Translational Oncology 14 (2021) 100957

Contents lists available at ScienceDirect

Translational Oncology

journal homepage: www.elsevier.com/locate/tranon

Survival prediction and response to immune checkpoint inhibitors: A prognostic immune signature for hepatocellular carcinoma

Ying Xu[a,b,1], Zheng Wang[a,c,1], Fufeng Li[a,b,*]

[a] Shanghai University of Traditional Chinese Medicine, No. 1200, Cailun Road, 200120 Shanghai, China
[b] Laboratory of TCM Four Processing, Shanghai University of Traditional Chinese Medicine, Shanghai, China
[c] Shuguang Hospital Affiliated to Shanghai University of Traditional Chinese Medicine, Shanghai, China

ARTICLE INFO

Keywords:
Immune signature
Hepatocellular carcinoma
Prognosis
Immune checkpoint inhibitor response
Tumor immune microenvironment

ABSTRACT

Hepatocellular carcinoma (HCC) is one of the most common cancers all over the world. Several studies have explored if immune-related genes and tumor immune microenvironment could play roles in HCC prognoses. This study is aimed at developing a prognostic signature of HCC based on immune-related genes or tumor immune microenvironment to predict survival and response to immune checkpoint inhibitors (ICIs). We constructed a prognostic signature using bioinformatics method and validated its predictive capability. The mechanisms of the signature prediction were explored with The Cancer Immunome Atlas (TCIA) and mutation analysis. We also explored the association between the signature and immunophenoscore (IPS), which is the marker of ICIs response. A 6 immune-related-gene (6-IRG) signature was developed. It was revealed in a multivariate analysis that the 6-IRG signature was an independent prognostic factor of overall survival and progression-free interval among HCC patients. In the high-risk group of 6-IRG signature score, macrophage M0 cells and regulatory T cells, which are observed associated with poor overall survival in our study, were higher. The low-risk group had a higher IPS, which meant a better response to ICIs. Taken together, we constructed a reliable 6-IRG signature for prediction of survival and response to ICIs. The signature needs further testing for clinical application.

Introduction

Hepatocellular carcinoma (HCC) is one of the most common cancers all over the world, which is affected by both of environment and diet [1]. It's believed that the development of HCC is a complex progression with a number of other factors. Epidemiology and experimental studies reveal that the occurrence of HCC is associated with the infection of HBV and HCV, aflatoxin, alcohol and liver cirrhosis. Early diagnosis of HCC is an important means to raise the survival rate of patients. With the advances of technology and of the understanding of HCC biology, new biomarkers have been constantly discovered. The application of chemotherapy, radiotherapy and surgical resection improved the prognosis of early HCC.

However, it is difficult to prevent metastasis and recurrence of HCC [2], which is responsible for most HCC deaths. To take a precision medicine approach in cancer immunotherapy, it is critical to identify and develop predictive biomarkers of checkpoint inhibitor-based immunotherapy responsiveness. Tumor immune microenvironment is an important variable relating to the progression of HCC. Additionally, several immune prognostic signatures have been reported to predict the prognosis of patients with cancer, such as hepatocellular carcinoma [3], lung cancer [4], ovarian cancer [5] and colorectal cancer [6], These studies indicate that IRGs or TIME can serve as promising biomarkers for estimating survival in corresponding cancer. However, limited studies have explored whether immune-related genes (IRGs) or TIME could be indicators for the prognosis of HCC. An improved prognostic signature of HCC based on IRGs or TIME is urgently needed to predict outcomes and response to ICI.

Our study was aimed at developing a new immune signature with predictive capability on the basis of IRGs or tumor immune microenvironment. Following the development of the immune signature, its association to clinicopathological characteristics, and prognosis overall survival and progression-free interval in HCC was explored. Moreover, immune cell infiltration, mutation data, and immunophenoscore linked to this signature in HCC was thoroughly investigated. This may help to provide a more complete understanding of, and more precise immunotherapy for, HCC.

* Corresponding author at: Shanghai University of Traditional Chinese Medicine, No. 1200, Cailun Road, 200120 Shanghai, China.
E-mail address: li_fufeng@aliyun.com (F. Li).
[1] Zheng Wang and Ying Xu contributed equally to this work and should be considered as co-first authors.

https://doi.org/10.1016/j.tranon.2020.100957
Received 14 July 2020; Received in revised form 7 November 2020; Accepted 13 November 2020

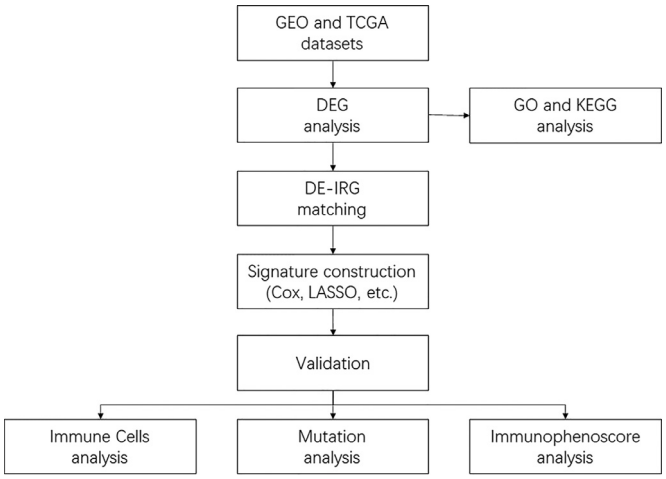

Fig. 1. Overall flowchart of this study.

Materials and methods

The flowchart of the whole study was presented in Fig. 1.

Patient data

We downloaded:

(1) 4 datasets (GSE45436 [7], GSE62232 [8], GSE84402 [9] and GSE101685 (unpublished)) of HCC expression microarrays form the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/gds/) [10];

(2) the clinical information and gene expression profiles of HCC from The Cancer Genome Atlas (TCGA) data portal (https://portal.gdc.cancer.gov/); and (3) the list of immune-related genes (IRGs) from the Immunology Database and Analysis Portal (ImmPort) database (https://immport.niaid.nih.gov) [11].

Differentially expressed genes analysis

In order to find the IRGs that involved in the development of HCC, we used the limma package [12] to screen the common differentially expressed genes (DEGs) between tumor samples and normal samples from the 4 GEO datasets with adjusted P-value <0.05 and $|\log_2$ (fold change) $|>1$. Then we selected the intersection of DEG list and IRG list as differentially expressed immune-related genes (DE-IRGs).

