

Preprints are preliminary reports that have not undergone peer review. They should not be considered conclusive, used to inform clinical practice, or referenced by the media as validated information.

Identification and Validation a Costimulatory Molecule Gene Signature to Predict the Prognosis and Immunotherapy Response for Hepatocellular Carcinoma

Yinan Hu Xijing Hospital **Jingyi** Liu Xijing Hospital Jiahao Yu Xijing Hospital **Fangfang Yang** Xijing Hospital Miao Zhang Xijing Hospital Yansheng Liu Xijing Hospital Shuoyi Ma Xijing Hospital Xia Zhou Xijing Hospital Jingbo Wang Xijing Hospital Ying Han (V hanying1@fmmu.edu.cn) Xijing Hospital https://orcid.org/0000-0003-3046-9507

Primary research

Keywords: Hepatocellular carcinoma, costimulatory molecule, prognostic signature, ssGSEA, IPS

Posted Date: November 23rd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1064833/v1

License: (a) (b) This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License **Version of Record:** A version of this preprint was published at Cancer Cell International on February 22nd, 2022. See the published version at https://doi.org/10.1186/s12935-022-02514-0.

Abstract

Background: Hepatocellular carcinoma (HCC) is one of the most common malignancies worldwide. Costimulatory molecules have been proven to be the foundation of immunotherapy. However, the potential roles of costimulatory molecule genes (CMGs) in HCC remain unclear. Our study is aimed to develop a costimulatory molecule-related gene signature that could evaluate the prognosis of HCC patients.

Methods: Based on The Cancer Gene Atlas (TCGA) database, univariate Cox regression analysis was applied in CMGs to identify prognosis-related CMGs. Consensus clustering analysis was performed to stratify HCC patients into different subtypes and compared them in OS. Subsequently, the LASSO Cox regression analysis was performed to construct the CMGs-related prognostic signature and Kaplan–Meier survival curves as well as ROC curve were used to validate the predictive capability. Then we explored the correlations of the risk signature with tumor-infiltrating immune cells, tumor mutation burden (TMB) and response to immunotherapy. The expression levels of prognosis-related CMGs were validated in HCC using qRT-PCR method.

Results: All HCC patients were classified into two clusters based on 11 CMGs with prognosis values and cluster 2 correlated with a poorer prognosis. Next, a prognostic signature of six CMGs was constructed, which was an independent risk factor for HCC patients. Patients with low-risk score were associated with better prognosis. The correlation analysis showed that the risk signature could predict the infiltration of immune cells and immune status of the immune microenvironment in HCC. The qRT-PCR indicated six CMGs with significantly differential expression in HCC tissues and normal tissues.

Conclusion: In conclusion, our CMGs-related risk signature could be used as a prediction tool in survival assessment and immunotherapy for HCC patients.

Introduction

Primary liver cancer is an aggressive malignant tumor with high mortality worldwide(1). Hepatocellular carcinoma (HCC) is the most common histological subtype and the fourth leading cause of cancerrelated mortality, accounts for approximately 90% of all primary liver cancer. At present, the traditional treatment methods for HCC are systemic chemotherapy, local ablation and surgical resection(2). However, the therapeutic effect of these methods is away from satisfactory. In recent years, some clinical trials related immunotherapy showed different outcomes in improving the prognosis of HCC patients(3– 5). Therefore, it is urgently required to explore novel prognostic signature for HCC that can predict survival and the response to immunotherapy.

The component of immune microenvironment in HCC is the target for many therapeutic advances, including immunotherapy(6). Most recently, immunotherapies targeting the adaptive immune system, specifically, T cells, have improved tumor control(7). Activating T cells involves many signals, among which costimulatory molecules are important(8, 9). HCC could utilize immune checkpoint and evade anti-

tumor immune responses by expressing the corresponding costimulatory ligands(10). B7-CD28 superfamily is a pivotal signal in co-stimulation of T cell activation, and PD-1/PD-L1 also belong to it, which demonstrated the critical effect of costimulatory molecules in HCC(11, 12). Besides, accumulating evidence has shown that TNF superfamily, another costimulatory signals, plays a central role in cancer immune regulation(13). The OX40-OX40L axis, a member of the TNF superfamily, has been shown to improve anti-tumor effects of immune cells and effect for cancer immunotherapy(14–16). Previous studies also have shown that costimulatory molecules can regulate the tumor immune microenvironment(TME), mainly affecting the activation and proliferation of T cells(17). Thus, these molecules possibly could provide novel insights in TME. However, the functions of costimulatory molecules in HCC remain unclear.

In this systematic study, we evaluated the expression levels of costimulatory molecules genes in HCC tissues and normal tissues from The Cancer Genome Atlas (TCGA) database. Then a costimulatory molecules-related prognostic signature was constructed for HCC patients and we explored the associations between the prognostic signature and clinicopathological features. Furthermore, we also analyzed the potential roles of this prognostic signature in the immune microenvironment, tumor mutation analysis and response to immunotherapy in different subgroups.

Materials And Methods

1. Data collection

The transcriptomic data and corresponding clinical information of HCC were downloaded from the public The Cancer Genome Atlas (TCGA) data portal (https://portal.gdc.cancer.gov/). A total of 50 normal samples and 374 HCC samples were obtained. Patients with incomplete overall survival (OS) information were excluded. Subsequently, the TCGA cohort was randomly divided into training set (n=186) and test set (n=184). There were no significantly differences in clinical characteristics between two sets (Table 1). Furthermore, a total of sixty costimulatory molecules genes (CMGs) were collected from prior reviews(18, 19).

2. Identification of differentially expressed genes (DEGs)

We utilized "limma" package in R software (version 4.0.4) to identify the differentially expressed genes (DEGs) between all HCC specimens and normal specimens according to the criteria of P-value < 0.05 and llog2 (fold change) | > 1. The DEGs were notated with *** if p<0.001, ** if p<0.01 and * if p<0.05. A PPI network was constructed using the Search Tool for the Retrieval of Interacting Genes (STRING) database (http://www.string-db.org/) to explore the interactions between these DEGs.

