An immune-related gene signature for predicting survival and immunotherapy efficacy in hepatocellular carcinoma

Yifei Dai1 · Weijie Qiang2 · Kequan Lin3 · Yu Gui4 · Xun Lan1 · Dong Wang5

Received: 1 July 2020 / Accepted: 12 October 2020 / Published online: 21 October 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract
Background Hepatocellular carcinoma (HCC) ranks the fourth in terms of cancer-related mortality globally. Herein, in this research, we attempted to develop a novel immune-related gene signature that could predict survival and efficacy of immunotherapy for HCC patients.

Methods The transcriptomic and clinical data of HCC samples were downloaded from The Cancer Genome Atlas (TCGA) and GSE14520 datasets, followed by acquiring immune-related genes from the ImmPort database. Afterwards, an immune-related gene-based prognostic index (IRGPI) was constructed using the Least Absolute Shrinkage and Selection Operator (LASSO) regression model. Kaplan–Meier survival curves as well as time-dependent receiver operating characteristic (ROC) curve were performed to evaluate its predictive capability. Besides, both univariate and multivariate analyses on overall survival for the IRGPI and multiple clinicopathologic factors were carried out, followed by the construction of a nomogram. Finally, we explored the possible correlation of IRGPI with immune cell infiltration or immunotherapy efficacy.

Results Analysis of 365 HCC samples identified 11 differentially expressed immune-related genes, which were selected to establish the IRGPI. Notably, it can predict the survival of HCC patients more accurately than published biomarkers. Furthermore, IRGPI can predict the infiltration of immune cells in the tumor microenvironment of HCC, as well as the response of immunotherapy.

Conclusion Collectively, the currently established IRGPI can accurately predict survival, reflect the immune microenvironment, and predict the efficacy of immunotherapy among HCC patients.

Keywords Hepatocellular carcinoma · Immune-related gene · Prognostic index · Immune microenvironment · Cancer immunotherapy

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00262-020-02743-0) contains supplementary material, which is available to authorized users.
Introduction

According to the Global Cancer Report of 2018, hepatocellular carcinoma (HCC) is among the most prevalent malignancies and ranks the fourth in terms of cancer-related mortality globally [1]. HCC accounts for nearly 90% of all primary liver cancer, which is considered as the most common type [2]. At present, the 5-year survival rate of this disease is as low as 14.1% in China [3]. Even for patients at the earliest stages, surgical resection, accepted as the optimal option, is also accompanied by a high recurrence rate [4, 5], making the overall prognosis of HCC patients far from satisfaction. Consequently, it is urgently demanded to predict survival and to improve the clinical outcome of HCC patients.

In recent years, rapid progress has been made in the treatment of liver cancer. Among the advent of a wealth of cutting-edge treatments, immunotherapy has gradually become a hot spot for liver cancer [6–8]. Immunotherapy is characterized by stimulating specific immune responses, inhibiting and killing tumor cells, thereby attenuating the rate of tumor recurrence and metastasis. International guidelines have clearly proposed that immunotherapy could be selected as an effective treatment for patients with advanced liver cancer [9]. However, only a small percentage of the population could benefit from immunotherapy. As an indispensable component of immunotherapy, the tumor immune microenvironment (TIME) has gradually acquired accumulative attention, and the analysis of TIME will contribute to the improvement of immunotherapy responsiveness. Some researchers revealed that the TIME could be taken as a main prognostic indicator, which could also enhance the potential of precision treatments [10, 11]. Therefore, it is suitable and feasible to construct an immune-related gene signature that is closely related to TIME, aiming at predicting immunotherapy efficacy.

Although a number of HCC signatures have been established based on immune-related genes [12–14], a more comprehensive and reliable index is urgently demanded, which can simultaneously predict survival and the efficacy of immunotherapy for HCC patients. To this end, herein, based on the cancer genomics and bioinformatics, we established an immune-related gene-based prognostic index (IRGPI), followed by the validation of its reliability through several data sets. Further, we explored the prognostic value of IRGPI, and the potential predictive role in immunotherapy efficacy. The experimental technical roadmap is summarized in Fig. 1. We expected the IRGPI could provide a foundation for cancer immunotherapy of HCC patients.

Methods

Collection of sample information

Clinical information and transcriptomic data of HCC samples were downloaded from The Cancer Genome Atlas (TCGA) data portal (https://portal.gdc.cancer.gov/) as well as Gene Expression Omnibus (GEO) (https://www.ncbi.nlm.nih.gov/gds/), which were named as entire TCGA cohort (n = 365) and GSE14520 cohort (n = 221), respectively [15, 16]. The entire TCGA cohort was randomly and equally categorized into TCGA training set (n = 183) and TCGA validation set (n = 182). In addition, the entire TCGA cohort and GSE14520 cohort were used as the internal testing set and external testing set, respectively. Patient demographics and clinical characteristics of the included datasets were summarized in Tab. S1. Furthermore, 1,811 unique immune-related genes (IRGs) were obtained from the Immunology Database and Analysis Portal (ImmPort) database (https://www.immport.org/home) [17].