Functional enrichment analysis

We conducted gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis on the Database for Annotation, Visualization and Integrated Discovery (DAVID) 6.8 [13] to explore the possible molecular mechanisms of the DE-IRGs. The terms of GO and KEGG with a threshold that false discovery rate (FDR) < 0.05 were considered significantly enriched.

Calculation and validation of the immune related signature of HCC

The HCC patients from TCGA were divided into a training set (TRS) and a testing set (TES). The TRS was used to calculate the prognostic immune related signature and to construct a prognostic immune related risk model of HCC, and the TES was used to valid the prognostic capability of this model [14]. We used a univariate Cox proportional hazard regression analysis to identify the relationship between DE-IRGs and overall survival in the TRS so that we could investigate the DE-IRGs related to prognosis for HCC patients. If $P < 0.05$, the corresponding DE-IRGs were considered as prognostic ones. To minimize overfitting and to find the best gene model, prognosis-related DE-IRGs were evaluated by the least absolute shrinkage and selection operator (LASSO) penalized Cox proportional hazards regression using glmnet package [15]. This method minimizes the residual sum of squares subject to the sum of the absolute value of the coefficients being less than a constant, because of which it tends to produce some coefficients that are exactly 0 and hence gives interpretable models. There should be two significant lambdas in the LASSO result, “min” one and “1se” one. The min lambda gives minimum cross-validated error while the 1se lambda is the largest value of lambda such that error is within 1 standard error of the minimum. For construction of a stable signature, genes with a coefficient > 0 under 1se lambda would be selected for further analysis. The risk score was calculated with the following model for each patient in the TRS, TES and total set: Risk score = exp (expression 1 * coefficient 1 + expression 2 * coefficient 2 + ... expression n * coefficient n) [16]. Then the patients were divided into high-risk and low-risk groups according to the median cutoff of the risk score. The area under the curve (AUC) was calculated both in high-risk and low-risk groups with survivalROC package to validate the prognostic capability of the immune related risk signature [17]. We plotted the Kaplan–Meier survival curves of the high-risk and low-risk groups using survival package [18], which demonstrated the overall survival of the patients.

Estimate of tumor-infiltrating immune cells

We used the Cancer Genome Atlas gene expression RNA-sequencing data to estimate the proportions of 22 types of infiltrating immune cells on the Cancer Immunome Database (TCIA, https://tcia.at/home), which provided results of comprehensive immunogenomic analyses of next generation sequencing data (NGS) data for 20 solid cancers from TCGA and other datasources [19].

Mutation analysis

The mutation data of TCGA HCC patients, if available, were downloaded from the TCGA data portal. Mutation data, which were stored in MAF form, were analyzed using maftools package [20]. And we calculated the tumor mutation burden (TMB) score of each HCC patient as follows: (total mutation/total covered bases) $\times 10^6$ [21].

Immunophenoscore analysis

The immunogenicity is determined by effector cells, immunosuppressive cells, MHC molecules, and immunomodulators adding up to four major categories of genes [22], through which machine learning can derived the Immunophenoscore of a patient without bias. We obtained the Immunophenoscores of patients with HCC from TCIA.

The main code used in this section can be founded on:
https://github.com/Ebonleo/pub_1/blob/main/rscript_1.r.
The code is available upon the request.

Results

Patient characteristics

As previously described, we obtained:

(1) 285 microarrays of samples including 214 tumor ones and 71 normal ones;
(2) clinical information and gene expression profiles of 377 HCC patients;

and (3) a list of IRGs that contained 2498 genes.

In order to remove system error from different platforms, the platform of the 4 datasets from GEO database is the same one (GPL570). The patients from TCGA were randomly divided into the TRS (n=188)

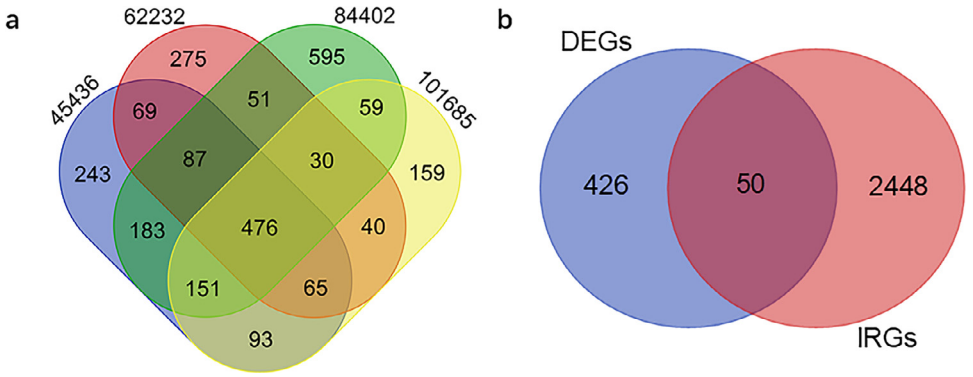

Fig. 2. Identification of differentially expressed immune-related genes. (a) Venn diagram for the intersections of HCC differentially expressed genes from GEO datasets. (b)Venn diagram for the intersections of HCC differentially expressed genes and immune-related genes.

Table 1
Clinical variables in the total, training and testing sets.

Variables	Group	Total set (n=377)	Training set (n=188)	Testing set (n=189)	P value	Method
Survival time (days)		821±729	862±760	781±696	0.278	Wilcoxon rank sum test
Vital status	Alive	245	123	122	0.859	Chi-squared test
	Dead	132	65	67		
Clinical stage	I	175	89	86	0.144	Chi-squared test
	II	87	35	52		
	III	86	45	41		
	IV	5	4	1		
	NA	24	15	9		
T stage	T1	185	98	87	0.316	Chi-squared test
	T2	95	39	56		
	T3	81	41	40		
	T4	13	8	5		
	TX	3	2	1		
N stage	N0	257	132	125	0.688	Chi-squared test
	N1	4	2	2		
	NX	115	54	61		
M stage	M0	272	136	136	0.581	Chi-squared test
	M1	4	3	1		
	MX	101	49	52		
Age (years)		59±14	60±13	59±14	0.678	t-test
Histological grade	G1	55	29	26	0.577	Chi-squared test
	G2	180	87	93		
	G3	124	63	61		
	G4	13	5	8		
	GX	5	4	1		

and the TES (n=189). There were no significant differences (α=0.05) of clinical variables between the two sets (Table 1).