3. Consensus clustering of prognosis-related CMGs

Univariate Cox regression analysis was performed to screen the CMGs with prognostic values in HCC with the cutoff value of P<0.05. To further elucidate the biological characteristics and prognostic values of CMGs, we employed the "ConsensusClusterPlus" package to cluster the HCC patients into different subgroups(20). Principal Component Analysis (PCA) was performed using R package to assess the distribution of gene expression among different subtypes. The OS difference between different clusters was verified by the Kaplan-Meier curves. Gene set enrichment analysis (GSEA) was conducted in gene set "h.all.v7.2.symbols.gmt" using Java GSEA software (version 4.1.0) to identify the potential biological processes among different clusters. An enrichment pathway with the normalized P < 0.05 and the false discovery rate (FDR) value < 0.05 were considered as statistically significant.

4. Construction of costimulatory molecule-related prognostic signature

Patients with HCC were randomly divided into a training set and a test set. The training set was used to construct a prognostic costimulatory molecule-related risk signature of HCC, and the test set and total set were used to validate the prognostic power of this risk signature. The least absolute shrinkage and selection operator (LASSO) penalized Cox proportional hazards regression was performed to narrow down the candidate genes and construct the risk model based on the prognosis-related costimulatory molecule genes using the R package "glmnet"(21). The penalty parameter (λ) was determined by the minimum criteria. The risk score was calculated with the following formular for each patient: Risk score = expression of gene 1 * coefficient 1 + expression of gene 2 * coefficient 2 + expression of gene 3 * coefficient 3 + ... + expression of gene n * coefficient n(22). Patients were divided into high- and low-risk groups with R package "survivalROC" to validate the prognostic capability. The Kaplan–Meier survival curves of the high- and low-risk groups were plotted using R package "survival" and "survminer" to demonstrate the OS of the patients.

5. Construction and validation of a nomogram

The nomogram and calibration curves were constructed with R package "rms". The consistency between the predicted and actual survival of the calibration curves was used to evaluate the accuracy of the nomogram. Meanwhile, the nomogram was examined using the ROC curves.

6. Functional enrichment analysis

HCC patients were stratified into high- and low-risk groups based on the median risk score. To explore the potential molecular mechanisms of the risk model genes, DEGs between the high- and low-risk groups were identified with the criteria of $|log2FC| \ge 1$ and FDR<0.05. Gene Ontology (GO) and Kyoto

Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were conducted using the "clusterProfler" package in R software according to the DEGs.

7. Assessment of immune cell infiltration

The Cell-type Identification by Estimating Relative Subsets of RNA Transcripts (CIBERSORT) analysis were performed to estimate the proportions of immune cells infiltration using R package "CIBERSORT" from RNA-sequencing data in TCGA(23). Wilcoxon rank-sum test was used to examine the differences of infiltrating immune cells in high- and low-risk groups. The tumor microenvironment score was calculated using R package "ESTIMATE"(24).

8. Mutation analysis

The mutation data for HCC patients were downloaded from the TCGA data portal (https://portal.gdc.cancer.gov/). Mutation data were further analyzed using the "maftools" package(25). We calculated the tumor mutation burden (TMB) score for each HCC patient as follows: (total mutation/total covered bases) ×10^6(26).

9. Immunophenoscore analysis

Immunophenoscore (IPS) could well predict the response of immune checkpoint inhibitors (ICIs). The immunogenicity is determined by four major categories of genes, including effector cells, major histocompatibility complex (MHC) molecules, immunomodulators and immunosuppressive cells. The IPS of a patient can be derived using machine learning without bias. The scores of IPS were calculated using a scale ranging from 0–10 based on representative cell type gene expression z-scores. The IPS of every HCC patient was obtained from The Cancer Immunome Atlas (TCIA) (https://tcia.at/home).

10. Verification of prognosis-related CMGs expression

Total RNA was extracted from tissue samples using Trizol reagent (Sigma, USA), and then, RNA was reverse transcribed into cDNA with the Evo M-MLV RT Premix (Accurate Biotechnology (Hunan) Co.,Ltd). Quantitative real-time PCR (qRT-PCR) analyses were performed by SYBR Green premix pro Taq HS qRT-PCR kit (Accurate Biotechnology (Hunan) Co.,Ltd) to validate gene expression, and the level of β -Actin served as an internal control. The relative expression was calculated based on the comparative Ct (2^{- $\Delta\Delta$ Ct) method. The primers' sequences for qRT-PCR are shown in Table 2.}

11. Tissue collection

Forty-three matched tumorous and non-tumorous tissue specimens of HCC were collected from The Xijing Hospital of Air Force Medical university during 2017-2018. The clinicopathological details are shown in Table 3. The research was approved by the Institutional Research Ethics Committees of the Xijing Hospital. Informed consent for publication was obtained from all patients for collection of tissue samples prior to the surgery.

Results

1. Identification of DEGs between normal and HCC tissues

The flowchart of this study was illustrated in Figure 1. The expression data of 59 CMGs, including 13 well-defined B7-CD28 family costimulatory molecules and 46 TNF family costimulatory molecules genes, were extracted from The Cancer Genome Atlas (TCGA) database after excluding TNFRSF6B for its low expression. The 59 costimulatory molecule-related genes expression levels were compared between HCC tumor and normal tissues, we identified 40 differentially expressed genes (DEGs) (P < 0.05). Among these DEGs, 11 genes (NGFR, TNFSF11, PDCD1LG2, CD274, TNFRSF1A, TNFRSF11B, TMIGD2, FAS, TNFRSF10D, TNFSF13 and CD86) were down-regulated while 29 genes (TNFRSF17, TNFRSF13B, CD276, TNFRSF12A, LTBR, TNFSF18, EDAR, TNFRSF14 , ICOSLG, RELT, CD28, ICOS, LTA, TNFRSF21, TNFRSF10C, VTCN1, TNFRSF11A, LTB, EDA2RC, TLA4, TNFSF9, TNFRSF25, PDCD1, CD70, TNFSF4, TNFRSF9, TNFRSF18, TNFSF15 and TNFRSF4) were up-regulated in tumor tissues (Figure 2A and Supplementary Table 1). The correlation among CMGs were analyzed. The relationships between each two of them were almost positively correlated, TNFRSF13C and TNFRSF13B were most correlated (Cor = 0.93) (Figure 2B). A protein-protein interaction (PPI) network was performed to further explore the interactions among these CMGs (Figure 2C). The minimum required interaction score was set at highest confidence 0.9. The resulted showed that TNF, CD28, CD40, CD80, CTLA4, LTA, TNFRSF10A and TNFSF13B were hub genes.

