Integrative Characterization of Immune-relevant Genes in Hepatocellular Carcinoma

Wei-Feng Hong1#, Yu-Jun Gu2#, Na Wang3#, Jie Xia3, Heng-Yu Zhou14, Ke Zhan5, Ming-Xiang Cheng6 and Ying Cai3,7,*

1Department of Medical Imaging, The First Affiliated Hospital of Guangdong Pharmaceutical University, Guangzhou, Guangdong, China; 2Department of Ultrasound Medicine, The First Affiliated Hospital of Sun Yat-Sen University, Guangzhou, Guangdong, China; 3Key Laboratory of Molecular Biology for Infectious Diseases (Ministry of Education), Institute for Viral Hepatitis, Department of Infectious Diseases, The Second Affiliated Hospital, Chongqing Medical University, Chongqing, China; 4College of Nursing, Chongqing Medical University, Chongqing, China; 5Department of Gastroenterology, The Second Affiliated Hospital, Chongqing Medical University, Chongqing, China; 6Department of Hepatobiliary Surgery, The Second Affiliated Hospital, Chongqing Medical University, Chongqing, China; 7Department of Intensive Care Medicine, The Second Affiliated Hospital, Chongqing Medical University, Chongqing, China

Received: 27 November 2020 | Revised: 18 February 2021 | Accepted: 21 February 2021 | Published: 8 March 2021

Abstract

Background and Aims: Tumor microenvironment plays an essential role in cancer development and progression. Cancer immunotherapy has become a promising approach for the treatment of hepatocellular carcinoma (HCC). We aimed to analyze the HCC immune microenvironment characteristics to identify immune-related genetic changes. Methods: Key immune-relevant genes (KIRGs) were obtained through integrating the differentially expressed genes of The Cancer Genome Atlas, immune genes from the Immunoology Database and Analysis Portal, and immune differentially expressed genes determined by single-sample gene set enrichment analysis scores. Cox regression analysis was performed to mine therapeutic target genes. A regulatory network based on KIRGs, transcription factors, and immune-related long non-coding RNAs (IRLncRNAs) was also generated. The outcomes of risk score model were validated in a testing cohort and in clinical samples using tissue immunohistochemistry staining. Correlation analysis between risk score and immune checkpoint genes and immune cell infiltration were investigated. Results: In total, we identified 21 KIRGs, including programmed cell death-1 (PD-1) and cytotoxic T-lymphocyte associated protein 4 (CTLA4), and found IKBKE, IL2RG, EDNRA, and IGHA1 may be equally important to PD-1 or CTLA4. Meanwhile, KIRGs, various transcription factors, and IRLncRNAs were integrated to reveal that the NRF1-AC127024.5-IKBKE axis might be involved in tumor immunity regulation. Furthermore, the immune-related risk score model was established according to KIRGs and key IRLncRNAs, and verified more obvious discriminating power in the testing cohort. Correlation analysis indicated TNFSF4, LGALS9, KIAA1429, JDD2, and CD276 were closely related to the risk score, and CD4 T cells, macrophages, and neutrophils were the primary immune infiltration cell types. Conclusions: Our results highlight the importance of immune genes in the HCC microenvironment and further unravel the underlying molecular mechanisms in the development of HCC.

Citation of this article: Hong WF, Gu YJ, Wang N, Xia J, Zhou HY, Zhan K, et al. Integrative characterization of immune-relevant genes in hepatocellular carcinoma. J Clin Transl Hepatol 2021;9(3):301–314. doi: 10.14218/JCTH.2020.00132.

Introduction

Liver cancer is the most common malignancy and the fourth leading cause of cancer-related mortality worldwide.1 Hepatocellular carcinoma (HCC), a predominant type of primary liver cancer, has become a major public health problem. Although surgical resection is a potentially curative modality for a minority of early-stage HCC patients, as many as 70% of these patients will experience disease recurrence within 5 years.2 Due to the occult onset of HCC, most advanced patients are not eligible for the timely administration of effective treatment.4 Sorafenib, as a multi-kinase
inhibitor, has been approved by the Food and Drug Administration and recommended as the first-line treatment in this population based on the “SHARP” trial, with median overall survival of 6.5 months.3,5,6 Then, the “REFLECT” trial demonstrated that lenvatinib was non-inferior to sorafenib for overall survival in untreated advanced HCC patients.7 It only showed some clinical benefits for secondary endpoints in progression-free survival, time to progression, and objective response rate. The current status of HCC recurrence and metastasis are not optimistic; therefore, novel and effective treatment options are desperately in need of further exploration to decrease recurrence rates.