Identification of DE-IRGs

Referring to the threshold in part 2.2, a total of 476 DEGs were identified, consisting of 149 up-regulated ones and 327 down-regulated ones. Then we extracted 50 DE-IRGs (Fig. 2). GO and KEGG enrichment analysis of the DE-IRGs were performed on DAVID. Top 6 enriched terms of biological process (BP), cellular component (CC), and molecular function (MF) are shown in Fig. 3 (a). The most enriched terms for BP, CC, and MF were immune response, extracellular region, and peptide hormone binding. The 6 pathways enriched in were cytokine-cytokine receptor interaction, staphylococcus aureus infection, pathways in cancer, neuroactive ligand-receptor interaction, complement and coagulation cascades, chemokine signaling pathway (Fig. 3 (b)).

Table 2
The univariate Cox regression analysis result with significance.

Gene	Beta value	HR (95% CI for HR)	Wald test	P value
BIRC5	0.22	1.3 (1.1-1.4)	10.0	0.001**
CCL14	-0.22	0.8 (0.70-0.93)	8.8	0.003**
FCN2	-0.10	0.9 (0.82-0.99)	4.3	0.038*
FYN	-0.35	0.7 (0.53-0.92)	6.4	0.012*
GHR	-0.24	0.79 (0.69-0.9)	12.0	<0.001***
IGF1	-0.15	0.86 (0.77-0.97)	6.6	0.010*
KLKB1	-0.16	0.85 (0.74-0.98)	5.1	0.024*
MASP1	-0.17	0.85 (0.75-0.96)	6.4	0.012*
MDK	0.14	1.10 (1.00-1.30)	4.5	0.034*
NR3C2	-0.34	0.71 (0.59-0.86)	13.0	<0.001***
TGFBR3	-0.20	0.82 (0.68-0.99)	4.4	0.035*
VIPR1	-0.17	0.85 (0.74-0.96)	6.6	0.010*

HR: hazard ratio.

Construction of immune related risk signature

We performed a univariate Cox regression analysis to investigate the prognostic value of the 50 DE-IRGs. 12 DE-IRGs were significantly associated with the overall survival of HCC patients in the TRS (Table 2). The 12 DE-IRGs underwent LASSO analysis to minimize overfitting, and 6 of the 12 DE IRGs were identified (Table 3). Then we used the 6 DE-IRGs to construct the immune signature (Table 4). The model for prediction was defined as a linear combination of the expression levels of the 6 DE-IRGs weighted by their relative coefficient in the multivariate Cox regression as follows: exp (0.07482 * BIRC5) + (-0.18477 * FYN) + (-0.03359 * IGF1) + (-0.09906 * MASP1) + (-0.21387 * NR3C2) + (-0.04448 * TGFBR3). One (BIRC5) of the 6 DE-IRGs were

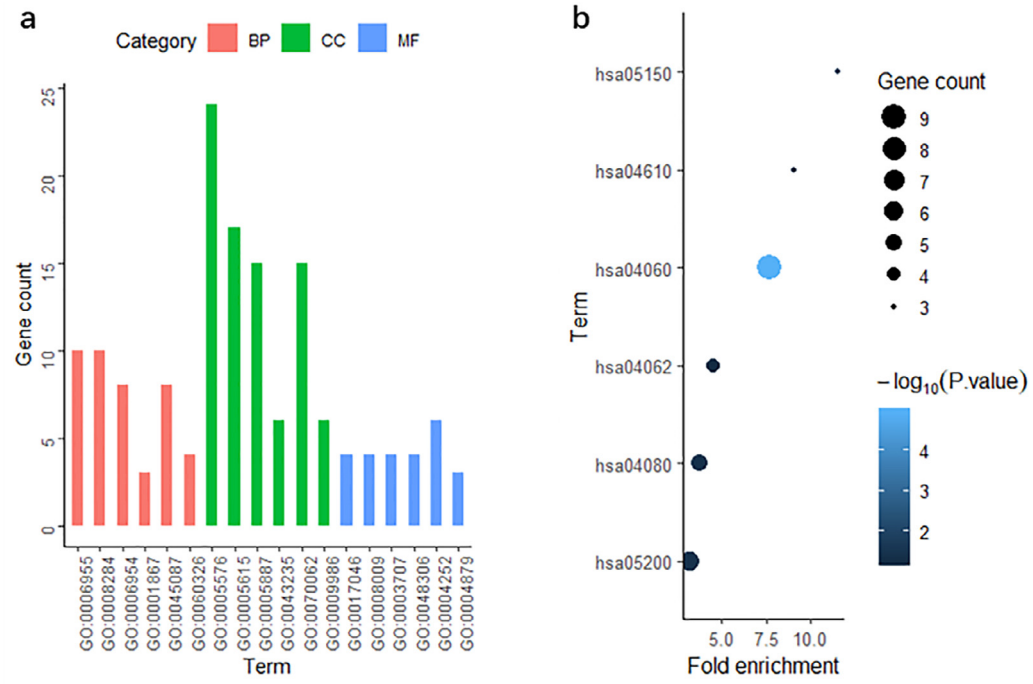

Fig. 3. Functional enrichment analyses of differentially expressed immune-related genes. (a) Gene ontology analysis. (b) The top six most significant Kyoto Encyclopedia of Genes and Genomes pathways.

Table 3
Genes with a coefficient >0 under "1se" lambda.

Gene	Coefficient
BIRC5	0.07893
FYN	-0.09003
IGF1	-0.10624
MASP1	-0.08650
NR3C2	-0.02509
TGFBR3	-0.00082

Table 4
Coefficients and multivariable Cox model results of each gene in 6-IRG risk signature.

Gene symbol	Log FC	Regulation	Coefficient	HR	Z score	P value
BIRC5	1.416	Up	0.07482	1.08	0.798	0.425
FYN	-1.440	Down	-0.18477	0.83	-1.190	0.234
IGF1	-2.304	Down	-0.03359	0.97	-0.499	0.618
MASP1	-1.237	Down	-0.09906	0.91	-1.242	0.214
NR3C2	-1.692	Down	-0.21387	0.81	-1.877	0.061
TGFBR3	-1.250	Down	-0.04448	0.96	-0.415	0.678

HR: hazard ratio; FC: fold change.

linked to high risk and five (FYN, IGF1, MASP1, NR3C2, TGFBR3) were protective ones.