2. Consensus clustering of prognosis-related CMGs in HCC

A univariate Cox regression analysis was performed to primary selecting of the survival-related genes from 59 CMGs. A total of 11 CMGs were significantly linked to the prognosis of HCC patients (P < 0.05). Two genes were protective genes with hazard ratio (HR) < 1, while 9 genes were risk factors with HR > 1 among them (Figure 3A). To explore the associations between the expression of 11 prognosis-related CMGs and HCC subtypes, HCC patients were divided into two subgroups (Cluster 1: n=197, Cluster 2: n=173) according to consensus clustering analysis (Figure 3B and Supplementary Figure 1A). The PCA were analyzed to verify the reliability between different subgroups, and Cluster 1 and Cluster 2 could gather together and non-overlapped with each other (Figure 3C). We compared the OS between two subtypes to better understand the relationships between clustering results and survival outcomes, the Kaplan-Meier curves indicated that Cluster 1 had a better prognosis than Cluster 2 (P = 0.002, Figure 3D). The clinical features and two clusters were compared with a heatmap. The majority of 11 prognosisrelated CMGs had higher expression in Cluster 2. These two clusters were different in grade (P < 0.01), but not with tumor stage, age and gender (Figure 3E). Furthermore, GSEA analysis showed that oncogenic pathways (apoptosis, G2M-checkpoint, IL2-STAT5-signaling, IL6-JAK-STAT3-signaling, inflammatory response, PI3K-AKT-MTOR-signaling, TNFA-signaling via NF-κB, unfolded protein response) were significantly enriched in Cluster 2 (Supplementary Figure 1B).

3. Construction and verification of costimulatory moleculerelated risk signature

To narrow down candidate genes and construct the risk signature, the least absolute shrinkage and selection operator (LASSO) Cox regression analysis was performed in the training set, and 6 of 11 prognosis-related CMGs were identified (Supplementary Figure 2). The formula to calculate the risk score as follows: risk score = (0.25584 * TNFSF4) + (-0.29002 * TMIGD2) + (0.13379 * TNFRSF4) + (0.22009 * TNFRSF11B) + (0.40207 * TNFRSF11A) + (-0.78099 * CD40LG). We calculated the risk scores for every HCC patient in the training set according to the above formular. Patients in the training set were divided into high- and low-risk groups based on their median risk sore. A significant difference of OS was observed in different subgroups. High-risk patients had a poorer OS than low-risk groups (P < 0.001) (Figure 4D). Time-dependent receiver operating characteristic (ROC) analysis was used to evaluate the sensitivity and specificity of the risk signature. The areas under the curve (AUC) were 0.756 at 1-year survival, 0.791 at 3-year survival and 0.729 at 5-year survival (Figure 4E). We ranked the risk scores of patients and analyzed their distribution in the training set (Figure 4A). The survival status of HCC patients in the training set was showed on the dot plot (Figure 4B). The heatmap displayed the expressions of 6 prognosis-related CMGs between two risk groups (Figure 4C).

To determine the stability of the risk signature, we further verified the predictive capability in the test set and total set. The risk score was calculated for each patient in the test set and total set by the same formular obtained from the training set and the patients were classified into high- and low-risk groups. Similarly, Kaplan-Meier survival curve showed significantly difference in two risk groups among the test set. The OS of the high-risk groups was poorer than that of the low-risk groups (P=0.019) (Figure 4I). The 1-year AUC was 0.728, the 3-year AUC was 0.644 and the 5-year AUC was 0.654 (Figure 4J). The survival status, the distribution of the risk score and the expression heatmap of 6 prognosis-related CMGs in the test set were presented in Figure 4F-H.

The results in the total set were similar to the training set and test set. Patients in the high-risk group had a significantly shorter prognosis than patients in the low-risk group (P<0.001) (Figure 4N). In the total set, the AUC was 0.739 at 1 year, 0.708 at 3 years and 0.662 at 5years (Figure 4O). The distribution of the risk score, survival status and the expression patterns of 6 prognosis-related CMGs were showed in Figure 4K-M.

4. Independent prognostic value of the risk signature

We performed the univariate and multivariate Cox regression analyses to examine whether the risk score could act as an independent prognosis variable of HCC. Univariate Cox regression analysis showed that pathological tumor stage and risk score were significantly associated with the prognosis (Figure 5A, C, E). Multivariate Cox regression analyses further identified that the risk score was an independent prognostic factor for OS in the training set, test set and total set (Figure 5B, D, F).

5. Correlations between the risk signature and clinicopathological factors

The association between the risk model and clinical characteristics were analyzed. The heatmap displayed the expressions of 6 prognosis-related CMGs and the clinicopathological characteristics in high- and low-risk groups. The risk score was significantly correlated with histological grade, pathological T stage and clinical stage (Figure 6A). Patients were divided into subgroups according to the clinical variables. Differences in clinicopathological factors between high- and low-risk groups were showed in Figure 6B. The risk score was significantly higher in advanced grade, T stage and clinical stage. Meanwhile, the correlation between the risk score and clinicopathological features on OS was also explored. Survival analysis manifested that higher risk score were correlated with poor prognosis in most subgroups (Figure 6C).

6. Construction of a novel nomogram

We constructed a nomogram to predict the survival rates for HCC patients based on age, gender, histological grade, TNM stage, tumor stage and risk score (Figure 7A). The calibration curves of the nomogram indicated good consistency between the predicted survival rate and actual 1-, 3- and 5-year survival rate (Figure 7B). The AUCs of risk score and tumor stage were 0.739 and 0.671 in 1-year, 0.698 and 0.680 in 3-year, 0.638 and 0.663 in 5-year, respectively (Figure 7C). These findings suggested that the risk signature might be reliable to predict the OS for HCC patients.