Differentially expressed immune-related genes (DEIRGs)

R package “limma” was utilized to identify differentially expressed genes (DEGs) between 365 HCC specimens and 50 normal specimens according to the criteria of \( |\log_2(\text{Fold Change})| > 1 \) and false discovery rate (FDR) < 0.05 [18], followed by extraction of DEIRGs from DEGs. The volcano plot of DEIRGs was plotted using R package “ggplot2” [19]. Additionally, a Venn diagram was drawn by an online tool (https://bioinformatics.psb.ugent.be/webtools/Venn/) for visualization of the intersections between DEGs and IRGs.
Afterwards, functional enrichment analysis was performed to examine the biological functions of these DEIRGs, including Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) via the Database for Annotation, Visualization, and Integrated Discovery (DAVID) 6.8 [20]. GO terms included biological process (BP), molecular function (MF) as well as cellular component (CC) [21, 22]. The enrichment of GO terms and KEGG signaling pathways was based on the criteria of FDR < 0.05, followed by visualization of the top 10 most significant GO terms as well as KEGG signaling pathways via R package “ggplot2” [19].

**Signature development and reliability evaluation**

The prognosis-related IRGs were identified and an IRGPI was established based on the training set, followed by validation of its predictive performance in other datasets. To be specific, during the exploration of prognosis-related
IRGs, univariate Cox proportional hazard regression analysis was conducted to evaluate the correlation of DEIRGs with overall survival (OS) in the training set. With the cutoff value of \( P < 0.05 \), the prognosis-related IRGs were identified. The optimal model based on prognosis-related IRGs was subsequently identified by the Least Absolute Shrinkage and Selection Operator (LASSO) penalized Cox proportional hazards regression via R package “glmnet” [23]. The IRGs that incorporated into the model were referred to as hub IRGs, and the differential expression of these genes was validated using the Oncomine database [24]. Moreover, this model was used to construct the IRGPI to predict the prognosis of HCC patients. The risk score of each HCC patient was calculated by the following formula: risk score = \([\text{Expression level of Gene 1} \times \text{coefficient}] + [\text{Expression level of Gene 2} \times \text{coefficient}] + \ldots + [\text{Expression level of Gene} n \times \text{coefficient}]\). Patients were further categorized into low- and high-risk groups based on the median value of risk score.

For further validation of the predictive performance of IRGPI, the Kaplan–Meier (K–M) survival curves were applied for survival comparison between low- and high-risk groups via R package “survival” [25]. Additionally, the time-dependent receiver operating characteristic (ROC) curve analysis (including 1-, 3-, and 5-year survival) was established to reflect the sensitivity and specificity of IRGPI using R package “survivalROC” [26]. Meanwhile, the ROC curve was also used to compare the performance of our IRGPI with other published immune-related signatures and widely used biomarkers of cancer immunotherapy. Thereinto, a TP53-associated immune prognostic model established on two genes was named as “Long signature” [12], while a 10 gene-based signature that was associated with tumor microenvironments was named as “Pan signature” [13]. And a risk score prognostic model based on eight genes was named as “Zhang signature” [14].

**Association between IRGPI and clinicopathologic factors**

Univariate and multivariate analyses on OS for IRGPI and clinicopathologic factors were carried out in the entire TCGA cohort and GSE14250 cohort using R package “survival” [25]. Moreover, independent \( t \) tests were applied to evaluate the association of IRGPI with different clinicopathological factors.

**Construction of prognostic nomogram**

For providing a quantitative analysis tool to predict the survival risk of HCC patients, the nomogram was further constructed on the basis of IGRPI as well as clinical parameters. Meanwhile, calibration curves were drawn for comparison of the predictive and actual survival to evaluate the predictive performance of nomograms. The nomogram and calibration curves were plotted via R package “rms” [27].

**Assessment of immune cell infiltration**

Immune cell infiltration was estimated from RNA-sequencing data using CIBERSORT, which is an excellent tool for analyzing the expression matrix of immune cell subtypes based on the principle of linear support vector regression [28].

**Analysis of immunotherapy efficacy**

Immunophenoscore (IPS) can well predict the response of immune checkpoint inhibitors (ICIs), whose scores are based on the expression of important components of tumor immunity, including MHC molecules, immunoregulatory factors, effector cells, and suppressor cells. In addition, the calculation of IPS score is based on representative cell-type gene expression \( z \)-scores with a scale ranging from 0 to 10. The IPS of each HCC patient was derived from The Cancer Immunome Atlas (TCIA) (https://tcia.at/home) [29], followed by an analysis of expression on several prominent checkpoints. Moreover, tumor mutation burden (TMB) can reflect the total number of mutations in tumor cells, which could be utilized for assessing the therapeutic effect of immunotherapy [30]. To explore the correlation between the IRGPI and TMB, we analyzed the available somatic mutation data in the entire TCGA cohort. The mutation data of HCC patients were downloaded and stored as MAF format in the TCGA data portal [31]. And TMB analysis was conducted by R package “maftools” [32].

**Statistical analysis**

Univariate and multivariate Cox regression analyses were conducted via R package “survival” [25], along with hazard ratios (HRs) and 95% confidence intervals (CIs). Moreover, the difference of various clinical factors was compared by the independent \( t \) test. A \( P < 0.05 \) indicated statistical significance.