We calculated the risk scores for each patient in the TRS based on the above model. Then the patients were divided into high-risk (n=94) and low-risk (n=94) groups according to part 2.4. There was a significant difference between the two risk groups (P<0.001) (Fig. 4 (a)). High-risk patients had a poorer overall survival than those in the low-risk group. The AUC of the 6-IRG risk signature was 0.71 at 5-year survival (Fig. 4 (d)). The risk scores of the patients in the TRS were ranked, and we analyzed their distribution in Fig. 4 (g). The heatmap revealed expression comparison of 6 DE-IRGs between two risk groups (Fig. 5).

Evaluating the predictive capability of the 6-IRG risk signature

In order to valid the stability of the 6-IRG signature, we further verified its predictive capability in the TES and the total set. The risk scores of each patient were calculated, according to which the patients were divided into high-risk and low-risk groups in the two sets. There were 94 patients from the TES in the high-risk group and 94 in the low-risk group. Similarly, Kaplan-Meier survival curve showed significant difference between the two risk groups (P = 0.026), and the overall survival of the high-risk patients was lower than that of the low-risk ones. In the TES, the AUC was 0.655, and the distribution of the risk score, survival status, and expression heatmap of 6 DE-IRGs were presented in Fig. 4 and Fig. 5.

The results in the total set were similar to the previous two sets (Fig. 4). There were 188 patients from the total set were in the high-risk group and 189 in the low-risk group. According to the Kaplan-Meier survival curves, it was significantly different in the two risk groups (P<0.001). In the total set, the AUC was 0.675, and the distribution of risk score, survival status, and expression heatmap of 6 DE-IRGs for CC patients in the total set were showed in Fig. 5.

Relationship between the 6-IRG risk signature and the prognosis

A univariate Cox regression was performed to analyze the relationship between overall survival, progression-free interval, clinicopathological factors, and 6-IRG risk signature in the total set (Table 5). This signature was able to act as an independent prognosis variable of HCC in the total set referring to the multivariate analysis (P <0.001, Table 6). So was it in progression-free interval prediction (P = 0.0179, Table 5).

Association between the 6-IRG risk signature and clinicopathological factors

We analyzed the association between the immune signature and clinicopathological variables. Significant differences were observed in tumor burden, T stage, grade and clinical stage. However, the 6-IRG risk score is not significant in M stage, and N stage (Fig. 6).

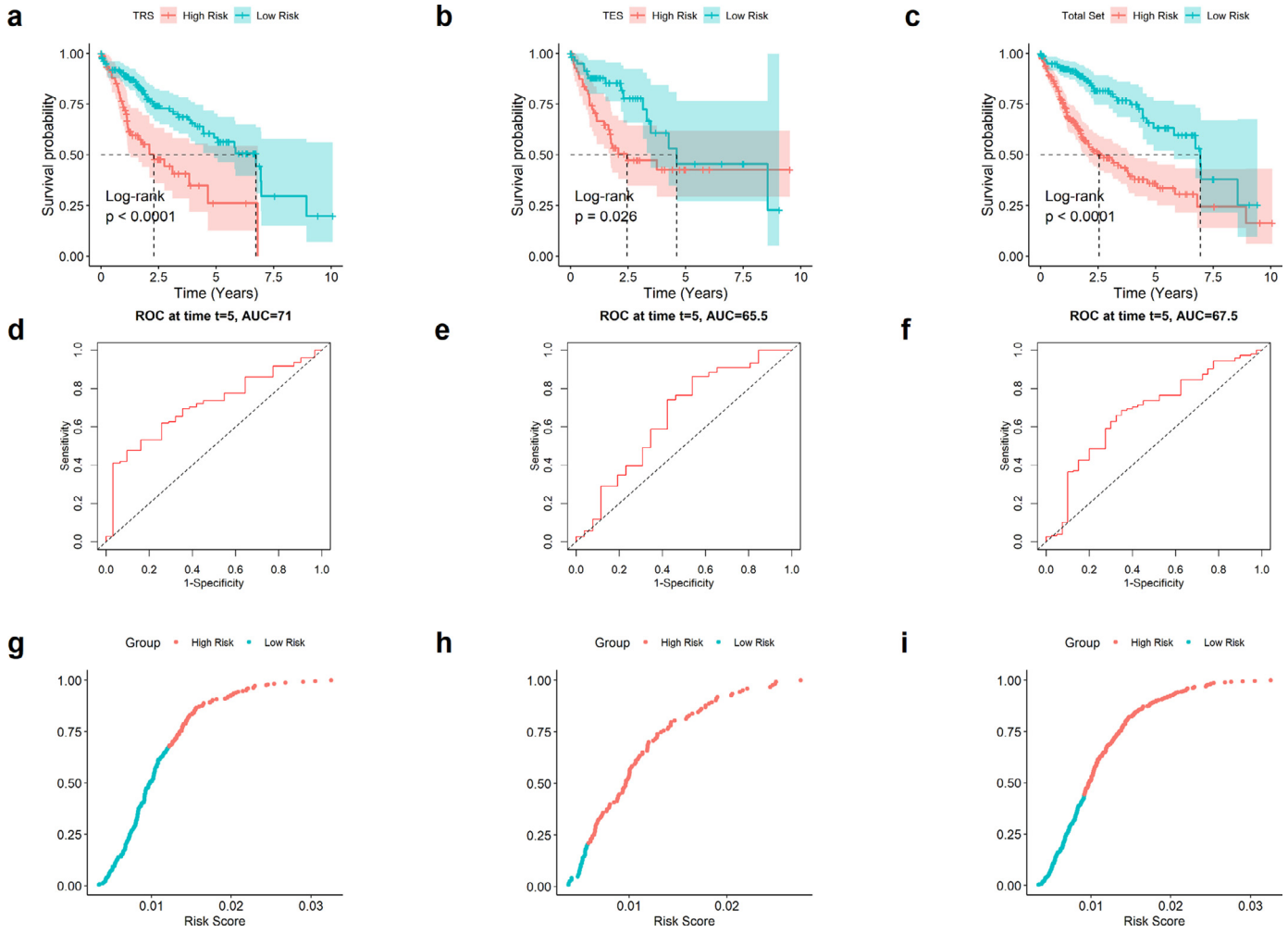

Fig. 4. Identification of the 6-IRG signature in the 3 sets. (a) Kaplan-Meier curve analysis of overall survival of HCC patients in TRS. (b) Kaplan-Meier curve analysis of overall survival of HCC patients in TES. (c) Kaplan-Meier curve analysis of overall survival of HCC patients in the total set. (d)Time-dependent ROC curves analysis of TRS. (e)Time-dependent ROC curves analysis of TES. (f)Time-dependent ROC curves analysis of the total set. (g)Risk score distribution in TRS. (h)Risk score distribution in TES. (i)Risk score distribution in the total set.

