An Immune-Related Gene Signature for Predicting Survival and Immunotherapy Efficacy in Hepatocellular Carcinoma

Yifei Dai
School of Medicine, Tsinghua University
Weijie Qiang
Chinese Academy of Medical Sciences & Peking Union Medical College Hospital of Skin Diseases and Institute of Dermatology
Kequan Lin
Tsinghua University School of Life Sciences
Yu Gui
Chengdu University of Traditional Chinese Medicine
Xun Lan (xlan@tsinghua.edu.cn)
School of Medicine, Tsinghua University
Dong Wang (dwang@cdutcm.edu.cn)
Chengdu University of Traditional Chinese Medicine

Research

Keywords: Hepatocellular carcinoma, immune-related gene, prognostic index, immune microenvironment, cancer immunotherapy

Posted Date: June 25th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-36927/v1

Version of Record: A version of this preprint was published at Cancer Immunology, Immunotherapy on October 21st, 2020. See the published version at https://doi.org/10.1007/s00262-020-02743-0.

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 acquisition of 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 analysis on overall survival for the IRGPI and multiple clinicopathologic factors were carried out, followed by the construction of 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 genes, which were selected to establish the IRGPI. Notably, it can predict 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.

Background

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 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 immune microenvironment 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.

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 a training cohort and a validation cohort. 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 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 (http://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 were 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 prognosis of HCC patients. The risk score of each HCC patient was calculated by the following formula: risk score = [Expression level of Gene 1 * coefficient] + [Expression level of Gene 2 * coefficient] + … + [Expression level of Gene n * 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 analysis 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 IRGPI 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 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 analysis was 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 7,667 DEGs, including 7,273 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. 1A, 1B). Functional enrichment analysis revealed that the most relevant signaling pathways to the DEIRGs was “cytokine-cytokine receptor interaction” (Fig. 1C). 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. 1D).

In the training set, 81 DEIRGs were significantly relevant to the OS of HCC patients ($P < 0.05$). After minimizing overfitting by LASSO regression model, 11 genes were selected as hub IRGs: *NDRG1*, *FABP6*, *MAPT*, *HSP90AA1*, *CD320*, *CACYBP*, *BRD8*, *OSGIN1*, *NRAS*, *ISG20L2*, and *PSMD14* (Fig. 1E). The expression levels of these 11 IRGs were significantly increased in a wide variety of tumor tissue than normal tissue (Fig. S1). 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* * 0.014227] + [Expression level of *BRD8* * 0.003685] + [Expression level of *OSGIN1* * 0.001297] + [Expression level of *NRAS* * 0.003575] + [Expression level of *ISG20L2* * 0.018457] + [Expression level of *PSMD14* * 0.02678].

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. 2A). Significantly worse OS was observed in high-risk patients than low-risk patients (Fig. 2B, $P < 0.05$). Afterwards, the reliability of IRGPI was determined by time-dependent ROC curves (Fig. 2C). 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 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. 2D-F). These 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 analysis revealed 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.

Table 1
Univariate and multivariate Cox regression analysis 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	1.107–2.415	0.013*
BCLC staging	2.144	1.693–2.714	< 0.001***	1.439	0.991–2.090	0.056

Overall survival	Univariate analysis			Multivariate analysis		
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*

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 ($P < 0.05$, Fig. 3A). 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) ($P < 0.05$, Fig. 3B). 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 (Fig. 3C). 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 (Fig. 3D).

IRGPI predicts the infiltration of immune cells into HCC microenvironment

For further exploration of the indicative roles of IPGRI 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 22 types of immune cells in all HCC specimens. Among the 22 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 ($P < 0.05$, Fig. 4A). In addition, survival analysis was conducted in 22 types of immune cells, showing that the relative proportion of M0 macrophages (Fig. 4B), M2 macrophages (Fig. 4C), activated memory CD4 T cells (Fig. 4D) as well as CD8 T cells (Fig. 4E) were significantly related to OS ($P < 0.05$). A higher proportion of M0 macrophages was associated with poorer OS, while a higher proportion of activated memory CD4 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. 5A, IRGPI was positively correlated with releasing of cancer cell antigens, Treg cell recruiting, and MDSC recruiting, but negatively with CD4 T cell recruiting, infiltration of immune cells into tumors, and killing of cancer cells. 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. S2). Moreover, to predict the response of immune checkpoint inhibitors (ICIs), the correlation between IRGPI 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. 5B, the IPS and IPS-CTLA4 scores were significantly elevated in the low-risk group which was categorized by the IRGPI, implying more immunogenicity on ICIs in the low-risk group. Besides, the expression of some critical immune checkpoints was investigated, showing that the expression of CTLA-4, LAG-3, PD-1, TIGIT, and TIM-3 was significantly higher in high-risk group than low-risk group (Fig. 5C). These results suggested that low-risk group was more likely to have an immune response and respond to immunotherapy.