Increasing evidence has shown that the tumor microenvironment plays an essential role in cancer development and progression. With improved understanding of biological interactions within the tumor microenvironment, immune system and tumor cells, cancer immunotherapy has appeared to provide tremendous promise as a cancer treatment modality in recent years. Typically, in the liver, large quantities of innate and adaptive immune cells play a critical role in immune surveillance to detect and eliminate pathogens and participate in immune response and regulation of host defenses.8 However, the inflammatory state, due to risk factors that contribute to HCC, such as chronic infection with hepatitis B virus or hepatitis C virus, will change the tumor microenvironment and facilitate evasion of immune surveillance, leading to tumor tolerance and promoting the development of HCC.8 Hence, targeted immunotherapy is actively researched with the goal of inhibiting aberrant oncogenic pathways and improving prognosis.

At present, immune checkpoint inhibitors are considered one of the immunotherapeutic agents for rapid development to promote immune reconstitution and restore immune cell function.9 Programmed cell death-1 (PD-1) and cytotoxic T-lymphocyte associated protein 4 (CTLA4) blockade therapies have become promising approaches for the treatment of HCC.10–13 Unfortunately, in some studies, overall survival and improved recurrence-free survival did not achieve the pre-defined statistical significance criteria.14 It has been suggested that the HCC microenvironment can form a potent immune tolerance system, which greatly hinders the efficacy of immune checkpoint therapy. Therefore, remodeling the immune tolerant microenvironment of HCC could be of great significance for HCC immunotherapy.

Given the complexity of immunotherapy and tumor heterogeneity, extensive genomic analysis could provide clinical grounds for rapid development to promote immune reconstitution and restore immune cell function.9 Programmed cell death-1 (PD-1) and cytotoxic T-lymphocyte associated protein 4 (CTLA4) blockage therapies have become promising approaches for the treatment of HCC.10–13 Unfortunately, in some studies, overall survival and improved recurrence-free survival did not achieve the pre-defined statistical significance criteria.14 It has been suggested that the HCC microenvironment can form a potent immune tolerance system, which greatly hinders the efficacy of immune checkpoint therapy. Therefore, remodeling the immune tolerant microenvironment of HCC could be of great significance for HCC immunotherapy.

Given the complexity of immunotherapy and tumor heterogeneity, extensive genomic analysis could provide clinical grounds for rapid development to promote immune reconstitution and restore immune cell function.9 Programmed cell death-1 (PD-1) and cytotoxic T-lymphocyte associated protein 4 (CTLA4) blockage therapies have become promising approaches for the treatment of HCC.10–13 Unfortunately, in some studies, overall survival and improved recurrence-free survival did not achieve the pre-defined statistical significance criteria.14 It has been suggested that the HCC microenvironment can form a potent immune tolerance system, which greatly hinders the efficacy of immune checkpoint therapy. Therefore, remodeling the immune tolerant microenvironment of HCC could be of great significance for HCC immunotherapy.

**Methods**

**Datasets acquisition and pre-processing**

Fragments per kilobase million upper quartile RNA-Seq gene expression profile of liver hepatocellular carcinoma (LIHC) were downloaded from The Cancer Genome Atlas (TCGA) (https://portal.gdc.cancer.gov/), including 50 normal tissue and 374 primary tumor samples.15 A list of 1,811 immune genes (IGs), including 17 immune categories according to different molecular function, were obtained from the Immunology Database and Analysis Portal (ImmPort) (https://www.immpor.org/home)16 after eliminating reiterated genes. One of the major collections (C7: immune signatures) in the molecular signatures database (MSigDB) of Gene Set Enrichment Analysis (GSEA) (https://www.gsea-msigdb.org/gsea/index.jsp)17 is a collection of 4,872 annotated gene sets that represent cell types, states, and perturbations within the immune system.18 The immunologic gene sets and 28 immune signaling pathways were also collected and processed for subsequent analyses. The data of tumor immune infiltration in TCGA-LIHC was downloaded from the tumor immune estimation resource (known as TIMER) (https://cistrome.shinyapps.io/timer/).19

**Screening immune-relevant genes (IRGs)**

To evaluate the enrichment scores of every sample on each immune-related term in RNA-Seq data of TCGA-LIHC, single-sample gene set enrichment analysis (ssGSEA) was used to generate ssGSEA-score20 implemented in the R package “GSVA” and “GSEABase”. Each score was corrected between 0 and 1. Patients were divided into a high-immune score (Immunity_H) and a low-immune score (Immunity_L) group using unsupervised hierarchical clustering of R package “sparcl”.

Differentially expressed genes (DEG) analysis was performed via the R package “Limma” based on the empirical Bayes method.21 Immune differentially expressed genes (IDEGs) were determined by Immunity_H to Immunity_L ratio, with cutoff criteria of absolute log2 fold-change (|log(2) FC|) >1.0 and false discovery rate <0.05. Significant DEGs of RNA-Seq data from TCGA-LIHC project were also identified by the same approach. Then, overlapping DEGs, IDEGs, and IGs were assessed and provided 77 interacting genes as the IRGs in the present study.