Table 5
Univariate analysis with Cox proportional hazard model.

	Overall survival		Progression-free interval	
Covariates	HR (95% CI for HR)	P value	HR (95% CI for HR)	P value
Age	1.02 (0.90-1.23)	0.045*	1.00 (0.93-1.20)	0.352
Histological grade				
G1/2	1.00		1.00	
G3/4	1.13 (1.03-1.23)	0.526	1.16 (1.10-1.19)	0.335
Clinical stage				
Stage I/II	1		1.00	
Stage III/IV	2.44 (2.31-2.64)	0.526	2.29 (2.22-2.38)	<0.001***
N stage (N0 vs. N1)	2.29 (2.09-2.39)	0.250	1.45 (1.37-1.54)	0.603
M stage (M0 vs. M1)	3.39 (3.26-3.54)	0.089	5.14 (0.91-5.20)	0.006
T stage				
T1/2	1		1	
T3/4	2.53 (2.44-2.62)	<0.001***	2.28 (2.18-2.38)	<0.001***
Risk score (high vs. low)	0.53 (0.44-0.63)	0.001**	0.67 (0.62-0.73)	0.010*

HR: hazard ratio.

Tumor immune microenvironment changing associated with the 6-IRG risk signature

Immune cells play the main roles of tumor immune microenvironment. Thus, we tried to find out which classes of immune cells were linked to the 6-IRG risk signature. TCIA could assess the relative proportion of the 22-type immune cells according to the RNA-sequencing data. Table 6 presented the immune cell type abundance between the 6-IRG signature low-risk group and the high-risk group from the total set. Immune cell type abundance between high risk and low risk group

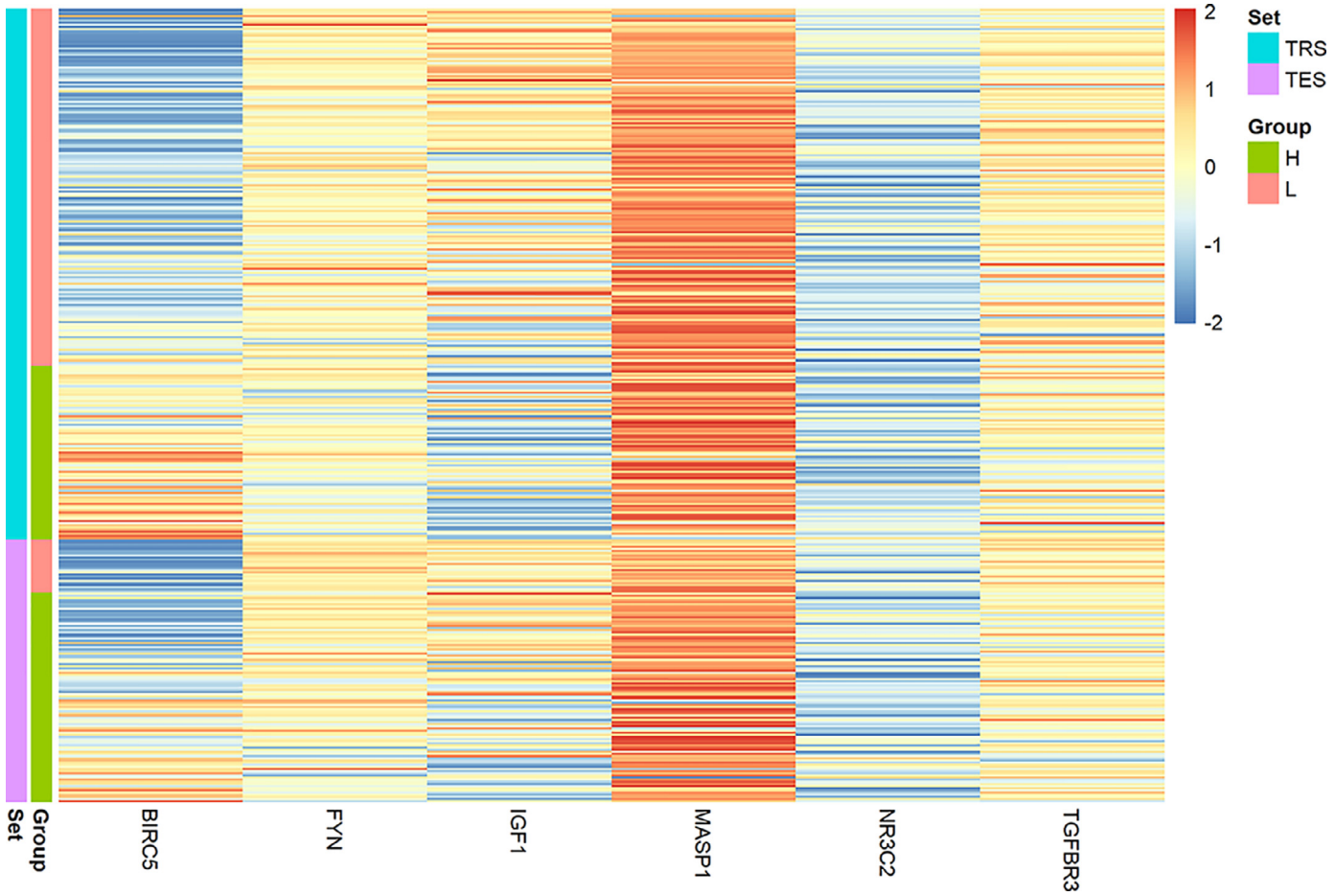

Fig. 5. The heatmap of 6 IRGs between two groups in TRS, TES and the total sets. H stands for high group and L stands for low group.

Table 6
Multivariate analysis with Cox proportional hazard model.