7. Functional enrichment analyses based on the risk signature

To explore the potential biological processes for the prognostic risk signature, a total of 474 DEGs were obtained in high- and low-risk groups with the criteria FDR < 0.05 and $|log2FC| \ge 1$. Among them, 63 genes were downregulated in high-risk group, while 411 genes were upregulated (Supplementary Table 2). Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways analyses were carried out based on the DEGs. The results of GO analysis indicated that the DEGs were mainly related to

nuclear division. KEGG analysis showed that DEGs were mostly enriched in cell cycle (Supplementary Figure 3).

8. Association between risk signature and tumor immune microenvironment

The differences of tumor-infiltrating immune cells between high- and low-risk groups were analyzed to explore the correlations between the prognostic risk signature and tumor immune microenvironment (TIME). Supplementary Figure 4 displayed the abundance of 22 immune cells between high- and low-risk subgroups. Among 22 immune cell types, memory B cells and macrophage M0 were positively correlated with the risk score, while the abundance of naïve B cells, plasma cells and regulatory T cells (Tregs) were significantly enriched in low-risk group (Figure 8A). Furthermore, the relative proportion of naïve B cells, resting memory CD4 T cells, activated memory CD4 T cells, regulatory T cells, gamma delta T cells, macrophage M1 and resting mast cells were significantly associated with OS (Figure 8B).

To further explore the relationship between the risk signature and immune status, we performed the expression profiles of 29 immune signature gens sets (16 types of immune cells and 13 immune-related pathways) in high- and low-risk groups using the single-sample gene set enrichment analysis (ssGSEA). The heatmap displayed the significant differences in immune status between high- and low-risk samples (Figure 9A). The low-risk subgroup showed higher levels of infiltration of immune cells and higher activity of immune-related pathways (Figure 9B). We found that the immune score and stromal score were higher in low-risk groups, while the tumor purity was significantly lower in low-risk subgroup (Figure 9C).

9. Differences in molecular characteristics between highand low-risk groups

We evaluated the relationship between mutation characteristics and the risk signature in TCGA HCC patients with available somatic mutation data. TMB was higher in high-risk patients in spite of no significant difference (Figure 10B). We also identified the top 20 genes with the highest mutation rates in high- and low-risk subgroups (Figure 10A). Additionally, we explored the association between immunophenoscore (IPS) and risk signature to predict the potential clinical efficacy and the response to ICI therapy in HCC patient. The IPS, IPS-CTLA4, IPS-PD1-PD-L1-PD-L2, and IPS-PD1-PD-L1-PD-L2-CTLA4 blocker were significantly higher in low-risk group, implying that HCC patients with low-risk score could benefit more from ICI therapy than high-risk patients (Figure 10C).

10. Verification of prognostic CMGs

We verified the expression of the CMGs (TNFSF4, TNFRSF4, TMIGD2, TNFRSF11A, TNFRSF11B, CD40LG) in 43 pairs of tumorous and non-tumorous tissue specimens from patients with HCC using qRT-PCR analysis. The results of qRT-PCR showed that the expression of TNFSF4, TNFRSF4, TNFRSF11A and CD40LG was higher in HCC tissues compared to normal tissues. However, the mRNA expression of

TNFRSF11B and TMIGD2 was higher in normal tissues, which was consistent with the results of bioinformatic analysis.

Discussion

Preliminary data from trials of ICIs in the treatment of HCC led to encouraging results. Nevertheless, with the rapid augment in the utilization of ICIs, immune-related adverse events of HCC arose(27, 28). Multiple studies found that the usage of ICIs strategies targeting costimulatory molecules for HCC management was promising(29). Therefore, it is necessary to improve the effect on immunotherapy by selecting the suitable HCC patients according to costimulatory molecules expression patterns. In this study, we analyzed the mRNA expression patterns of costimulatory molecules-related in HCC and selected six genes with prognostic values. Then, we constructed the first costimulatory molecule-related prognostic signature for HCC patients. We found that prognostic signature was strongly associated with clinical characteristics. Additionally, our signature was significantly correlated with tumor immune microenvironment and the response to immunotherapy. Univariate and multivariate cox regression analysis indicated our signature could be an independent prognostic factors for the survival of HCC patients. These findings suggested that CMGs risk signature may indicate some insights to personalized targeted treatment in clinical practice.

Costimulatory molecules played an important role in immunotherapy(17). Recent findings demonstrated that CD28 co-stimulation was necessary for responses to PD-1 blockade in tumor rejection(30). Thus, understanding the states of costimulatory molecules in HCC patients will help us determine which patients might benefit in immunotherapy. To explore the expression levels of costimulatory molecules in HCC, we acquired 13 members of the B7-CD28 family and 46 members of the TNF family for HCC patients. Six costimulatory molecular genes (TNFSF4, TNFRSF4, TMIGD2, TNFRSF11A, TNFRSF11B, CD40LG) with prognostic values were selected. The TNFRSF4-TNFSF4 pathway provided crucial costimulatory signals for CD4⁺T cell responses(31). Previous study showed that TNFSF4 was closely related to the unfavorable prognosis of HCC patients(32). In addition, TNFRSF4 was overexpressed in HCC, associated with a more aggressive phenotype and the activation of multiple immunosuppressive pathways(33). Studies showed TMIGD2 was mainly expressed in tissue-resident lymphocyte T cells, related to improved tumor prognosis(34). The different interaction between TMIGD2 and B7-H5 have been identified in certain cancers, such as lung cancer, osteosarcoma, oral squamous cell carcinoma (OSCC), colorectal cancer (CRC) and glioma(35). Of note, our results firstly revealed that TMIGD2 was highly expressed in HCC with favorable prognosis. TNFRSF11A, also known as RANK, was significantly up-regulated in HCC, and can lead directly to migration and invasion by its ligand(36). Interestingly, genetic deletion of TNFRSF11A in thymic epithelial cells resulted in impaired thymic involution and blunted expansion of natural regulatory T (Treg) cells(37). Additionally, study showed that HCC patients with high serum TNFRSF11B, also known as osteoprotegerin(OPG), level had poorer survival rates compared with HCC patients with low OPG level(38). CD40 ligand-expressing dendritic cells could induce regression of HCC(39). Moreover, the expression levels of six prognostic genes were verified using gRT-

PCR and immunohistochemistry. However, the expression of CMGs was not completely similar with our previous results, may partly owing to different race and clinical characteristics. With these six costimulatory molecular genes elucidated in immunity, we hope that the signature constructed by these could predict the response to immunotherapy for HCC patients.