**Construction of weighted gene co-expression networks and identification of key IRGs**

Based on the expression of 77 IRGs with complete clinical data in tumor tissues from TCGA-LIHC project, weighted gene co-expression networks analysis (WGCNA) was carried out to create expression modules and analyze the correlation of each module with immune traits (ImmuneGroup, StromalScore, ImmuneScore, ESTIMATEScore, and TumorPurity) using the R package WGCNA by hierarchical clustering of adjacency-based dissimilarity.22 Modules and eigengenes were defined as the first principle component of each gene module and regarded as representative of genes in each module. Gene significance was calculated to measure the Pearson correlation between gene expression and sample traits and to identify the significance of each module. We selected higher gene significance and defined survival-related modules according to p≤0.001 for further analysis. A scale-free topology fit index (scale-free R2) >0.95 was implemented to verify the soft threshold power and maintain optimal mean connectivity. A dynamic hybrid cut method, using a bottom-up algorithm, was applied to identify crucial modules, with cut height-off of 0.25. At last, key immune-relevant genes (KIRGs) in crucial modules were selected by those with module membership >0.5. General characteristics of KIRGs were analyzed using GSCALite (http://bioinfo.life.hust.edu.cn/web/GSCALite/).23

**Construction of co-expression network based on KIRGs**

To analyze the function of KIRGs, the relationship between transcription factors and KIRGs was explored. The transcription factors set was extracted from the Cistrome Project (http://www.cistrome.org/)24 and subjected to expression
differential analysis by the R package "Limma", with cut-off value of $|\log FC| > 1.0$ and adjusted $p$-value of $< 0.05$. The significant correlations between the differentially expressed transcription factors (DETFs) and KIRGs (DETFs-KIRGs) were calculated by Pearson's test, with $p$-value of $<0.0001$ and correlation coefficient of $>0.30$, which was visualized by Cytoscape software (3.7.1) in the appropriate type of correlation network.

Meanwhile, considering immune-related long non-coding RNAs (IRLncRNAs) involved in the mRNA transcript process of KIRGs, the network of IRLncRNAs-KIRGs was established using the co-expression analysis approach as described above. The profile of long non-coding RNAs was taken from the annotation data of TCGA-LIHC. Subsequently, the network of DETFs and IRLncRNAs (DETFs-IRLncRNAs) was also generated (correlation coefficient $>0.30$ and $p<0.0001$). Finally, the DETFs-IRLncRNA-KIRGs regulatory network was integrated for the pairs of the above three networks' data.

**Risk score model construction and verification**

TCGA-LIHC patients were randomly divided at the ratio of 1:1 into two cohorts (training cohort and testing cohort). To avoid overfitting, least absolute shrinkage and selection operator (Lasso) Cox regression was applied to eliminate genes generated from univariate Cox analysis of all genes of IRLncRNAs-KIRGs. Then, we performed multivariate Cox regression to construct a risk score model, and the performance was evaluated in the testing cohort. Kaplan-Meier survival analysis and time-dependent receiver operating characteristic (commonly known as ROC) analysis were used to compare the survival states and evaluate the accuracy of the risk score model.

**Enrichment analysis and protein-protein interaction**

To determine the function and potential regulatory pathways of identified genes, GSEA was explored for biological functions and pathways using the R package "clusterprofiler" and "GSVA", with adjusted $p$-value of $<0.05$ and q-value of $<0.05$, as each was regarded as statistically significant. Gene Ontology analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis were performed by the web tool "MetaScape" (http://metascape.org/). The search tool for the retrieval of interacting genes/proteins database (commonly known as STRING) was utilized to assess protein-protein interactions, with information minimum required interaction score: medium confidence (0.40) (www.string-db.org/).

**Experimental section**

Paired tumor and peritumor tissues were obtained from patients diagnosed with HCC and who had undergone surgery at the Department of Hepatological Surgery of the Second Affiliated Hospital of Chongqing Medical University (Chongqing, China). All experiments were performed in accordance with the Declaration of Helsinki of the World Medical Association and were approved by the ethics committee of the Second Affiliated Hospital of Chongqing Medical University (No. 2020-004). All participants provided written informed consent.

Tissue proteins were lysed using radio immunoprecipitation assay lysis buffer (CWBio, Beijing, China) with protease inhibitor cocktail (CWBio). Protein concentration was measured with a bicinchoninic acid assay (CWBio). The same amount of total proteins was subjected to SDS-PAGE, followed by a standard western blotting procedure with the primary antibodies and secondary antibodies (Supplementary Table 1), and detection using an enhanced chemiluminescence system (Advanta, Menlo Park, CA, USA).