Discussion

An increasing body of evidence has suggested the close correlation of immune microenvironment 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 identified 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 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 22 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. Our study suggested that IRGPI was significantly and positively associated with the relative proportion of activated memory CD4 T cells and M0 macrophages, which are the only two types of immune cells significantly associated with OS. Some studies have shown that the selective loss or apoptosis of intrahepatic $CD4^+$ T lymphocytes would promote hepatocarcinogenesis [48, 49].

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]. In order 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-CTLA4 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. Firstly, this signature was sufficiently validated and evaluated in multiple datasets, indicating the robustness and reliability of the signature. Secondly, 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. Thirdly, 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.

Conclusion

In this study, 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.

Abbreviations

HCC	hepatocellular carcinoma
TIME	tumor immune microenvironment
TCGA	The Cancer Genome Atlas
IRG	immune-related gene
LASSO	Least Absolute Shrinkage and Selection Operator
IRGPI	immune-related gene-based prognostic index
K-M	Kaplan-Meier
ROC	receiver operating characteristic
OS	overall survival
TMB	tumor mutation burden
IPS	immunophenoscore
GEO	Gene Expression Omnibus
ImmPort	Immunology Database and Analysis Portal
DEG	differentially expressed gene
FDR	false discovery rate
DEIRG	differentially expressed immune-related gene
KEGG	Kyoto Encyclopedia of Genes and Genomes
GO	Gene Ontology
BP	biological process
MF	molecular function
CC	cellular component
DAVID	Database for Annotation, Visualization, and Integrated Discovery
MAF	Mutation Annotation Format
ICI	immune checkpoint inhibitor
TCIA	The Cancer Immunome Atlas
HRs	hazard ratios
CIs	confidence intervals
AUC	area under curve
TX	unknown T stage

MX	unknown M stage
NX	unknown N stage
ALT	alanine transferase
TNM	Tumor Node Metastasis
AFP	alpha-fetoprotein
CTLA4	cytotoxic T lymphocyte antigen 4
PD1	programmed cell death 1
PD-L1	programmed cell death-ligand 1
PD-L2	programmed cell death-ligand 2

Declarations

Authors' Contributions

YFD, WJQ and DW conceived the study; YFD, WJQ, KQL, XL and DW designed the experiments; YFD performed the experiments; YFD and WJQ wrote the manuscript; KQL, YG, XL and DW edited the manuscript; and all authors read and gave final approval to submit the manuscript.

Acknowledgments

The authors sincerely thank the TCGA, ImmPort, GEO, Oncomine, CIBERSORT and TCIA databases for the availability of the data.

Funding

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

Availability of data and materials

The datasets used during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

Not applicable.

Consent for publication

All authors have approved the publication.

Competing interests

The authors declare that they have no competing interests.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicians. 2018;68:394-424.
2. Llovet JM, Zucman-Rossi J, Pikarsky E, Sangro B, Schwartz M, Sherman M, et al. Hepatocellular carcinoma. Nature Reviews Disease Primers. 2016;2:16018.
3. Allemani C, Matsuda T, Di Carlo V, Harewood R, Matz M, Nikšić M, et al. 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. 2018;391:1023-75.
4. Fujiwara N, Friedman SL, Goossens N, Hoshida Y. Risk factors and prevention of hepatocellular carcinoma in the era of precision medicine. Journal of Hepatology. 2018;68:526-49.
5. Famularo S, Di Sandro S, Giani A, Lauterio A, Sandini M, De Carlis R, et al. Recurrence Patterns After Anatomic or Parenchyma-Sparing Liver Resection for Hepatocarcinoma in a Western Population of Cirrhotic Patients. Annals of Surgical Oncology. 2018;25:3974-81.
6. Llovet JM, Montal R, Sia D, Finn RS. Molecular therapies and precision medicine for hepatocellular carcinoma. Nature Reviews Clinical Oncology. 2018;15:599-616.
7. Iñarrairaegui M, Melero I, Sangro B. Immunotherapy of hepatocellular carcinoma: Facts and hopes. Clinical Cancer Research. 2018;24:1518-24.
8. Heinrich B, Czauderna C, Marquardt JU. Immunotherapy of Hepatocellular Carcinoma. Oncology Research and Treatment. 2018;41:292-7.
9. Galle PR, Forner A, Llovet JM, Mazzaferro V, Piscaglia F, Raoul JL, et al. EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma. Journal of Hepatology. 2018;69:182-236.
10. Taube JM, Galon J, Sholl LM, Rodig SJ, Cottrell TR, Giraldo NA, et al. Implications of the tumor immune microenvironment for staging and therapeutics. Modern Pathology. 2018;31:214-34.
11. Xu WH, Xu Y, Wang J, Wan FN, Wang HK, Cao DL, et al. Prognostic value and immune infiltration of novel signatures in clear cell renal cell carcinoma microenvironment. Aging. 2019;11:6999-7020.
12. Long J, Wang A, Bai Y, Lin J, Yang X, Wang D, et al. Development and validation of a TP53-associated immune prognostic model for hepatocellular carcinoma. EBioMedicine. 2019;42:363-74.