	Overall survival		Progression-free interval	
Covariates	HR	P value	HR	P value
Age	1.02 (0.90-1.10)	0.059	1.00 (0.93-1.10)	0.304
Histological grade				
G1/2	1.00		1.00	
G3/4	1.12 (1.02-1.29)	0.474	1.16 (10.9-1.26)	0.516
Clinical stage				
Stage I/II	1.00		1.00	
Stage III/IV	0.18 (0.10-0.26)	0.249	1.24 (1.16-1.35)	0.854
N stage (N0 vs. N1)	6.39 (6.32-6.45)	0.087	0.81 (0.73-0.91)	0.853
M stage (M0 vs. M1)	1.58 (1.50-1.69)	0.539	2.92 (2.83-2.99)	0.089
T stage				
T1/2	1.00		1.00	
T3/4	15.06 (14.96-15.19)	0.068	1.24 (1.17-1.30)	0.854
Risk score (high vs. low)	0.46	<0.001***	0.70 (0.59-.079)	0.036*

HR: hazard ratio.

of the 6-IRG signature from the total set are displayed in Table 7. Among the 22 immune cell types, CD8 T cells, macrophage M0, memory B cells, regulatory T cells, and T follicular helper cells were positively correlated with the risk score, while gamma delta T cells, macrophage M1, naive B cells, resting mast cells, and resting memory CD4 T cells were negatively associated with the risk score (Fig. 7). Moreover, after grouped according to the median of the fraction, the relative proportion of activated mast cells, macrophage M0, neutrophil, and regulatory T cells were significantly correlated with overall survival (Fig. 8).

The immune related risk signature and mutation profile

We assessed the association between mutation profile and the signature of the HCC patients. The top 10 mutated genes in the high-risk group were TP53, CTNNB1, RYR2, ARID1A, LRP1B, RB1, CSMD3, USH2A, XIRP2, ABCA13. And those in the low-risk group were CTNNB1, TP53, APOB, ABCA13, ALB, PCLO, BIRC6, CSMD1, DNAH6 and FAT2. And the TMB in the high-risk group was higher than that in the low-risk group in spite of no significance. Moreover, there was no significant rela-

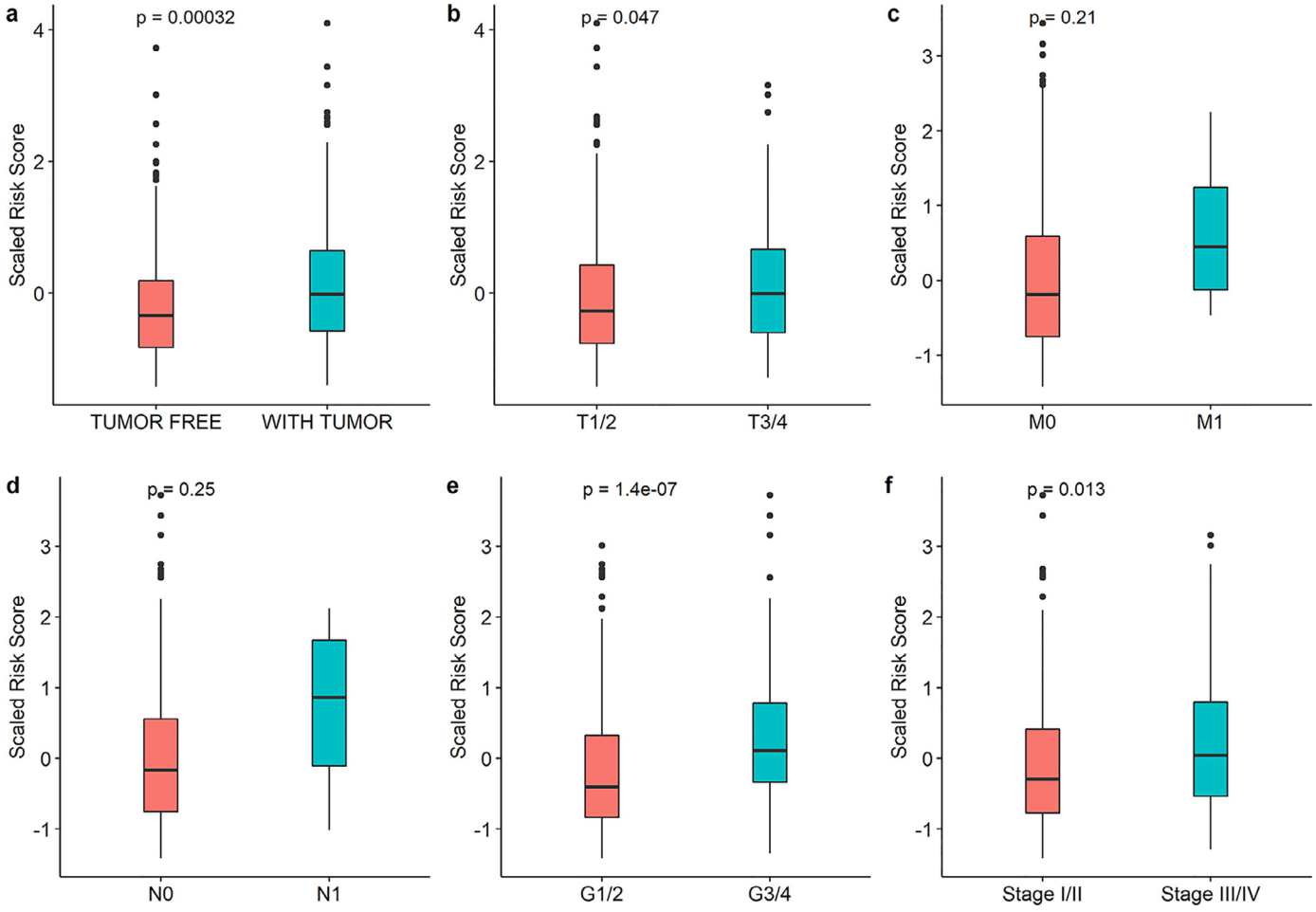

Fig. 6. The relationships between the signature and (a) tumor burden; (b) T stage; (c) M stage; (d) N stage; (e) grade; and (f) clinical stage.

Table 7
Immune cell type abundance between the 6-IRG signature high-risk group and low-risk group in the total set.