Tissue paraffin sections were prepared for immunohistochemistry staining, according to standard procedures. The slides were visualized using 3,3′-diaminobenzidine substrate kit (Abcam, Cambridge UK) and counterstained with hematoxylin (Solarbio, Beijing, China). Stained slides were scanned by the Motic Easyscanner (Motic, Hong Kong, China) and the images were captured with Motic DS Assistant Lite (Motic VM V1 Viewer 2.0).

Tissue total RNA was extracted using TRIzol reagent (Invitrogen, Carlsbad, CA, USA), following the manufacturer's instructions. The cDNA library was constructed by reverse transcription using a commercial reverse-transcription kit (Takara, Shiga, Japan) according to the manufacturer's protocol. The cDNA was amplified using SYBR Green Master Mix (Bimake, Houston, TX, USA) and normalized by GAPDH. Quantitative polymerase chain reaction primers (Supplementary Table 2) were synthesized by Invitrogen.

**Statistical analysis**

All statistical analyses were performed in R version 3.6.0 and GraphPad Prism software 8.0, and $p$-values $<0.05$ were considered significant. Continuous variables, such as difference analysis, were performed by t-test or Wilcoxon’s test, according to the normality of data distribution. Correlation analysis of categorical variables was carried out with the chi-squared test. Correlations between continuous variables were estimated using the Pearson correlation, while Spearman’s correlation was used to assess the association between non-normally distributed data.

**Results**

**Initial screening IRGs**

To screen significant immune biomarkers, we incorporated important databases in the initial workflow (Fig. 1A). A total of 7,667 DEGs were identified from TCGA-LIHC dataset (Supplementary Fig. 1A, B). To understand each tumor sample’s immune status, the corresponding enrichment scores on every immune-related term were calculated based on the ssGSEA method. By unsupervised hierarchical clustering, patients were clustered into the Immunity_H group ($n=170$) and the Immunity_L group ($n=204$) (Supplementary Fig. 1C). Stromal cells and immune cells were the non-tumor components of the immune microenvironment which reflected tumor immune infiltration and tumor purity. The stromal scores and immune scores were calculated and combined to estimate scores in order to display tumor purity by the ESTIMATE algorithm (Fig. 2A; Supplementary Table 3). Upon comparison with the Immunity_L group, stromal scores, immune scores, and ESTIMATE scores were significantly higher in the Immunity_H group, while the lower level of tumor purity represented the low activity of tumor cells ($p<0.001$) (Fig. 2B–E). Then, enrichment analysis of biological functions and pathways in the two groups was performed using GSEA. The Immunity_H group was identified when the enrichment score was $>0$, and found to be mainly enriched in complement activation (classical pathway), humoral immune response mediated by circulating immunoglobulin and MHC class II protein complex in the aspect of biological function (Fig. 2F). Allograft rejection, intestinal immune network for IgA production, and primary immunodeficiency embodied the pathway...
enrichment results (Fig. 2G). Next, the bubble chart showed the top 10 biological functions in the Immunity_H group and the Immunity_L group (Supplementary Table 4). Meanwhile, 1,950 IDEGs were obtained by differential analysis for the Immunity_H and Immunity_L groups (Supplementary Fig. 1D, E). Finally, 77 IRGs were screened out by overlapping the DEGs, IDEGs, and IGs in this present study. Gene Ontology analysis and KEGG pathway enrichment analysis results were statistically confirmed via Metascape (Supplementary Fig. 1F, G).

Identification of the modules associated with immune traits by construction of WGCNA

WGCNA was applied to find important modules most associated with immune traits based on 77 IRGs. A total of seven modules were identified by setting soft threshold power and cut height-off value. IRGs in the gray module were identified as non-clustering genes (Fig. 3A, B). In the heatmap of module-trait relationships, the blue module (12 IRGs) and brown module (9 IRGs) manifested significant correlation of immune traits, especially by ImmuneScore and StromalScore (correlation coefficient of >0.73, p<0.05) (Fig. 3C). Genes in these two modules with module membership of >0.5 were defined as KIRGs (Fig. 3C).

Functional analysis of the KIRGs signature

We analyzed the fundamental functional of the KIRGs (Fig. 1B). Differential analysis of KIRGs’ expressions was performed, with 11 genes (APOBEC3H, CD3D, CTLA4, CXCR3, EDNRA, inhibitor of nuclear factor kappa-B kinase subunit epsilon (IKBKE), IL2RG, LTA, LTBPI2, PDCD1, and SYTL1) showing high expression and 9 genes (CD244, COLLEC10, CXCL12, FOS, GDF2, IGHA1, IGHA2, MARCO, NGFR, and THBS1) showing low expression in tumor tissue (Fig. 4A). Correlation analysis of 21 KIRGs showed that CXCR3 was highly associated with CD3D and LTA (correlation coefficient of >0.80, p<0.05) (Fig. 4B). To display interactive relationships among proteins of KIRGs, a protein-protein interaction network was constructed by the STRING database. LTA, CTLA4, FOS, and CXCL12 had the most interactive lines (>10) in the bar graph, which totaled 60 lines (Fig. 4C, D). High node degrees could indicate an essential role in tumor immune processes.