13. Pan L, Fang J, Chen MY, Zhai ST, Zhang B, Jiang ZY, et al. Promising key genes associated with tumor microenvironments and prognosis of hepatocellular carcinoma. World Journal of Gastroenterology. 2020;26:789-803.
14. Zhang FP, Huang YP, Luo WX, Deng WY, Liu CQ, Xu LB, et al. Construction of a risk score prognosis model based on hepatocellular carcinoma microenvironment. World Journal of Gastroenterology. 2020;26:134-53.
15. Weinstein JN, Collisson EA, Mills GB, Shaw KRM, Ozenberger BA, Ellrott K, et al. The cancer genome atlas pan-cancer analysis project. Nature Genetics. 2013;45:1113-20.
16. Wang Y, Gao B, Tan PY, Handoko YA, Sekar K, Deivasigamani A, et al. Genome-wide CRISPR knockout screens identify NCAPG as an essential oncogene for hepatocellular carcinoma tumor growth. FASEB journal : official publication of the Federation of American Societies for Experimental Biology. 2019;33:8759-70.
17. Bhattacharya S, Andorf S, Gomes L, Dunn P, Schaefer H, Pontius J, et al. ImmPort: Disseminating data to the public for the future of immunology. Immunologic Research. 2014;58:234-9.
18. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Research. 2015;43:e47.
19. Ginestet C. ggplot2: Elegant Graphics for Data Analysis. Journal of the Royal Statistical Society: Series A (Statistics in Society). 2011;174:245-6.
20. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature Protocols. 2009;4:44-57.
21. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Research. 2000;28:27-30.
22. Gene Ontology Consortium. The Gene Ontology (GO) database and informatics resource. Nucleic Acids Research. 2004;32:D258-61.
23. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. Journal of Statistical Software. 2010;33:1-22.
24. Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D, et al. ONCOMINE: A Cancer Microarray Database and Integrated Data-Mining Platform. Neoplasia. 2004;6:1-6.
25. Therneau TM. A Package for Survival Analysis in S. Version 2.38. CRAN website - http://cran.r-project.org/package=survival. 2015.
26. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. Biometrics. 2000;56:337-44.
27. Harrell Jr FE. rms: Regression Modeling Strategies. R package version 5.0-0. CRAN. 2016.
28. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. Nature Methods. 2015;12:453-7.
29. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, et al. Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of

Response to Checkpoint Blockade. Cell Reports. 2017;18:248-62.