Tumors were complex ecosystems, defined by spatiotemporal interactions among heterogeneous cell types(40). Subsequently, we compared the associations between our signature and tumor immune microenvironment, the immune cell infiltration and tumor mutation profiles in high-risk and low-risk patients. Our results showed that naïve B cells, plasma cells and regulatory T cells (Tregs) were significantly enriched in low-risk groups. Much of researches regarded Tregs as an immunosuppressive cell, posing anti-tumor immunity in various cancers(41). Nevertheless, some scholars stated that inhibiting the expression of PD-1 promoted other immune checkpoints, resulting in impaired immune killing ability(42, 43). Accordingly, fewer Tregs cells in HCC patients with poor prognosis, indicated that those were more likely to be inhibited by PD-1 and activated more immune checkpoints. Correspondingly, high-risk subgroup manifested lower levels of infiltration of immune cells, implicating less process in immune activation.

Tumor mutation burden (TMB) is emerging as a potential biomarker, and participated in immunotherapyrelated pathway(44, 45). We found that the tumor mutation burden (TMB) in the high-risk group was higher than that in the low-risk group with no significant, partly due to the small sample size. TP53 mutation frequency was evidently higher in high-risk group than low-risk group, suggesting more increases genomic instability and complicated major pathway signaling changes in HCC. Additionally, it was important to highlight that different microenvironment-based immune subtypes, based on gene profiling or signatures, and other molecular features, may help identify subgroups of patients more likely to benefit from specific therapies(46). Some scholars have found that immune-excluded tumors in HCC were proposed to be primarily resistant to ICIs(47). IPS could predict the response to immunotherapy in cervical cancer and HCC. The prediction of IPS has been demonstrated in different studies(48, 49). In the present study, we found low-risk group tended to have higher IPS-CTLA4, IPS-PD1/PD-L1/ PD-L2, and IPS-PD1/PD-L1/PD-L2⁺CTLA4, implying that HCC patients with low-risk score could benefit more from immunotherapy than high-risk patients. Therefore, our signature was of great help to clinical immunotherapy decision.

However, there were some limitations in this study. Firstly, we did not explore the exact function of six costimulatory molecule genes in HCC. Thus, it was still necessary to clarify the mechanism of them in the future. Secondly, it was inevitable that there were limited clinical information for HCC patients in public datasets, so the values of the prognostic signature needed to be determined by experimental and prospective studies. Moreover, the risk signature for evaluating the response to immunotherapy was restricted to costimulatory molecule genes and tumor immune microenvironment was highly heterogeneous. Therefore, the prognostic information for HCC patients with immunotherapy were needed to validate the prediction power of our signature clinically.

Conclusion

In our study, we first elucidated the expression of costimulatory molecules for HCC patients, and constructed a six CMGs prognostic signature. The costimulatory molecular-related signature could stratify patients into different subsets with adverse clinical outcomes. In addition, immunotherapy response prediction by our signature explained disparate effect on HCC patients. Consequently, we believed our research manifested the capacity of costimulatory molecules and provided clinicians with applicable treatment.

Declarations

Ethics approval and consent to participate

The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study conformed to the provisions of the Declaration of Helsinki (as revised in 2013). The research was approved by the Institutional Research Ethics Committees of the Xijing Hospital. Informed consent for publication was obtained from all patients for collection of tissue samples prior to the surgery.

Consent for publication

Not applicable.

Competing interests

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

Availability of data and materials

Gene expression profiles, clinical information and mutation data of THCA in this study are available from the public database (TCGA, https:// portal. gdc. cancer. gov/).

Funding

This study was funded by the National Natural Science Foundation of China (Nos. 81770569, 81870421, and 81773072), the International Cooperation and Exchange of the National Natural Science Foundation of China (No. 81820108005).

Authors' contributions

Study concept and design (YH), acquisition of data (YNH,JYL), analysis and interpretation of data (YNH,JHY, JYL), drafting of the manuscript (YNH, JHY), critical revision of the manuscript for important intellectual content (MZ,YSL,JBW,YH), administrative, technical, or material support, study supervision (XZ, FFY,SYM).

Acknowledgements

Not applicable.

References

- 1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021; doi:10.3322/caac.21660.
- 2. Llovet JM, Zucman-Rossi J, Pikarsky E, et al. Hepatocellular carcinoma. Nat Rev Dis Primers. 2016; doi:10.1038/nrdp.2016.18.
- 3. El-Khoueiry AB, Sangro B, Yau T, et al. Nivolumab in patients with advanced hepatocellular carcinoma (CheckMate 040): an open-label, non-comparative, phase 1/2 dose escalation and expansion trial. The Lancet. 2017; doi:10.1016/s0140-6736(17)31046-2.
- 4. Zhu AX, Finn RS, Edeline J, et al. Pembrolizumab in patients with advanced hepatocellular carcinoma previously treated with sorafenib (KEYNOTE-224): a non-randomised, open-label phase 2 trial. The Lancet Oncology. 2018; doi:10.1016/s1470-2045(18)30351-6.
- 5. Lee MS, Ryoo B-Y, Hsu C-H, et al. Atezolizumab with or without bevacizumab in unresectable hepatocellular carcinoma (GO30140): an open-label, multicentre, phase 1b study. The Lancet Oncology. 2020; doi:10.1016/s1470-2045(20)30156-x.
- 6. Sperandio RC, Pestana RC, Miyamura BV, et al. Hepatocellular Carcinoma Immunotherapy. Annu Rev Med. 2021; doi:10.1146/annurev-med-042220-021121.
- 7. Egen JG, Ouyang W, Wu LC. Human Anti-tumor Immunity: Insights from Immunotherapy Clinical Trials. Immunity. 2020; doi:10.1016/j.immuni.2019.12.010.
- 8. McHayleh W, Bedi P, Sehgal R, et al. Chimeric Antigen Receptor T-Cells: The Future is Now. J Clin Med. 2019; doi:10.3390/jcm8020207.
- 9. Chapman NM, Boothby MR, Chi H. Metabolic coordination of T cell quiescence and activation. Nat Rev Immunol. 2020; doi:10.1038/s41577-019-0203-y.
- 10. Chen L, Flies DB. Molecular mechanisms of T cell co-stimulation and co-inhibition. Nat Rev Immunol. 2013; doi:10.1038/nri3405.