	Abundance		
Immune cell type	High risk	Low risk	P value
Activated dendritic cells	0.020	0.022	0.463
Activated mast cells	0.009	0.010	0.179
Activated memory CD4 T cells	0.008	0.006	0.356
Activated Natural killer cells	0.001	0.002	0.192
CD8 T cells	0.039	0.025	0.001**
Eosinophil	0.000	0.000	0.571
Gamma delta T cells	0.111	0.124	0.040*
Macrophage M0	0.088	0.048	<0.001***
Macrophage M1	0.067	0.080	0.003**
Macrophage M2	0.170	0.186	0.110
Memory B cells	0.025	0.016	0.001**
Monocyte	0.011	0.015	0.101
Naive B cells	0.019	0.021	0.006**
Naive CD4 T cells	0.001	0.007	0.051
Neutrophil	0.023	0.020	0.280
Plasma cells	0.046	0.042	0.403
Regulatory T cells	0.038	0.018	<0.001***
Resting Dendritic cells	0.031	0.030	0.763
Resting mast cells	0.044	0.059	0.005**
Resting memory CD4 T cells	0.124	0.157	0.002**
Resting Natural killer cells	0.059	0.058	0.884
T follicular helper cells	0.064	0.055	0.005**

The data are presented as mean. Method: Wilcoxon rank sum test.

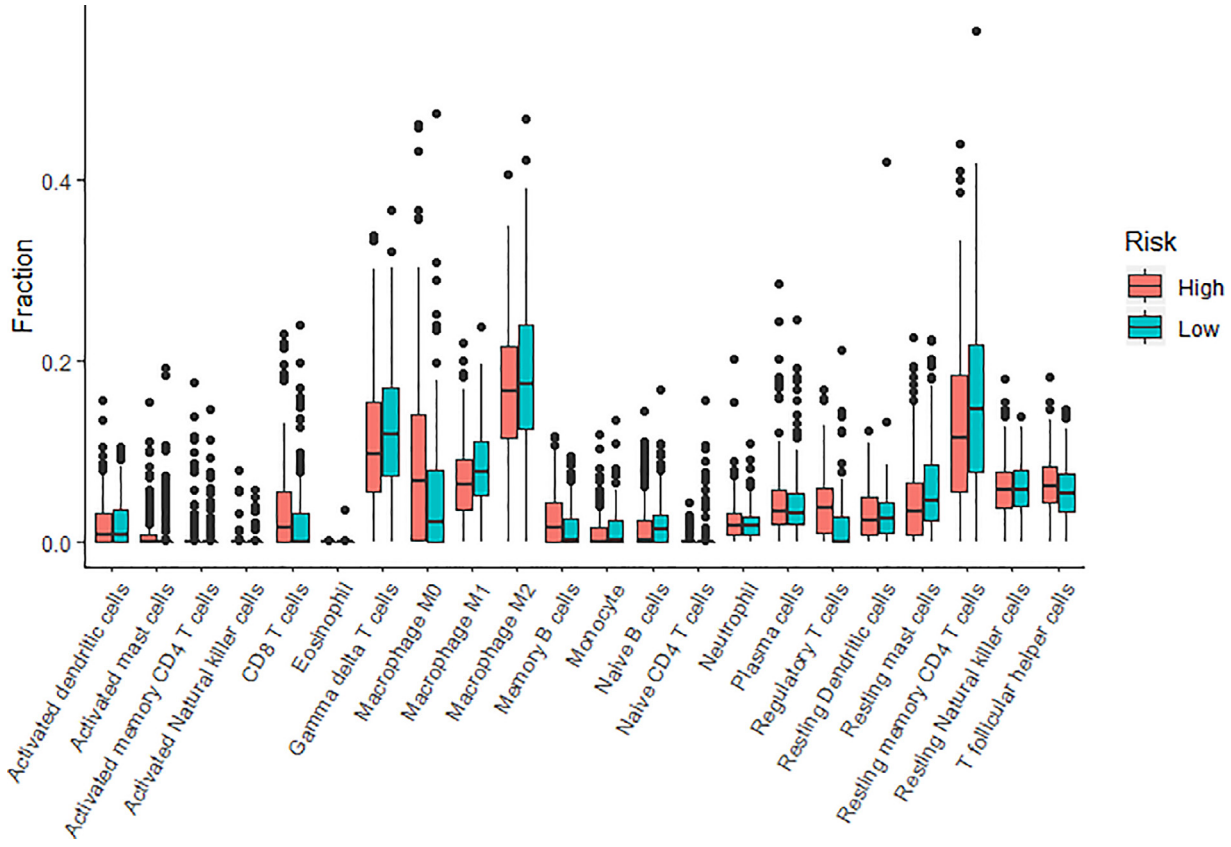

Fig. 7. The association of immune cells infiltration and the signature in HCC. Warm-colored boxplots represent the 6-IRG signature high-risk group. Cold-colored boxplots represent the 6-IRG signature low-risk group

tionship between TMB and overall survival. Neither was there between TMB and progression-free interval (Fig. 9).

The immune related risk signature and response to ICI

Recent studies have reported the roles that immunophenoscore played in the prediction of the response to ICI of melanoma patients. This is on the basis of the high pre-existing immunogenic potential. The relationship between IPS and the 6-IRG risk signature was investigated in our study (Fig. 10). We used the IPS, IPS-PD1/PD-L1/PD-L2, IPS-CTLA4, and IPS-PD1/PD-L1/PD-L2 + CTLA4 to assess the potential of ICI application for HCC. The IPS and IPS-PD1/PD-L1/PD-L2 were significantly higher in the low-risk group ($P<0.05$). Meanwhile, the expression of PD-L1 and PD-L2 was significantly higher in the low-risk group. The results suggested that 6-IRG signature low-risk patients had a better opportunity for ICI application.

Discussion

There has been evidence that ICI has potential in the treatment of HCC. However, only a small proportion of the patients benefit from it. So, it's important to identify the biomarkers for ICI. One-biomarker-one-disease style has gradually been abandoned, which is unable to meet the requirements of precision medicine. On the contrary, accurate prediction calls for several indicators, genomics and transcriptomics included. In addition, an immune-related model, according to the roles that tumor immune environment play in cancers, contributes to identification of the patients who can benefit from immunotherapy.

The prognostic value of this signature

We used the data from GEO database and TCGA to construct and to validate the IRG risk signature, which consisted of six prognosis DE-IRGs. One (BIRC5) of the six genes was related to high risk, while five (FYN, IGF1, MASP1, NR3C2 and TGFBR3) were protective factors. Compared to the normal samples, 149 genes were up-regulated in the tumor samples, and 327 were down-regulated.