Gene Ontology analysis and KEGG pathway enrichment analysis were performed using the R package “clusterprofiler”. To demonstrate more information from the results, the top 20 Gene Ontology terms were enriched as a Gene Ontology heatmap, mainly in “leukocyte migration” and “humoral immune response” (Fig. 4E). We also found that the three terms of “response to oxygen levels”, “response to hypoxia” and “response to decreased oxygen levels” enriched from the same genes (Fig. 4E). In the KEGG circle map, KIRGs were enriched in some common and critical pathways referring to tumorigenesis and progression, such as “T cell receptor signaling pathway”, “Cytokine-cytokine receptor interaction” and “PI3K-Akt signaling pathway” and so on (Fig. 4F).

To further verify whether KIRGs were closely related to immune function, 29 terms from the ImmPort database served as the candidate profile and a gene set including 21 KIRGs was found to be enriched in eight immune
Fig. 2. Microenvironment signatures of TCGA-LIHC cohort and enrichment analysis. (A) Heatmaps showed expression profiles for microenvironment signatures of stromal scores, immune scores, ESTIMATE scores, and tumor purity with unsupervised hierarchical clustering analyses. (B–E) Stromal scores, immune scores, ESTIMATE scores and tumor purity in the Immunity_H group and the Immunity_L group. (***p<0.001) (F–G) Enrichment analysis of biological functions and pathways in the Immunity_H group (top 5). (H) Bubble chart of top 10 biological functions in the Immunity_H and Immunity_L groups. (I) Heatmap of Gene Ontology and KEGG enrichment analysis for IRGs via Metascape.
terms, namely Check-point, Th1 cells, Tfh, T cell coinhibition, plasmacytoid dendritic cells, CCR, TIL, and regulatory T cells, based on the ssGSEA-score of the R package "GSVA" (Supplementary Fig. 2A–C). Next, we extracted all the genes of some critical immune terms from TCGA-LHC project and differential analysis of expression was implemented between the Immunity_H group and the Immunity_L group (Supplementary Fig. 3A–E). Besides, due to multiple filters, IKBKE, IL2RG, EDNRA, and IGH A1 were screened out more significantly with p<0.2 using...
Hong W.F. et al: Immune markers in hepatocellular carcinoma

Univariate Cox regression and multivariate Cox regression of 21 KIRGs (Fig. 5A, B). We examined protein expression of these four markers in clinical HCC and corresponding peritumor tissues. Western blotting indicated that IKBKE, IL2RG, and EDNRA were highly expressed in most HCC tissue specimens, while IGHA1 had low expression in tumors (Fig. 5C). Immunohistochemical assay showed IKBKE and IL2RG were significantly higher in representative HCC tissue (Fig. 5D).

Analysis of KIRGs in a web-based platform of GS-CALite

To understand gene set mutations, we observed single nucleotide variation frequency and analyzed the variant types.
Of the 21 KIRGs, the vast majority of variants occurred as single nucleotide polymorphisms and missense mutations in the included samples (Supplementary Fig. 4A, B). The top 10 frequently mutated genes were GDF2, LTBP2,
Further, differential methylation between tumor and paired normal tissues were found, along with that of other tumor types (Supplementary Fig. 4G). PDCD1, CTLA4, and MARCO indicated high levels of methylation in HCC tissues, while CXCL12 and EDNRA showed a high level in paired normal tissues. We also observed the relationships between methylation levels and corresponding gene expression, which displayed that PDCD1 methylation had a positive correlation with expression, whereas IKBKE and EDNRA methylation showed negative relationships in most tumors (Supplementary Fig. 4H).

We also investigated the difference of gene expression between pathway activity groups (activation and inhibition). These KIRGs were also involved in apoptosis, cell cycle, DNA damage response, epithelial to mesenchymal transition, the innate immune response, and MAPK signaling pathways (Supplementary Fig. 4I). PDLD1, CTLA4, IKBKE, and IL2RG could promote apoptosis and activate DNA damage response and epithelial to mesenchymal transition (Supplementary Fig. 4J). Eventually, drug sensitivity analysis was performed for 21 KIRGs that potentially participate in HCC pathogenesis (Supplementary Fig. 4F).