- 11. Esensten JH, Helou YA, Chopra G, et al. CD28 Costimulation: From Mechanism to Therapy. Immunity. 2016; doi:10.1016/j.immuni.2016.04.020.
- 12. Sangro B, Sarobe P, Hervas-Stubbs S, et al. Advances in immunotherapy for hepatocellular carcinoma. Nat Rev Gastroenterol Hepatol. 2021; doi:10.1038/s41575-021-00438-0.
- 13. Croft M. The role of TNF superfamily members in T-cell function and diseases. Nat Rev Immunol. 2009; doi:10.1038/nri2526.
- Lu Y, Zhang M, Wang S, et al. p38 MAPK-inhibited dendritic cells induce superior antitumour immune responses and overcome regulatory T-cell-mediated immunosuppression. Nat Commun. 2014; doi:10.1038/ncomms5229.
- Aberg KM, Radek KA, Choi EH, et al. Psychological stress downregulates epidermal antimicrobial peptide expression and increases severity of cutaneous infections in mice. J Clin Invest. 2007; doi:10.1172/JCl31726.
- 16. Shibahara I, Saito R, Zhang R, et al. OX40 ligand expressed in glioblastoma modulates adaptive immunity depending on the microenvironment: a clue for successful immunotherapy. Mol Cancer. 2015; doi:10.1186/s12943-015-0307-3.
- 17. Wei SC, Duffy CR, Allison JP. Fundamental Mechanisms of Immune Checkpoint Blockade Therapy. Cancer Discov. 2018; doi:10.1158/2159-8290.CD-18-0367.
- Aye L, Song X, Yang J, et al. Identification of a Costimulatory Molecule Gene Signature to Predict Survival and Immunotherapy Response in Head and Neck Squamous Cell Carcinoma. Front Cell Dev Biol. 2021; doi:10.3389/fcell.2021.695533.
- 19. Hua X, Ge S, Zhang J, et al. A costimulatory molecule-related signature in regard to evaluation of prognosis and immune features for clear cell renal cell carcinoma. Cell Death Discov. 2021; doi:10.1038/s41420-021-00646-2.
- 20. Wang W, Sun B, Xia Y, et al. RNA N6-Methyladenosine-Related Gene Contribute to Clinical Prognostic Impact on Patients With Liver Cancer. Front Genet. 2020; doi:10.3389/fgene.2020.00306.
- 21. Wang H, Lengerich BJ, Aragam B, et al. Precision Lasso: accounting for correlations and linear dependencies in high-dimensional genomic data. Bioinformatics. 2019; doi:10.1093/bioinformatics/bty750.
- 22. Hsuan-Yu Chen MS, Sung-Liang Yu, Ph.D., Chun-Houh Chen, Ph.D., Gee-Chen Chang, M.D., Ph.D., Chih-Yi Chen, M.D., Ang Yuan, M.D., Ph.D., Chiou-Ling Cheng, M.Sc., Chien-Hsun Wang, M.Sc., Harn-Jing Terng, Ph.D., Shu-Fang Kao, M.Sc., Wing-Kai Chan, M.D., Han-Ni Li, M.Sc., Chun-Chi Liu, M.Sc., Sher Singh, Ph.D., Wei J. Chen, M.D., Sc.D., Jeremy J.W. Chen, Ph.D., and Pan-Chyr Yang, M.D., Ph.D. A Five-Gene Signature and Clinical Outcome in Non–Small-Cell Lung Cancer. N Engl J Med. 2007; doi:10.1056/NEJMoa060096.
- 23. Becht E, Giraldo NA, Lacroix L, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol. 2016; doi:10.1186/s13059-016-1070-5.

- 24. Yoshihara K, Shahmoradgoli M, Martinez E, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun. 2013; doi:10.1038/ncomms3612.
- 25. Mayakonda A, Lin DC, Assenov Y, et al. Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res. 2018; doi:10.1101/gr.239244.118.
- 26. Robinson DR, Wu YM, Lonigro RJ, et al. Integrative clinical genomics of metastatic cancer. Nature. 2017; doi:10.1038/nature23306.
- 27. Richard S. Finn MB-YR, MD, PhD; Philippe Merle, MD, PhD; Masatoshi Kudo, MD, PhD; Mohamed Bouattour, MD; Ho Yeong Lim, MD; Valeriy Breder, MD, PhD; Julien Edeline, MD, PhD; Yee Chao, MD, PhD; Sadahisa Ogasawara, MD; Thomas Yau, MD; Marcelo Garrido, MD; Stephen L. Chan, MD; Jennifer Knox, MD; Bruno Daniele, MD; Scot W. Ebbinghaus, MD; Erluo Chen, MPH; Abby B. Siegel, MD; Andrew X. Zhu, MD, PhD; and Ann-Lii Cheng, MD, PhD; o n behalf of the KEYNOTE-240 investigators. Pembrolizumab As Second-Line Therapy in Patients With Advanced Hepatocellular Carcinoma in KEYNOTE-240: A Randomized, Double-Blind, Phase III Trial. Journal of Clinical Oncology. 2020; doi:10.1200/JC0.19.01307.
- 28. Lyon AR, Yousaf N, Battisti NML, et al. Immune checkpoint inhibitors and cardiovascular toxicity. The Lancet Oncology. 2018; doi:10.1016/s1470-2045(18)30457-1.
- 29. Chinnadurai R, Scandolara R, Alese OB, et al. Correlation Patterns Among B7 Family Ligands and Tryptophan Degrading Enzymes in Hepatocellular Carcinoma. Front Oncol. 2020; doi:10.3389/fonc.2020.01632.
- 30. Alice O. Kamphorst AW, Tahseen Nasti, Shu Yang, Ruan Zhang, Daniel L. Barber, Bogumila T. Konieczny, Candace Z. Daugherty, Lydia Koenig, Ke Yu, Gabriel L. Sica, Arlene H. Sharpe, Gordon J. Freeman, Bruce R. Blazar, Laurence A. Turka, Taofeek K. Owonikoko, Rathi N. Pillai, Suresh S. Ramalingam, Koichi Araki, Rafi Ahmed. Rescue of exhausted CD8 T cells by PD-1–targeted therapies is CD28-dependent. Science. 2017; doi:10.1126/science.aaf0683.
- 31. Gajdasik DW, Gaspal F, Halford EE, et al. Th1 responses in vivo require cell-specific provision of OX40L dictated by environmental cues. Nat Commun. 2020; doi:10.1038/s41467-020-17293-3.
- 32. Hong WF, Gu YJ, Wang N, et al. Integrative Characterization of Immune-relevant Genes in Hepatocellular Carcinoma. J Clin Transl Hepatol. 2021; doi:10.14218/JCTH.2020.00132.
- Xie K, Xu L, Wu H, et al. OX40 expression in hepatocellular carcinoma is associated with a distinct immune microenvironment, specific mutation signature, and poor prognosis. Oncoimmunology. 2018; doi:10.1080/2162402X.2017.1404214.
- 34. Tian Y, Sun Y, Gao F, et al. CD28H expression identifies resident memory CD8 + T cells with less cytotoxicity in human peripheral tissues and cancers. Oncoimmunology. 2019; doi:10.1080/2162402X.2018.1538440.
- 35. Zhong C, Lang Q, Yu J, et al. Phenotypical and potential functional characteristics of different immune cells expressing CD28H/B7-H5 and their relationship with cancer prognosis. Clin Exp Immunol. 2020; doi:10.1111/cei.13413.