**Results**

**Construction of IRGPI**

The analysis of 365 HCC samples and 50 normal samples gave rise to 7667 DEGs, including 7273 up-regulated as well as 394 down-regulated genes. In addition, 329 DEIRGs were extracted from DEGs, including 267 up-regulated and 62 down-regulated genes (Fig. 2a, b). Functional enrichment...
analysis revealed that the most relevant signaling pathways to the DEIRGs was “cytokine–cytokine receptor interaction” (Fig. 2c). Meanwhile, the most enriched term in the aspect of biological process (BP), molecular function (MF), and cellular component (CC) was “immune response”, “extracellular space”, and “growth factor activity”, respectively (Fig. 2d).

In the training set, 81 DEIRGs were significantly relevant to the OS of HCC patients ($P < 0.05$, Table S2). After minimizing overfitting by LASSO regression model (Fig. S1), lambda.min was selected to prevent overfitting, which helped to produce the model with a higher accuracy rate. 11 genes were selected as hub IRGs: NDRG1, FABP6, MAPT, HSP90AA1, CD320, CACYBP, BRD8, OSGIN1, NRAS, ISG20L2, and PSMD14 (Fig. 2e).

The expression levels of these 11 IRGs were significantly increased in a wide variety of tumor tissue than normal tissue (Fig. S2). IRGPI was, therefore, established by means of expression data of hub IRGs multiplied by the Cox regression coefficient as follows: risk score = [Expression level of NDRG1 × 0.007898] + [Expression level of FABP6 × 0.032016] + [Expression level of MAPT × 0.04243] + [Expression level of HSP90AA1 × 0.000435] + [Expression level of CD320 × 0.014474] + [Expression level of CACYBP × 0.00145] + [Expression level of BRD8 × 0.003685] + [Expression level of OSGIN1 × 0.001297] + [Expression level of NRAS × 0.003575] + [Expression level

---

Fig. 2 Establishment of IRGPI for HCC. a Volcano plot of DEIRGs between HCC and normal tissues. b Venn diagram visualizing the intersections of DEGs with IRGs. c KEGG enrichment analysis of 329 DEIRGs. d GO enrichment analysis of 329 DEIRGs. Green, blue, and orange parts indicated biological process, cellular component, and molecular function, respectively. e Forest plot demonstrating the multivariable Cox model results of each gene in IRGPI.
IRGPI predicts survival of HCC patients

HCC patients were categorized into low- and high-risk groups based on the median value of risk score of IRGPI (shown in Fig. 3a), and the IRGPI threshold is 0.879126. Significantly worse OS was observed in high-risk patients than low-risk patients (Fig. 3b, P < 0.05). Afterwards, the reliability of IRGPI was determined by time-dependent ROC curves (Fig. 3c). As a result, the area under curve (AUC) was 0.809, 0.717 and 0.622 in 1-year, 3-year and 5-year survival, respectively, in TCGA training set, which indicated the good potential of the constructed IRGPI in monitoring survival. These curves were also applied in TCGA validation set, and the AUC was 0.767, 0.663 and 0.721 for 1-year, 3-year and 5-year survival, respectively. Meanwhile, we found that IRGPI also had a high predictive accuracy of survival in the entire TCGA cohort and GSE14520 cohort. Moreover, ROC curves were used to compare the prediction performance of IRGPI with other signatures. As a result, IRGPI achieved consistently superior performance, whether in comparison with other published immune-related signatures or widely used biomarkers of cancer immunotherapy (Fig. 3d–f). These

![Fig. 3 IRGPI accurately predicts survival of HCC patients. Risk score distribution, survival status, and expression of 11 IRGs for patients in low- and high-risk groups (a), KM survival analyses (b), time-dependent ROC curve analyses (c) in TCGA training set, TCGA validation set, entire TCGA cohort, and GSE14520 cohort. d ROC curves for the performance of IRGPI and other published immune-related signatures. e ROC curves for the performance of IRGPI and widely used biomarkers of cancer immunotherapy. f The area under the ROC curve (AUC) for all candidate biomarkers](image-url)
results indicated that IRGPI was a highly reliable index and superior to other signatures.

**IRGPI is an independent prognostic indicator**

To prove the independence of IRGPI, Cox proportional hazards regression analysis was conducted in the entire TCGA cohort and GSE14520 cohort. As shown in Table 1, univariate and multivariate analyses revealed a significant correlation between IRGPI and OS \( (P < 0.05) \). Therefore, IRGPI was considered as an independent prognostic indicator in entire TCGA cohort \[ HR \ (95\% \ CI) = 2.973 \ (1.966–4.496), \ P < 0.001 \]. After elimination of cases with unknown M stage \( (Mx, n = 99, > 27\%) \) and unknown N stage \( (Nx, n = 113, > 31\%) \), the sample size of entire TCGA cohort was small, thus M stage and N stage were not included in the analysis. In addition, this index was also capable of independently predicting OS in the GSE14520 cohort \[ HR \ (95\% \ CI) = 2.090 \ (1.034–4.225), \ P = 0.040 \]. Taken together, the above outcomes strongly indicated that IRGPI was an independent prognostic factor.