30. Liu L, Bai X, Wang J, Tang XR, Wu DH, Du SS, et al. Combination of TMB and CNA stratifies prognostic and predictive responses to immunotherapy across metastatic cancer. Clinical Cancer Research. 2019;25:7413-23.
31. National Cancer Institute. Genomic Data Commons Data Portal. Harmonized Cancer Datasets. 2019.
32. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: Efficient and comprehensive analysis of somatic variants in cancer. Genome Research. 2018;28:1747-56.
33. Yang S, Wu Y, Deng Y, Zhou L, Yang P, Zheng Y, et al. Identification of a prognostic immune signature for cervical cancer to predict survival and response to immune checkpoint inhibitors. OncoImmunology. 2019;8:e1659094.
34. Berraondo P, Minute L, Ajona D, Corrales L, Melero I, Pio R. Innate immune mediators in cancer: between defense and resistance. Immunological Reviews. 2016;274:290-306.
35. Elola MT, Ferragut F, Méndez-Huergo SP, Croci DO, Bracalente C, Rabinovich GA. Galectins: Multitask signaling molecules linking fibroblast, endothelial and immune cell programs in the tumor microenvironment. Cellular Immunology. 2018;333:34-45.
36. Gardner A, Ruffell B. Dendritic Cells and Cancer Immunity. Trends in Immunology. 2016;37:855-65.
37. Cariani E, Missale G. Immune landscape of hepatocellular carcinoma microenvironment: Implications for prognosis and therapeutic applications. Liver International. 2019;39:1608-21.
38. Banerjee K, Kumar S, Ross KA, Gautam S, Poelaert B, Nasser MW, et al. Emerging trends in the immunotherapy of pancreatic cancer. Cancer Letters. 2018;417:35-46.
39. Sanmamed MF, Chen L. A Paradigm Shift in Cancer Immunotherapy: From Enhancement to Normalization. Cell. 2018;175:313-26.
40. Jiang Y, Han QJ, Zhang J. Hepatocellular carcinoma: Mechanisms of progression and immunotherapy. World Journal of Gastroenterology. 2019;25:3151-67.
41. Zhao QJ, Zhang J, Xu L, Liu FF. Identification of a five-long non-coding RNA signature to improve the prognosis prediction for patients with hepatocellular carcinoma. World Journal of Gastroenterology. 2018;24:3426-39.
42. Bing Z, Tian J, Zhang J, Li X, Wang X, Yang K. An Integrative Model of miRNA and mRNA Expression Signature for Patients of Breast Invasive Carcinoma with Radiotherapy Prognosis. Cancer Biotherapy and Radiopharmaceuticals. 2016;31:253-60.
43. Jiang X, Hao Y. Analysis of expression profile data identifies key genes and pathways in hepatocellular carcinoma. Oncology Letters. 2018;15:2625-30.
44. Cheng J, Xie HY, Xu X, Wu J, Wei X, Su R, et al. NDRG1 as a biomarker for metastasis, recurrence and of poor prognosis in hepatocellular carcinoma. Cancer Letters. 2011;310:35-45.
45. Lian YF, Huang YL, Zhang YJ, Chen DM, Wang JL, Wei H, et al. CacYBP enhances cytoplasmic retention of p27Kip1 to promote hepatocellular carcinoma progression in the absence of RNF41 mediated degradation. Theranostics. 2019;9:8392-408.

46. Liu M, Li Y, Chen L, Man Chan TH, Song Y, Fu L, et al. Allele-specific imbalance of oxidative stress-induced growth inhibitor 1 associates with progression of hepatocellular carcinoma. Gastroenterology. 2014;146:1084-96.
47. Lv J, Zhang S, Wu H, Lu J, Lu Y, Wang F, et al. Deubiquitinase PSMD14 enhances hepatocellular carcinoma growth and metastasis by stabilizing GRB2. Cancer Letters. 2020;469:22-34.
48. Ma C, Kesarwala AH, Eggert T, Medina-Echeverz J, Kleiner DE, Jin P, et al. NAFLD causes selective CD4+ T lymphocyte loss and promotes hepatocarcinogenesis. Nature. 2016;531:253-7.
49. Brown ZJ, Fu Q, Ma C, Kruhlak M, Zhang H, Luo J, et al. Carnitine palmitoyltransferase gene upregulation by linoleic acid induces CD4+ T cell apoptosis promoting HCC development. Cell Death and Disease. 2018;9:620.
50. Choi C, Yoo GS, Cho WK, Park HC. Optimizing radiotherapy with immune checkpoint blockade in hepatocellular carcinoma. World Journal of Gastroenterology. 2019;25:2416-29.
51. Longo V, Brunetti O, Gnoni A, Licchetta A, Delcuratolo S, Memeo R, et al. Emerging role of immune checkpoint inhibitors in hepatocellular carcinoma. Medicina (Lithuania). 2019;55:698.
52. Ying HQ, Deng QW, He BS, Pan YQ, Wang F, Sun HL, et al. The prognostic value of preoperative NLR, d-NLR, PLR and LMR for predicting clinical outcome in surgical colorectal cancer patients. Medical Oncology. 2014;31:305.
53. Zhang G, Wu Y, Zhang J, Fang Z, Liu Z, Xu Z, et al. Nomograms for predicting long-term overall survival and disease-specific survival of patients with clear cell renal cell carcinoma. OncoTargets and Therapy. 2018;11:5535-44.

Supplementary Files

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

- supplementarymaterials.docx
- supplementarymaterials.docx
- LIHCmanuscriptPDFversion.pdf
- LIHCmanuscriptPDFversion.pdf