- 36. Song FN, Duan M, Liu LZ, et al. RANKL promotes migration and invasion of hepatocellular carcinoma cells via NF-kappaB-mediated epithelial-mesenchymal transition. PLoS One. 2014; doi:10.1371/journal.pone.0108507.
- 37. Paolino M, Koglgruber R, Cronin SJF, et al. RANK links thymic regulatory T cells to fetal loss and gestational diabetes in pregnancy. Nature. 2021; doi:10.1038/s41586-020-03071-0.
- Zhang C, Lin J, Ni X, et al. Prognostic Value of Serum Osteoprotegerin Level in Patients With Hepatocellular Carcinoma Following Surgical Resection. Front Oncol. 2021; doi:10.3389/fonc.2021.731989.
- Gonzalez-Carmona MA, Lukacs-Kornek V, Timmerman A, et al. CD40ligand-expressing dendritic cells induce regression of hepatocellular carcinoma by activating innate and acquired immunity in vivo. Hepatology. 2008; doi:10.1002/hep.22296.
- 40. Sun Y, Wu L, Zhong Y, et al. Single-cell landscape of the ecosystem in early-relapse hepatocellular carcinoma. Cell. 2021; doi:10.1016/j.cell.2020.11.041.
- 41. Zheng C, Zheng L, Yoo JK, et al. Landscape of Infiltrating T Cells in Liver Cancer Revealed by Single-Cell Sequencing. Cell. 2017; doi:10.1016/j.cell.2017.05.035.
- 42. Ellestad KK, Thangavelu G, Ewen CL, et al. PD-1 is not required for natural or peripherally induced regulatory T cells: Severe autoimmunity despite normal production of regulatory T cells. Eur J Immunol. 2014; doi:10.1002/eji.201444688.
- 43. Huang RY, Francois A, McGray AR, et al. Compensatory upregulation of PD-1, LAG-3, and CTLA-4 limits the efficacy of single-agent checkpoint blockade in metastatic ovarian cancer. Oncoimmunology. 2017; doi:10.1080/2162402X.2016.1249561.
- 44. Huang T, Yan T, Chen G, et al. Development and Validation of a Gene Mutation-Associated Nomogram for Hepatocellular Carcinoma Patients From Four Countries. Front Genet. 2021; doi:10.3389/fgene.2021.714639.
- 45. Donehower LA, Soussi T, Korkut A, et al. Integrated Analysis of TP53 Gene and Pathway Alterations in The Cancer Genome Atlas. Cell Rep. 2019; doi:10.1016/j.celrep.2019.07.001.
- 46. Daniela Sia YJ, Iris Martinez-Quetglas, Olga Kuchuk, Carlos Villacorta-Martin, Manuel Castro de Moura, Juan Putra, Genis Camprecios, Laia Bassaganyas, Nicholas Akers, Bojan Losic, Samuel Waxman, Swan N. Thung, Vincenzo Mazzaferro, Manel Esteller, Scott L. Friedman, Myron Schwartz, Augusto Villanueva, Josep M. Llovet. Identification of an immune-specific class of hepatocellular carcinoma, based on molecular features. Gastroenterology. 2017; doi:10.1053/j.gastro.2017.06.007.
- 47. Ruiz de Galarreta M, Bresnahan E, Molina-Sanchez P, et al. beta-Catenin Activation Promotes Immune Escape and Resistance to Anti-PD-1 Therapy in Hepatocellular Carcinoma. Cancer Discov. 2019; doi:10.1158/2159-8290.CD-19-0074.
- 48. Yang S, Wu Y, Deng Y, et al. Identification of a prognostic immune signature for cervical cancer to predict survival and response to immune checkpoint inhibitors. Oncoimmunology. 2019; doi:10.1080/2162402X.2019.1659094.

49. Xu Y, Wang Z, Li F. Survival prediction and response to immune checkpoint inhibitors: A prognostic immune signature for hepatocellular carcinoma. Transl Oncol. 2021; doi:10.1016/j.tranon.2020.100957.