**IRGPI significantly correlates with disease progression**

To explore the possible relationships between IRGPI and multiple clinicopathologic factors, correlation analysis was conducted via independent \( t \) tests. In the entire TCGA cohort, the risk score was significantly higher in patients with advanced tumor grade, advanced pathological stage, and advanced T stage \( (P < 0.05, \text{Fig. 4a}) \). In the GSE14520 cohort, higher risk score was more commonly detected in male patients, and those with larger tumor size, advanced TNM staging, and increased alpha-fetoprotein (AFP) \( (P < 0.05, \text{Fig. 4b}) \). These findings demonstrated that IRGPI was statistically correlated with multiple clinicopathological factors, and a higher risk score generally indicated poorer clinical pathological status.

Additionally, based on IRGPI and some clinicopathological factors, we constructed a prognostic nomogram, aiming at providing a quantitative analysis tool that can predict the survival risk of individual patients \( (\text{Fig. 4c}) \). More importantly, the calibration curves of the prognostic nomogram showed the good consistency between predictive and actual 1-, 3-, and 5-year survival in the entire TCGA cohort \( (\text{Fig. 4d}) \).

**Table 1** Univariate and multivariate Cox regression analyses of IRGPI and other clinicopathologic factors for OS in the entire TCGA cohort and GSE14520 cohort

| Overall survival | Univariate analysis | Multivariate analysis |
|-----------------|--------------------|----------------------|
|                 | HR  | 95% CI | \( P \) value | HR  | 95% CI | \( P \) value |
| **Entire TCGA cohort** | | | | | | |
| Age             | 1.012 | 0.996–1.029 | 0.139 | 1.003 | 0.987–1.020 | 0.714 |
| Gender (male vs. female) | 0.779 | 0.516–1.174 | 0.232 | 0.801 | 0.515–1.246 | 0.325 |
| Tumor status (with tumor vs. tumor free) | 1.600 | 1.074–2.383 | 0.021* | 1.669 | 1.093–2.549 | 0.018* |
| Tumor grade     | 1.085 | 0.831–1.416 | 0.551 | 0.961 | 0.709–1.302 | 0.796 |
| Pathological stage | 1.693 | 1.362–2.104 | <0.001*** | 0.828 | 0.349–1.962 | 0.667 |
| T stage         | 1.680 | 1.369–2.063 | <0.001*** | 1.741 | 0.781–3.881 | 0.175 |
| IRGPI (high risk vs. low risk) | 3.253 | 2.280–4.641 | <0.001*** | 2.973 | 1.966–4.496 | <0.001*** |
| **GSE14520 cohort** | | | | | | |
| Age             | 0.990 | 0.971–1.010 | 0.321 | 0.995 | 0.972–1.019 | 0.690 |
| Gender (male vs. female) | 1.658 | 0.800–3.436 | 0.174 | 1.125 | 0.527–2.403 | 0.761 |
| ALT(> / < = 50U/L) | 1.085 | 0.704–1.671 | 0.713 | 0.824 | 0.510–1.329 | 0.426 |
| Main tumor size (> / < = 5 cm) | 2.087 | 1.354–3.215 | <0.001*** | 0.769 | 0.427–1.387 | 0.383 |
| Multinodular (yes/no) | 1.553 | 0.961–2.510 | 0.073 | 0.304 | 0.160–0.576 | <0.001*** |
| Cirrhosis (yes/no) | 4.757 | 1.170–19.351 | 0.029* | 3.722 | 0.889–15.589 | 0.072 |
| TNM staging     | 2.238 | 1.685–2.971 | <0.001*** | 1.635 | 1.072–2.415 | 0.013* |
| BCLC staging    | 2.144 | 1.693–2.714 | <0.001*** | 1.439 | 0.991–2.090 | 0.056 |
| CLIP staging    | 1.892 | 1.531–2.337 | <0.001*** | 2.080 | 1.396–3.099 | <0.001*** |
| AFP (> / < = 300 ng/ml) | 1.655 | 1.078–2.542 | 0.021* | 0.534 | 0.264–1.081 | 0.081 |
| IRGPI (high risk vs. low risk) | 2.724 | 1.405–5.281 | 0.003** | 2.090 | 1.034–4.225 | 0.040* |

* \( P < 0.05; ** \( P < 0.01; *** \( P < 0.001 \)
IRGPI predicts the infiltration of immune cells into HCC microenvironment

For further exploration of the indicative roles of IRGPI on TIME, it is necessary to investigate the types of infiltrating immune cells in HCC patients. As an excellent tool to estimate immune cell infiltration, CIBERSORT was adopted for evaluation of the relative proportion of 21 types of immune cells in all HCC specimens. As shown in Fig. 5a, the highly infiltrating immune cells include M0 macrophages, resting memory CD4 T cells, CD8 T cells and M2 macrophages. Among the 21 types of immune cells, the relative proportion of naive B cells, resting memory CD4 T cells, and monocytes had a significant negative correlation with risk score, while the relative proportion of activated memory CD4 T cells and M0 macrophages had a significant positive correlation with risk score.
correlation with risk score ($P < 0.05$, Fig. 5a). Besides, we also analyzed the correlation between 11 IRGs contained in IRGPI and the infiltration levels of 21 types of immune cells in all HCC specimens, respectively. As shown in Fig. S3, most of these 11 IRGs are negatively related to the infiltration levels of multiple immune cells, such as monocytes and M1 macrophages, while positively related to the infiltration levels of M0 macrophages and activated mast cells. These results reveal the potential of these 11 genes to reflect tumor microenvironment.