Construction of the DFTFs-IRLncRNA-KIRGs network

Subsequently, in order to observe transcription factors involved in the regulation of KIRGs, the correlations of KIRGs to DFTFs were generated according to co-expression analysis and visualized as a regulatory network (Fig. 1C). A total of 162 positive or negative correlation pairs of DETFs-KIRGs were found; among these, the pairs of FOS-ERG1 (correlation coefficient of 0.63, \( p = 5.2E-43 \)), CIITA-CTLA4 (correlation coefficient of 0.62, \( p = 6.3E-43 \)) and CXCR3-CIITA (correlation coefficient of 0.6, \( p = 9.1E-43 \)) showed very high correlation (Fig. 5E). Similarly, the IRLncRNAs-KIRGs network and DETFs-IRLncRNAs network were also constructed. Eventually, combining the three networks, the DFTFs-IRLncRNAs-KIRGs regulatory network was built, comprised of 103 DETFs, 175 LncRNAs, and 15 KIRGs, to elucidate underlying regulatory mechanisms (Fig. 5F). Most of the pairs represented positive correlations within the network.

IKBKE, having a prominent role among the KIRGs, was related to numerous IRLncRNAs, most obviously AC127024.5. Interestingly, in the network, NRF1 was regarded as the most relevant transcription factor to AC127024.5 (correlation coefficient of 0.63, \( p = 4.9E-43 \)) which indicated that the NRF1-AC127024.5-IKBKE axis might be involved in regulating many biological processes. We also explored the NRF1-AC127024.5-IKBKE axis for immune cell infiltration in TCGA-LIHC patients from the TIMER database, and found an involvement in the vast majority of immune cell infiltration, especially those related to B cells, neutrophils and macrophages (\( p < 0.01 \)) (Supplementary Fig. 7).

Furthermore, univariate and Lasso Cox regressions were implemented to optimize the parameters for screening risk genes among the KIRGs and IRLncRNAs (Fig. 6A, B). In total, IL2RG and eight key IRLncRNAs (LINC01871, AC145207.5, LINC0294, LINC01138, AL031985.3, AC083799.1, SNHG1, and SNHG3) were obtained by multivariate Cox regression in the training cohort. The risk score model was constructed and verified the performance in the testing cohort. Patients in each cohort were divided into a low risk group and a high risk group, according to the median risk score of the training cohort. Survival status, risk scores, and expression patterns of each patient were reflected in Fig. 6C and 6D. Kaplan-Meier curve analysis showed that the low risk group had a significantly better prognosis than that of the high risk group in the training cohort (\( p = 4.70E-06 \)) and the testing cohort (\( p = 4.70E-05 \)), respectively (Fig. 6E, G). Time-dependent ROC curve analysis of the risk score model in the training and testing cohorts all determined good predictive accuracy with area under the curve values of 0.826 and 0.724 for 1-year survival, and 0.822 and 0.736 for 3-year survival, respectively (Fig. 6F, H). Therefore, it appeared that the risk score model successfully stratified HCC from TCGA.

In addition, we found that the risk score model was associated with clinicopathological parameters, such as tumor-stage, clinical stage, and survival state by chi-squared test (\( p < 0.01 \)). Further analysis of univariate and multivariate Cox regression revealed that risk score could be an independent prognostic indicator for HCC patients (\( p < 0.001 \)) (Fig. 7B, C). Representative markers (IL2RG, SNHG1, SNHG3, and LINC01138) showed obvious high expression in HCC samples at the mRNA level (Fig. 7D). Meanwhile, the forecast performance of the risk score model was superior to those models based on tumor-stage, node-stage, metastasis-stage, grade, American Joint Committee on Cancer stage, KIRGs and key IRLncRNAs (Supplementary Fig. 8A–G). The multiplex model seemed a dependable indicator for predicting the prognosis of HCC patients.

We analyzed the expression of each immune checkpoint gene of ImmPort for correlation with the integration of risk score resulting from the two cohorts. Pearson’s correlation indicated TNFSF4, LGALS9, KIAA1429, IDO2, and CD276 were closely related to risk score (\( p < 0.05 \)), shown as a circle map (Fig. 8A). Ultimately, we investigated the measurement of the integrated risk score for immune cell infiltration in TCGA-LIHC patients from the TIMER database and observed that effects of risk score appeared to be concentrated among the CD4 T cells, macrophages, and neutrophils (\( p < 0.05 \)), while the B cells, CD8 T cells and dendritic cells did not show significant correlation (Fig. 8B–G).

Discussion

Rapidly emerged immunotherapy has demonstrated increasing promise in the application of treatment for human cancers. On account of tumor complexity and heterogeneity, only a minority of patients could have benefitted from immunotherapy. Interestingly, the tumor immune microenvironment is closely related to patients’ responsiveness after receiving the therapy of immune-checkpoint blockade. Thus, understanding the tumor immune microenvironment’s diversity will likely uncover novel biomarkers and provide effective therapeutic targets.