Tables

Table 1

Characteristic	Туре	Total set	Test set	Training set	p value
Age	<=65	232(62.7%)	116(63.04%)	116(62.37%)	0.9782
	>65	138(37.3%)	68(36.96%)	70(37.63%)	
Gender	Female	121(32.7%)	61(33.15%)	60(32.26%)	0.9422
	Male	249(67.3%)	123(66.85%)	126(67.74%)	
Grade	G1-2	232(62.7%)	116(63.04%)	116(62.37%)	0.9214
	G3-4	133(35.95%)	65(35.33%)	68(36.56%)	
	Unknow	5(1.35%)	3(1.63%)	2(1.08%)	
Stage	Stage I-II	256(69.19%)	130(70.65%)	126(67.74%)	0.2024
	Stage III-IV	90(24.32%)	38(20.65%)	52(27.96%)	
	Unknow	24(6.49%)	16(8.7%)	8(4.3%)	
T -	T1-2	274(74.05%)	140(76.09%)	134(72.04%)	0.2946
	ТЗ-4	93(25.14%)	41(22.28%)	52(27.96%)	
	Unknow	3(0.81%)	3(1.63%)	0(0%)	
M	M0	266(71.89%)	128(69.57%)	138(74.19%)	0.573
	M1	4(1.08%)	3(1.63%)	1(0.54%)	
	Unknow	100(27.03%)	53(28.8%)	47(25.27%)	
N	N0	252(68.11%)	123(66.85%)	129(69.35%)	0.1515
	N1-3	4(1.08%)	0(0%)	4(2.15%)	
	Unknow	114(30.81%)	61(33.15%)	53(28.49%)	

The clinical information in training set, test set and total set.

Table 2

Page 18/30

The primer sequences for qRT-PCR analysis.

Premier	Sequences (5'-3')	
β-Actin-F	CTCCATCCTGGCCTCGCTGT	
β-Actin-R	GCTGTCACCTTCACCGTTCC	
TNFSF4-F	CCCTGGGACCTTTGCCTATT	
TNFSF4-R	GGGGTTGGACCCTTTCCATC	
TNFRSF4-F	AAGCCTGGAGTTGACTGTGC	
TNFRSF4-R	CCTGTCCTCACAGATTGCGT	
TNFRSF11A-F	GTTGCAGCTCAACAAGGACAC	
TNFRSF11A-R	CAGAGAAGAACTGCAAACCGC	
TNFRSF11B-F	CTGGAACCCCAGAGCGAAAT	
TNFRSF11B-R	GCCTCCTCACACAGGGTAAC	
TMIGD2-F	AGAACAGAAACCGGATCGCA	
TMIGD2-R	GGCTGTTACCTGAGTCCCTT	
CD40LG-F	ATGGGAAACAGCTGACCGTT	
CD40LG-R	GATTGTTGCCCGCAAGGTTT	

Table 3

The clinical features of the HCC (n=43).

Characteristics	Samples(N=43)	Percentage(%)
Gender		
Male	31	72
Female	12	28
Age		
≤ 60	35	81
> 60	8	19
Aetiology		
HBV	41	95
Others	2	5
AFP, ng/ml		
<400	37	86
≥400	6	14
Т		
T1-2	39	91
T3-4	4	9
Ν		
N0	42	98
N1	1	2
М		
M0	43	100
M1		
Stage		
Stage 🛛 - 🕅	39	91
Stage 🛛 - 🕅	4	9



The flowchart of the study



Expressions of the CMGs in HCC. (A) The expression levels of 59 CMGS in HCC. *P < 0.05, **P < 0.01, ***P < 0.001. (B) The correlation of the 59 CMGs. (C) PPI network showed the interactions of the CMGs (the highest confidence:0.9).



Consensus clustering analysis on the CMGs. (A) Univariate Cox regression analysis identified prognosisrelated CMGs. (B) Consensus clustering matrix for k=2. (C) Principal Component Analysis (PCA) plot for clusters. (D) Kaplan–Meier overall survival (OS) curves of Cluser1 and Cluster 2. (E) Heatmap and clinical factors of the two clusters.



Construction of CMGs risk signature for HCC. The distribution of the risk score, survival status, expression of 6 prognosis-related CMGs in high- and low-risk groups, Kaplan-Meier survival curve, time-dependent ROC curve analyses of the CMGs risk signature in the training set (A-E), in the test set (F-G), in the total set (K-O), respectively.



Univariate and multivariate Cox regression analyses. Univariate and Multivariate Cox regression analysis of the correlation between the risk score and clinicopathological features in the training set (A-B), test set (C-D) and total set (E-F), respectively.



Clinical characteristics of CMGs prognostic signature in different subgroups. (A) The heatmap and clinicopathological factors of high- and low-risk subgroups. (B) The relationship between the CMGs prognostic signature and age, gender, stage, grade, TMN stage. (C) Survival analyses of CMGs risk model in different clinical subgroups.



Figure 7

Construction and validation of a novel nomogram. (A) The nomogram for predicting 1-year, 3-year and 5year OS of HCC in total set. (B) The calibration curves for internal validation of the nomogram in total set. (C) The time-dependent ROC of the nomogram and clinical factors for 1-year, 3-year and 5year OS prediction in total set.



Analysis of tumor infiltrating immune cells. (A) The association of CMGs prognostic signature and immune cells infiltration. (B) The relationship between OS and immune cells infiltration.



Comparison of the immune status in high- and low-risk groups. (A) The immune status of HCC patients in high- and low-risk groups. Tumor purity, ESTIMATE score, immune score and stromal score of every sample were showed in the heatmap. (B) The box plot displayed the differences of enrichment scores of 16 types of immune cells and 13 immune-related pathways in high- and low-risk groups. (C) The differences of stromal score, immune score, ESTIMATE score and tumor purity in high- and -risk groups with violin plots. *** p < 0.001.



Evaluation of tumor mutation burden and the response to immunotherapy. (A) Mutation profiles of highand low-risk groups. (B) The relationship between CMGs risk signature and TMB. (C) The association between IPS and risk signature for HCC patients.



Verification of prognostic genes in HCC tissues and normal tissues.®A) TNFSF4,®B®TNFRSF4, (C) TNFRSF11A, (D) TNFRSF 11B, (E) TMIGD2, (F) CD40LG. ns : not significant,*P <0.05,**P <0.01.

Supplementary Files

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

- SFigure1.tif
- SFigure2.tif
- SFigure3.tif
- SFigure4.tif