In addition, we explored the prognostic value of 21 types of immune cells, among which the infiltrating abundance of M0 macrophages (Fig. 5b), M2 macrophages (Fig. 5c), activated memory CD4 T cells (Fig. 5d) as well as CD8 T cells (Fig. 5e) were significantly related to OS ($P < 0.05$). The higher infiltrating abundance of M2 macrophages was associated with poorer OS, while the higher infiltrating abundance of CD8 T cells was related to better OS. Collectively, IRGPI was statistically correlated with the infiltration level of most immune cells, implying that our IRGPI could potentially reflect the state of TIME.

**IRGPI predicts responses of immunotherapy**

To further explore the association of IRGPI with immunity, the correlation analysis was conducted between IRGPI and immune functions. As shown in Fig. 6a, IRGPI was positively correlated with MDSC recruiting, but negatively with infiltration of immune cells into tumors ($P < 0.05$). As a well-known biomarker of immunotherapy, we also analyzed the relationship between tumor mutation burden (TMB) and IRGPI, revealing the positive correlation of IRGPI with TMB (Fig. S4). Moreover, to predict the response of immune checkpoint inhibitors (ICIs), the correlation between IRGPI

---

**Fig. 5** Analysis of immune cell infiltration. a Correlations of IRGPI with immune cell infiltration. The blue and red violin represented the IRGPI low- and high-risk group, respectively. The white points inside the violin implicated median values. The relationships between OS and M0 macrophages (b), M2 macrophages (c), memory activated CD4 T cells (d), and CD8 T cells (e)
and immunophenoscore (IPS) in HCC patients was explored. IPS has been proved excellent in predicting the response of ICIs in several studies [29, 33]. The major immune checkpoints include cytotoxic T lymphocyte antigen-4 (CTLA-4), programmed cell death protein 1 (PD-1), programmed death ligand-1 (PD-L1) as well as programmed death ligand-2 (PD-L2). Thus, the scores of IPS, IPS-CTLA4 blocker, IPS-PD1/PD-L1/PD-L2 blocker, and IPS-CTLA4 + PD1/PD-L1/PD-L2 blocker were used for evaluating the potential application of ICIs. As shown in Fig. 6b, the IPS and IPS-CTLA4 blocker were significantly increased in the low-risk group. The score of IPS and IPS-CTLA4 blocker were significantly increased in the low-risk group. The expression of CTLA-4, LAG-3, PD-1, TIGIT, and TIM-3 was significantly elevated in the high-risk group.

Fig. 6 Analysis of immunotherapy responses. a Correlation matrix of IRGPI and anticancer immune responses. The blue dots represent a negative correlation, and the red dots represent a positive correlation. *, ** and *** represent $P < 0.05$, $P < 0.01$ and $P < 0.001$, respectively. b The relationship between the IRGPI and IPS. The score of IPS and IPS-CTLA4 blocker were significantly increased in the low-risk group. c The relationship between the IRGPI and several prominent immune checkpoints. The expression of CTLA-4, LAG-3, PD-1, TIGIT, and TIM-3 was significantly elevated in the high-risk group.

Discussion

An increasing body of evidence has suggested the close correlation of TIME with tumorigenesis and cancer progression [34–36]. By analyzing the immune landscape of HCC
microenvironment, some researchers pointed out that the immune contexture could be a major prognostic indicator, and should not be disregarded to enhance the potential of precision treatments [37]. At present, immunotherapy has been widely recognized to treat a variety of cancers including HCC [38–40]. However, not all patients can benefit from it. Therefore, it is necessary to establish an IRG signature for survival prediction of HCC patients and enriching the effective population of cancer immunotherapy.

During the past years, genomics and bioinformatics have enabled the identification of molecular signatures. For example, several signatures have been developed for prognostic prediction based on lncRNA, miRNA, and mRNA [41, 42]. In this study, IRGPI was constructed by integrating the clinical information and transcriptomic data of HCC samples in the TCGA cohort and GSE14520 cohort. A total of 329 DEIRGs were identified, of which the most relevant biological process and signaling pathway was “immune response” and “cytokine–cytokine receptor interaction”, respectively. This result was closely associated with immune, which was consistent with some existing literature reports [43]. Subsequently, Cox regression analysis and LASSO regression model identified 11 out of 81 prognosis-related IRGs, which were used to construct IRGPI, including NDRG1, FABP6, MAPT, HSP90AA1, CD320, CACYBP, BRD8, OSGIN1, NRAS, ISG20L2, and PSMD14. Among them, NDRG1 has been reported to be an essential molecule in controlling the metastasis and recurrence of HCC [44]. In addition, the deletion of CACYBP has also been reported to increase apoptosis of HCC cells [45], while the variants of OSGIN1 could reduce apoptosis and are associated with shorter survival [46]. Besides, knockdown and overexpression assays have demonstrated that PSMD14 could promote migration and invasion of HCC cells in vitro, and facilitate tumor growth and metastasis in vivo [47]. Although the direct association between the other seven genes and HCC has not been discovered, we think that the underlying correlations deserve further experimental validation.