According to the natural and fundamental immunological properties of the liver and the current dilemma of immunotherapy for HCC, in this study, we systematically identified 21 KIRGs that potentially participate in HCC pathogenesis and progression. To obtain more robust and reliable IGs, our results were analyzed by more comprehensive and
Fig. 6. Construction of risk score model based on DETFs-IRLncRNAs-KIRGs network. (A–B) KIRGs and IRLncRNAs were narrowed down by the Lasso algorithm. (C) Survival status scatter plots, risk score distribution, and expression patterns in the training cohort. (D) Survival status scatter plots, risk score distribution, and expression patterns in the testing cohort. (**p < 0.001). (E) Kaplan-Meier curve analysis of the high risk and low risk groups in the training cohort. (F) Time-dependent ROC curve analysis of the risk score model in the training cohort. (G) Kaplan-Meier curve analysis of the high risk and low risk groups in the testing cohort. (H) Time-dependent ROC curve analysis of the risk score model in the testing cohort.
Fig. 7. Relationship between risk score model and clinicopathological parameters. (A) Expression profiles of risk genes and association with clinicopathological characteristics. (B–C) Univariate and multivariate Cox regression analyses of risk score and clinicopathological characteristics. (D) Representative key IRLcRNAs were detected by real-time polymerase chain reaction of HCC and paired peritumor tissues from 18 patients. Abbreviations: M, metastasis; N, node; T, tumor.
strict screening methods on the basis of DEGs and IDEGs in TCGA database, and IGs retrieved from ImmPort. In its methodology, ssGSEA is a widely accepted algorithm to analyze statistical enrichment. The ESTIMATE algorithm was used to calculate the stromal score, immune score and estimate score in the immune microenvironment for further immune assessment. The Lasso algorithm, as one of the regression regularization methods, applies a Lasso Cox regression model and ensures optimal risk model, even when variables in the dataset have high dimensions and multicolinearity.

As is known, binding of ligand and inhibitory receptors on immune cells will weaken the T cell mediated immune response. Antibodies against immune checkpoint proteins, such as CTLA4 or PD-1 (also called PDCD1), the first generation of antibody-based immunotherapy, has been implemented for the treatment of HCC patients. In consideration of partial immune response to these inhibitors, early results of a recent study based on the CheckMate-040 trial (nivolumab plus ipilimumab) were obtained for patients of advanced HCC who previously received sorafenib, and demonstrated an objective response rate of 33% (95% confidence interval of 20–48) in these patients. The Food and Drug Administration has accelerated the approval of this combination therapy strategy. Similar treatment regimens emerged for melanoma, non-small-cell lung cancer, and renal cell carcinoma. Evidently, our results showed reliability and high practical utility from a clinical perspective.

Meanwhile, it is also imperative to develop novel immune therapeutic approaches. In the present study, we found IKBKE, IL2RG, EDNRA, and IGH A1 to be statistically more significant with \( p < 0.2 \) than PDCD1 (\( p > 0.2 \)) among the 21 KIRGs by univariate Cox regression and multivariate Cox regression. Due to multi-step screening, we boldly defined the critical value as 0.2. IKBKE, a member of the nonclassical IKK family, is considered to be a potential target for

---

**Fig. 8.** Correlation analysis between risk score and immune checkpoint gene and immune cell infiltration. (A) Circle map of the relationships between immune checkpoint genes and integration of risk score. (B–G) Risk score for immune cell infiltration of CD4 T cells, macrophages, CD8 T cells, B cells, dendritic cells, and neutrophils.
cancer therapy. Typically, IKBKE was classified as an on-tumor network to reveal the possible functional relationship. DETFs, and constructed a DETFs-IRlncRNAs-KIRGs regulatory network. We identified IRLncRNAs associated with KIRGs and their roles as crucial components of cancer regulatory networks. Implementation of combination therapy with immune also presents several obstacles and faces many challenges in clinical practice. This field of cancer immunotherapy still lacks a deep understanding.

In summary, our analysis results highlight the importance of IGS in the HCC microenvironment. Moreover, sufficient information on novel biomarkers, networks, and pathways further unravels the underlying molecular mechanisms in the development of HCC. Understanding the immune microenvironment signatures will be advantageous to provide permissive justification to improve the clinical efficacy of the immunotherapy.

Conclusions

In summary, our analysis results highlight the importance of IGs in the HCC microenvironment. Moreover, sufficient information on novel biomarkers, networks, and pathways further unravels the underlying molecular mechanisms in the development of HCC. Understanding the immune microenvironment signatures will be advantageous to provide permissive justification to improve the clinical efficacy of the immunotherapy.

Acknowledgments

We sincerely appreciate the guidance of Xinyuan Lao and Zhujun Tan of Hellixlife, and the invaluable assistance of Huimei Wang, Yuanning Guo, and Liewang Qiu during the current study.

