW, Nagalla S, Black MA, D’Agostino RJ, Miller LD. Dual roles for immune metagenes in breast cancer prognosis and therapy prediction. Genome Med. 2014;6(10):80. https://doi.org/10.1186/s13073-014-0080-8.

44. Wouters MCA, Nelson BH. Prognostic significance of tumor-infiltrating B cells and plasma cells in human cancer. Clin Cancer Res. 2018;24(24):6125-35. https://doi.org/10.1158/1078-0432.CCR-18-1481.

45. Mosser DM, Edwards JP. Exploring the full spectrum of macrophage activation. Nat Rev Immunol. 2008;8(12):958-69. https://doi.org/10.1038/nri2448.

46. Xiang W, Shi R, Kang X, Zhang X, Chen P, Zhang L, Hou A, Wang R, Zhao Y, Zhao K, Liu Y, Ma Y, Luo H, Shang S, Zhang J, He F, Yu S, Gan L, Shi C, Li Y, Yang W, Liang H, Miao H. Monocarboxylic acid lipase regulates can-nabinoid receptor 2-dependent macrophage activation and cancer progression. Nat Commun. 2018;9(1):2574. https://doi.org/10.1038/s41467-018-04999-8.

47. Fu XT, Song K, Zhou J, Shi YH, Liu WR, Shi GM, Gao Q, Wang XY, Ding ZB, Fan J. Tumor-associated macrophages mediate resistance to oxaliplatin via inducing autophagy in hepatocellular carcinoma. Cancer Cell Int. 2015;15:71. https://doi.org/10.1186/1475-2891-15-71.

48. Gordon MS, Kanegai CM, Doerr JR, Wall R. Somatic hypermutation of the immunoglobulin heavy chain gene in human chronic lymphocytic leukemia. Proc Natl Acad Sci U S A. 2001;98(2):97-106. https://doi.org/10.1073/pnas.011290398.

49. Chu PG, Arber DA. CD79: a review. Appl Immunohistochem Mol Morphol. 2003;11(4):237-44. https://doi.org/10.1097/00129039-200312000-00001.

50. Musette P, Bouaziz JD. B cell modulation strategies in autoimmune diseases: new concepts. Front Immunol. 2018;9:622. https://doi.org/10.3389/fimmu.2018.00622.

51. Zhang X, Wei L, Wang J, Qin Z, Wang J, Lu Y, Zheng X, Peng Q, Ye Q, Ai F, Liu P, Wang S, Li G, Shen S, Ma J. Suppression Colitis and colitis-associated colon cancer by Anti-S100a9 antibody in mice. Front Immunol. 2017;8:1774. https://doi.org/10.3389/fimmu.2017.01774.

52. Lee TH, Chang HS, Bae DJ, Song HJ, Kim MS, Park JS, Jun JA, Lee SY, UH ST, Kim SH, Park CS. Role of S100A9 in the development of neutrophilic inflammation in asthmatics and in a murine model. Clin Immunol. 2017;183:158-66. https://doi.org/10.1016/j.clim.2017.08.013.

53. Wu R, Duan L, Cui F, Cao J, Xiang Y, Tang Y, Zhou L. S100A9 promotes human hepatocellular carcinoma cell growth and invasion through RAGE-mediated ERK1/2 and p38 MAPK pathways. Exp Cell Res. 2015;334(2):226-38. https://doi.org/10.1016/j.yexcr.2015.04.008.

54. Wei R, Zhu WW, Yu GY, Wang X, Gao C, Zhou X, Lin ZF, Shao WQ, Wang SH, Lu M, Qin LX. S100 calcium-binding protein A9 from tumor-associated macrophages enhances cancer stem cell-like properties of hepatocellular carcinoma. Int J Cancer. 2021;148(5):1233-44. https://doi.org/10.1002/ijc.33371.

55. Emery JG, McDonnell P, Burke MB, Deen KC, Lyn S, Silverman C, Duf E, Appelbaum ER, Eichman C, DiPrinzio R, Dodds RA, James IE, Rosenberg M, Lee JC, Young PR. Osteoprotegerin is a receptor for the cytotoxic ligand TRAIL. J Biol Chem. 1998;273(23):14637-7. https://doi.org/10.1074/jbc.273.23.14637.

56. Jung K, Lein M, Ringsdorf M, Rojas J, Schnoor D, Loeing SA, Staack A. Diagnostic and prognostic validity of serum bone turnover markers in metastatic renal cell carcinoma. J Urol. 2006;176(4 Pt 1):1326-31. https://doi.org/10.1016/j.juro.2006.06.028.

57. Luan F, Li X, Cheng X, Huangfu L, Han J, Guo T, Du H, Wen X, Ji J. TNFRSF11B activates Wnt/beta-catenin signaling and promotes gastric cancer progression. Int J Biol Sci. 2020;16(11):1956-71. https://doi.org/10.7150/ijbs.46360.

58. Gao YB, Xiang ZL, Zhou HY, Wu ZF, Fan J, Zeng HY, Zeng ZC. Enhanced production of CTGF and IL-11 from highly metastatic hepatoma cells under hypoxic conditions: an implication of hepatocellular carcinoma metastasis to bone. J Cancer Res Clin Oncol. 2013;139(4):669-79. https://doi.org/10.1007/s00432-013-1370-4.

59. Huang Z, Chu L, Liang J, Tan X, Wang Y, Wen J, Chen J, Wu Y, Liu S, Liao J, Hou R, Ding Z, Zhang Z, Liang H, Song Y, Chang Z, Zhang J, Guo T, Chen X, Zhang B. H19 promotes HCC bone metastasis through reducing OPG expression in a PP1CA/p38MAPK-dependent manner and sponging miR-200b-3p. Hepatology. 2020. https://doi.org/10.1002/hep.31673.

60. Jiang R, Xia Y, Li J, Deng L, Zhao L, Shi J, Wang X, Sun B. High expression levels of IKKalpha and IKKbeta are necessary for the malignant properties of liver cancer. Int J Cancer. 2010;126(5):1263-74. https://doi.org/10.1002/ijc.24854.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.