In consideration of the importance of immune cell infiltration in tumors, CIBERSORT was further adopted for evaluating the relative proportion of 21 types of immune cells in every HCC specimen. Some evidence has indicated that the interplay between tumor and microenvironment plays a critical role in HCC progression and the probability of response to immunotherapies. So, we assessed the potential of IRGPI to reflect immune cell infiltration, as well as the prognostic value of different types of immune cells. Generally speaking, the high abundance of activated memory CD4+ T cells contributes to immune response and is associated with better prognosis, which also apply to CD8 T cells [48]. Besides, it has been reported that tumor-associated macrophages are mainly M2 macrophages which can promote tumor growth [49]. Consistent with this, our study suggested that the patients with higher infiltrating abundance of CD8 T cells have a better prognosis. Although some partial results seem to be not consistent, they should be acceptable from an overall perspective.

The advent of immunotherapy has shed novel light on HCC treatment, of which ICIs have become a potentially effective treatment [6]. Targeting immune checkpoint molecules such as PD-1 and CTLA-4 could reinvigorate anti-tumor immunity [50]. Recently, nivolumab and pembrolizumab, two therapeutics against PD-L1/PD1, have been recently approved for subsequent-line therapy [51]. To predict the reactivity of ICIs, the relationship between IRGPI and IPS was explored in HCC patients. The analysis indicated that the low-risk group had higher IPS and IPS-CTL4 scores, revealing that IRGPI has the potential to determine the specific HCC patients who are immunogenic and more responsive to ICIs. The predictive value of IRGPI on the response to ICIs provides a theoretical basis for the therapeutic selection of ICIs in clinical practice. Hopefully, this predictive model could assist to accelerate the pace of individualized cancer immunotherapy.

To further enhance the accuracy of prognostic prediction, we constructed and validated a nomogram by integrating IRGPI, age, gender, tumor status, tumor grade, pathological stage and T stage. Similarly, Ying et al. [52] combined inflammatory biomarkers with risk factors to form a nomogram, which could improve the accuracy for predicting clinical outcomes in CRC patients undergoing surgical resection. More importantly, these new prognostic tools could not only improve the accuracy of prognostic prediction, but also help to predict the specific survival risk of individual patients, which is of great significance in clinical practice [53].

There are several strengths in this study. First, this signature was sufficiently validated and evaluated in multiple datasets, indicating the robustness and reliability of the signature. Second, comprehensive and in-depth researches were carried out in various aspects, including discussions on the correlation of IRGPI with the immune cells, IPS and TMB. Third, a nomogram was further established for the quantitative calculation, which is conducive to clinical promotion and application. Nevertheless, several limitations still exist in our study. Thus, more HCC patients and validations are warranted to further test this signature by prospective studies in the future.

In conclusion, we have constructed an IRG-based index that is closely related to the immune microenvironment, which can better predict survival and reflect the efficacy of immunotherapy for HCC patients. In the era of precision medicine, the IRG-based index could hopefully provide an effective tool to meet the clinical requirements of HCC treatment to a certain extent.
Author contributions YD, WQ and DW conceived the study; YD, WQ, KL, XL and DW designed the experiments; YD performed the experiments; YD and WQ wrote the manuscript; KL, YG, XL and DW edited the manuscript; and all authors read and gave final approval to submit the manuscript.

Funding This work was supported by grants from the National Natural Science Foundation of China (Grant No. 81673460) and Science and Technology Department of Sichuan Province (Grant No. 2020JDTD0022).

Compliance with ethical standards

Competing interests The authors declare that they have no competing interests.

Availability of data and material The datasets used during the current study are available from the corresponding author on reasonable request.