Funding

This work was supported by the China Postdoctoral Science Foundation (grant number 2019M663445 to YC), National Natural Science Foundation of China (grant number 81602045 to JX, 81802454 to HYZ), Hong W.F. et al: Immune markers in hepatocellular carcinoma

Conflict of interest

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

Author contributions

Study concept and design (YC, WFH), analysis and interpretation of data (WFH, YJG), sample collection (HYZ, MXC), experimental processes and drafting of the manuscript (KZ, NW), critical revision of the manuscript for important intellectual content (YC and JX). All authors have read and approved the final manuscript.

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 Cancer J Clin 2018;68:394–424. doi: 10.3332/caac.21492.
[2] Omata M, Cheng AL, Kokudo N, Kudo M, Lee JM, Jia J, et al. Asia-Pacific clinical practice guidelines on the management of hepatocellular carcinoma: a 2017 update. Hepatol Int 2017;11:317–370. doi: 10.1111/heon.12702-017-9799-9.
[3] Llovet JM, Zucman-Rossi J, Pikarsky E, Sangro B, Schwartz M, Sherman M, et al. Hepatocellular carcinoma. Nat Rev Dis Primers 2016;2:16018. doi: 10.1038/nrdp.2016.18.
[4] Hollebecque A, Malka D, Ferté C, Ducreux M, Boige V. Systemic treatment of advanced hepatocellular carcinoma: from disillusions to new horizons. Eur J Cancer 2015;51:327–339. doi: 10.1016/j.ejca.2014.12.005.
[5] Llovet JM, Ricci S, Mazzaferro V, Hilgard P, Gane E, Blanc JF, et al. Sorafenib in advanced hepatocellular carcinoma. N Engl J Med 2009;359:378–390. doi: 10.1056/NEJMoa0708857.
[6] Cheng AL, Kang YK, Chen Z, Tsao CJ, Qin S, Kim JS, et al. Efficacy and safety of sorafenib in patients in the Asia-Pacific region with advanced hepatocellular carcinoma: a phase III randomised, double-blind, placebo-controlled trial. Lancet Oncol 2009;10:25–34. doi: 10.1016/S1470-2045(08)70285-7.
[7] Kudo M, Finn RS, Qin S, Han KH, Ikeda K, Piscaglia F, et al. Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial. Lancet 2018;391:1163–1173. doi: 10.1016/S0140-6736(18)30207-1.
[8] Bayrak C, Habib N, Huang KW, Podde M, Warwick J, Arasadanin R, Immunological basis of genesis of hepatocellular carcinoma: Unique challenges and potential opportunities through immunomodulation. Vaccines (Basel) 2020;8:247. doi: 10.3390/vaccines8020247.
[9] Zhou G, Sprengers D, Boor PPC, Doukas M, Schutz H, Mancham S, et al. Antibodies against immune checkpoint molecules restore functions of tumor-infiltrating T cells in hepatocellular carcinomas. Gastroenterology 2017;153:1107–1119.e10. doi: 10.1053/j.gastro.2017.06.017.
[10] El-Khoeiry AB, Sangro B, You T, Crocetti TS, Kudo M, Hsu C, et al. Nivolumab in patients with advanced hepatocellular carcinoma (CheckMate 040): an open-label, non-comparative, phase 1/2 dose escalation and expansion trial. Lancet 2017;389:2492–2502. doi: 10.1016/S0140-6736(17)31046-2.
[11] Zhu AX, Finn RS, Edelme J, Catian S, Ogawa S, Palmer D, et al. Pembrolizumab in patients with advanced hepatocellular carcinoma previously treated with sorafenib (KEYNOTE-224): a non-randomised, open-label phase 2 trial. Lancet Oncol 2018;19:940–952. doi: 10.1016/S1470-2045(18)30351-6.
[12] Sangro B, Gomez-Martín C, de la Mata M, Illarramendi M, Garralda E, Barrera E, et al. A clinical trial of CTLA-4 blockade with tremelimumab in patients with hepatocellular carcinoma and chronic hepatitis C.. J Hepatol 2012;59:81–88. doi: 10.1016/j.jhep.2012.02.022.
[13] You T, Kang YK, Kim Y, El-Khoeiry AB, Santoro A, Sangro B, et al. Nivolumab (NIVO) + ipilimumab (IPI) combination therapy in patients (pts) with advanced hepatocellular carcinoma (HCC): Results from CheckMate 040. J Clin Oncol 2019;37:4012. doi: 10.1200/JCO.2019.37.15_suppl.4012.
[14] Finn RS, Ryoo BY, Merle P, Kudo M, Bouattour M, Lim HY, et al. Pembrolizumab as second-line therapy in patients with advanced hepatocellular carcinoma in KEYNOTE-240: A randomized, double-blind, phase III trial. J