References

1. Bray F, Ferlay J, Soerjomataram I et al (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 68:394–424. https://doi.org/10.3322/caac.21492
2. Llovet JM, Zucman-Rossi J, Pikarsky E et al (2016) Hepatocellular carcinoma. Nat Rev Dis Prim 2:16018. https://doi.org/10.1038/nrdp.2016.18
3. Allemani C, Matsuda T, Di Carlo V et al (2018) Global surveillance of trends in cancer survival 2000–14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. The Lancet 391:1023–1075. https://doi.org/10.1016/S0140-6736(17)33326-3
4. Fujiwara N, Friedman SL, Goossens N, Hoshida Y (2018) Risk factors and prevention of hepatocellular carcinoma in the era of precision medicine. J Hepatol 68:526–549. https://doi.org/10.1016/j.jhep.2017.09.016
5. Fumaluro S, Di Sandro S, Giani A et al (2018) Recurrence patterns after anatomic or parenchyma-sparing liver resection for hepatocarcinoma in a western population of cirrhotic patients. Am Surg Oncol 25:3974–3981. https://doi.org/10.1045/s10434-018-6730-0
6. Llovet JM, Montal R, Sia D, Finn RS (2018) Molecular therapies and precision medicine for hepatocellular carcinoma. Nat Rev Clin Oncol 15:599–616. https://doi.org/10.1038/s41571-018-0073-4
7. Ish harairaegui M, Melero I, Sangro B (2018) Immunotherapy of hepatocellular carcinoma: facts and hopes. Clin Cancer Res 24:1518–1524. https://doi.org/10.1158/1078-0432.CCR-17-0289
8. Heinrich B, Czau druga C, Marquardt JU (2018) Immunotherapy of hepatocellular carcinoma. Oncol Res Treat 41:292–297. https://doi.org/10.1159/000488916
9. Galle PR, Forner A, Llovet JM et al (2018) EASL clinical practice guidelines: management of hepatocellular carcinoma. J Hepatol 69:182–236. https://doi.org/10.1016/j.jhep.2018.03.019
10. Taue JM, Galon J, Sholl LM et al (2018) Implications of the tumor immune microenvironment for staging and therapeutics. Mod Pathol 31:214–234. https://doi.org/10.1038/modpa.thol.2017.156
11. Xu WH, Xu Y, Wang J et al (2019) Prognostic value and immune infiltration of novel signatures in clear cell renal cell carcinoma microenvironment. Aging 11:6999–7020. https://doi.org/10.18632/aging.102233
12. Long J, Wang A, Bai Y et al (2019) Development and validation of a TP53-associated immune prognostic model for hepatocellular carcinoma. EBioMedicine 42:363–374. https://doi.org/10.1016/j.ebiom.2019.03.022
13. Pan L, Fang J, Chen MY et al (2020) Promising key genes associated with tumor microenvironments and prognosis of hepatocellular carcinoma. World J Gastroenterol 26:789–803. https://doi.org/10.3748/wjg.v26.i8.789
14. Zhang FP, Huang YP, Luo WX et al (2020) Construction of a risk score prognostic model based on hepatocellular carcinoma microenvironment. World J Gastroenterol 26:134–153. https://doi.org/10.3748/wjg.v26.i2.134
15. Weinstein JN, Collisson EA, Mills GB et al (2013) The cancer genome atlas pan-cancer analysis project. Nat Genet 45:1113–1120. https://doi.org/10.1038/ng.2764
16. Wang Y, Gao B, Tan PY et al (2019) Genome-wide CRISPR knockout screens identify NCA PG as an essential oncogene for hepatocellular carcinoma tumor growth. FASEB J Off Publ Fed Am Soc Exp Biol 33:8759–8770. https://doi.org/10.1096/fj.20182213RR
17. Bhattacharya S, Andorf S, Gomes L et al (2014) ImmPort: disseminating data to the public for the future of immunology. Immunol Rev 58:234–239. https://doi.org/10.1007/s12026-014-8516-1
18. Ritchie ME, Phipson B, Wu D et al (2015) Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43:e47. https://doi.org/10.1093/nar/gkv007
19. Ginestet C (2011) ggplot2: elegant graphics for data analysis. J Royal Stat Soc: Series A (Stat Soc) 174:245–246. https://doi.org/10.1111/j.1467-985X.2010.00676_9.x
20. Huang DW, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4:44–57. https://doi.org/10.1038/nprot.2008.211
21. Ogata H, Goto S, Sato K et al (2000) KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28:27–30. https://doi.org/10.1093/nar/28.1.27
22. Gene Ontology Consortium (2004) The Gene Ontology (GO) database and informatics resource. Nucleic Acids Res 32:D258–D261. https://doi.org/10.1093/nar/gkh036
23. Friedman J, Hastie T, Tibshirani R (2010) Regularization paths for generalized linear models via coordinate descent. J Royal Stat Soc: Series A (Stat Soc) 174:245–246. https://doi.org/10.1111/j.0006-341X.2008.00448.x
24. Rhodes DR, Yu J, Shanker K et al (2004) ONCOMINE: a cancer microarray database and integrated data-mining platform. Nucleic Acids Res 32:2213R R
25. Friedman J, Hastie T, Tibshirani R (2010) Regularization paths for generalized linear models via coordinate descent. J Royal Stat Soc: Series A (Stat Soc) 174:245–246. https://doi.org/10.1111/j.1467-985X.2010.00676_9.x
26. Therneau TM (2015) A package for survival analysis in S. Version 2.38. CRAN website—https://cran.r-project.org/package=survival
27. Heagerty PJ, Lumley T, Pepe MS (2000) Time-dependent ROC curves for censored survival data and a diagnostic marker. Biometrics 56:337–344. https://doi.org/10.1111/j.0006-341X.2000.00337.x
28. Harrell Jr FE (2016) rms: Regression Modeling Strategies. R package version 5.0–0. CRAN
29. Newman AM, Liu CL, Green MR et al (2015) Robust enumeration of a TP53-associated immune prognostic model for hepatocellular carcinoma. EBioMedicine 42:363–374. https://doi.org/10.1016/j.ebiom.2019.03.022
30. Liu L, Bai X, Wang J et al (2019) Combination of TMB and CNA stratifies prognostic and predictive responses to immunotherapy
across metastatic cancer. Clin Cancer Res 25:7413–7423. https://doi.org/10.1158/1078-0432.CCR-19-0558
31. Jensen MA, Ferretti V, Grossman RL, Staudt LM (2017) The NCI genomic data commons as an engine for precision medicine. Blood 130:453–459. https://doi.org/10.1182/blood-2017-03-735654
32. Mayakonda A, Lin DC, Assenov Y et al (2018) MafTools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res 28:1747–1756. https://doi.org/10.1101/gr.239244.118
33. Yang S, Wu Y, Deng Y et al (2019) Identification of a prognostic immune signature for cervical cancer to predict survival and response to immune checkpoint inhibitors. OncolImmunology 8:e1659094. https://doi.org/10.1080/2162402X.2019.1659094
34. Berraondo P, Minute L, Ajona D et al (2016) Innate immune mediators in cancer: between defense and resistance. Immunol Rev 274:290–306. https://doi.org/10.1111/imr.12464
35. Cariani E, Missale G (2019) Immune landscape of hepatocellular carcinoma microenvironment: implications for prognosis and therapeutic applications. Cell Immunol 333:34–45. https://doi.org/10.1016/j.cellimm.2018.03.008
36. Gardner A, Ruffell B (2016) Dendritic cells and cancer immunity. Trends Immunol 37:855–865. https://doi.org/10.1016/j.it.2016.09.006
37. Carieni E, Missale G (2019) Immune landscape of hepatocellular carcinoma microenvironment: implications for prognosis and therapeutic applications. Liver International 39:1608–1621. https://doi.org/10.1111/liv.14192
38. Banerjee K, Kumar S, Ross KA et al (2018) Emerging trends in the immunotherapy of pancreatic cancer. Cancer Lett 417:35–46. https://doi.org/10.1016/j.canlet.2017.12.012
39. Sanmamed MF, Chen L (2018) A paradigm shift in cancer immunotherapy: from enhancement to normalization. Cell 175:313–326. https://doi.org/10.1016/j.cell.2018.09.035
40. Jiang Y, Han QJ, Zhang J (2019) Hepatocellular carcinoma: mechanisms of progression and immunotherapy. World J Gastroenterol 25:3151–3167. https://doi.org/10.3748/wjg.v25.i25.3151
41. Zhao QJ, Zhang J, Xu L, Liu FF (2018) Identification of a five-long non-coding RNA signature to improve the prognosis prediction for patients with hepatocellular carcinoma. World J Gastroenterol 24:3426–3439. https://doi.org/10.3748/wjg.v24.i30.3426
42. Bing Z, Tian J, Zhang J et al (2016) An integrative model of miRNA and mRNA expression signature for patients of breast invasive carcinoma with radiotherapy prognosis. Cancer Biotherapy Radiopharm 31:253–260. https://doi.org/10.1089/cbr.2016.2059
43. Jiang X, Hao Y (2018) Analysis of expression profile data identifies key genes and pathways in hepatocellular carcinoma. Oncology Letters 15:2625–2630. https://doi.org/10.3892/ol.2017.7534
44. Cheng J, Xie HY, Xu X et al (2011) NDRG1 as a biomarker for metastasis, recurrence and of poor prognosis in hepatocellular carcinoma. Cancer Lett 310:35–45. https://doi.org/10.1016/j.canlet.2011.06.001
45. Liu YF, Huang YL, Zhang YJ et al (2019) CacYBP enhances cytoplastic retention of p27Kip1 to promote hepatocellular carcinoma progression in the absence of RNF41 mediated degradation. Theranostics 9:8392–8408. https://doi.org/10.7150/thno.36838
46. Lian YF, Huang YL, Zhang YJ et al (2019) CacYBP enhances cytoplastic retention of p27Kip1 to promote hepatocellular carcinoma progression in the absence of RNF41 mediated degradation. Theranostics 9:8392–8408. https://doi.org/10.7150/thno.36838
47. Lv J, Zhang S, Wu H et al (2020) Deubiquitinase PSMD14 enhances hepatocellular carcinoma growth and metastasis by stabilizing GRB2. Cancer Lett 469:22–34. https://doi.org/10.1016/j.canlet.2019.10.025
48. Matsumoto H, Thike AA, Li H et al (2016) Increased CD4 and CD8-positive T cell infiltrate signifies good prognosis in a subset of triple-negative breast cancer. Breast Cancer Res Treat 156:237–247. https://doi.org/10.1007/s10549-016-3743-x
49. Yao RR, Li JH, Zhang R et al (2018) M2-polarized tumor-associated macrophages facilitated migration and epithelial-mesenchymal transition of HCC cells via the TLR4/STAT3 signaling pathway. World J Surg Oncol 16; https://doi.org/10.1186/s12957-018-1312-y
50. Choi C, Yoo GS, Cho WK, Park HC (2019) Optimizing radiotherapy with immune checkpoint blockade in hepatocellular carcinoma. World J Gastroenterol 25:2416–2429. https://doi.org/10.3748/wjg.v25.i20.2416
51. Longo V, Brunetti O, Gnoni A et al (2019) Emerging role of immune checkpoint inhibitors in hepatocellular carcinoma. Medicina (Lithuania) 55:698. https://doi.org/10.3390/medicina55100698
52. Ying HQ, Deng QW, He BS et al (2014) The prognostic value of preoperative NLR, d-NLR, PLR and LMR for predicting clinical outcome in surgical colorectal cancer patients. Med Oncol 31:305. https://doi.org/10.1007/s12032-014-0305-0
53. Zhang G, Wu Y, Zhang J et al (2018) Nomograms for predicting long-term overall survival and disease-specific survival of patients with clear cell renal cell carcinoma. Oncotarget 11:5535–5544. https://doi.org/10.2217/OTX.17.1881

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