It's reported that all of the 6 genes play important roles in development and prediction in cancers. As a direct target, NR3C2 is involved in proliferation and metastasis of HCC under the regulation of miR-766. In our study, NR3C2 was down-regulated in the HCC tumor samples compared to the normal samples [23].Migration inhibitory factor up-regulated miR-301b that targeted NR3C2, reduced expression levels of which correlated with poorer survival in multiple independent cohorts of pancreatic cancer patients, and suppressed its expression mechanistically [24].Fyn is a member of the Src family kinases (SFKs) which is involved in signal transduction pathways in the nervous system, and the development and activation of T cells under normal physiological conditions [25]. Fyn controls the cell growth, death, morphogenic transformation and cellular motility resulting in the development and progression of several cancer types. Enhanced expression and/or activation of Fyn is observed in various cancers, including breast cancers, glioblastoma, melanoma, prostate, and squamous cell carcinoma. The importance of Fyn in the resistance or susceptibility of cancer cells to some anti-cancer treatments have been demonstrated in recent studies. Fyn, which plays a key role in the resistance mechanism, is up-regulated in tamoxifen-resistant breast cancer cell lines. There is prognostic significance of IGF-1 signaling pathway in patients with advanced NSCLC. It's found that reduction of IGF-1/IGF-BP3 ratio was statistically sig-

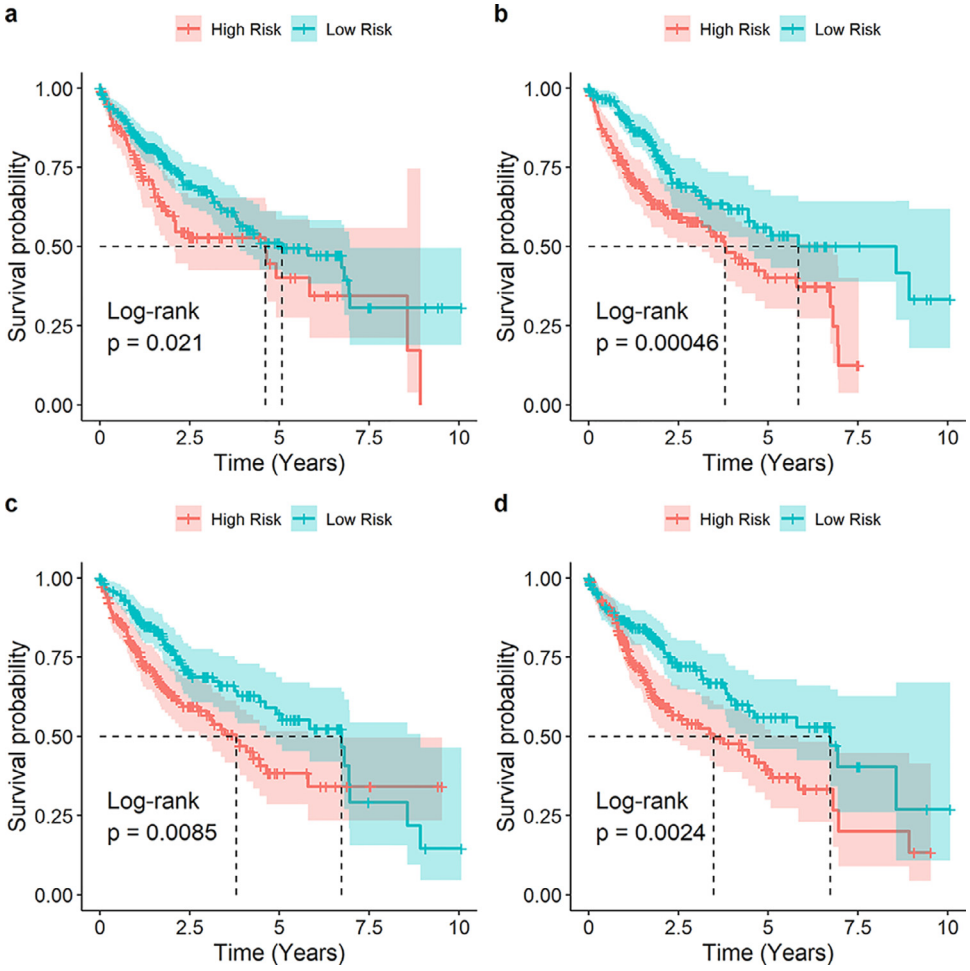

Fig. 8. The association of immune cells infiltration and overall survival in TCGA HCC dataset. (a) activated mast cells; (b) macrophage M0; (c) neutrophil, activated; (d) regulatory T cells.

nificant only among patients with NSCLC who responded to first-line treatment. IGF-1/IGFBP3 has opportunities to become a predictive tool for response to chemotherapy in NSCLC if validated in larger cohorts. Moreover, studies provide support that paracrine IGF1/IGF1R signaling promotes colorectal cancer progression [26, 27].Regulation of TFGBR3 contributes to the development of pancreatic ductal adenocarcinoma and metastasis of clear-cell renal cell carcinoma [28].The expression of MASP-1 is down-regulated in the tumor samples compared to normal samples from GEO datasets. It's reported that MASP-1 serum levels are associated with worse prognostic in cervical cancer progression [29].A meta-analysis reveals that several polymorphisms of BIRC5, a member of apoptosis inhibition gene family, might cause the different risk of urinary cancer [30].

Thus, the genes may reflect the changes of tumor immune environment in high-risk HCC patients, contributing to the progression of HCC and poor survival. Due to the obvious difference in the survival curves, the signature revealed outstanding predictive power. Furthermore, the signature was significantly associated with the prognosis of the patients, including overall survival and progression-free interval. When linked with clinicopathologic factors, the signature always played an independent prognostic role in the multivariate Cox regression. The results proved that the signature is a reliable predictive tool. The relationships between the signature and several pathological factors was validated, such as tumor grade and stage. Moreover, we estimated the relative proportions of 22-type immune cells each HCC sample.

The abundance of macrophage M0 and regulatory T cells were high in the high-risk group, which was associated with better survival. The relationship between the HCC prognosis and the infiltration of macrophage M0 and regulatory T cells is constructed.

This seemed an explanation for the predictive power of the signature. In order to explore more probable mechanisms of the signature, we performed mutation an analysis. There were differences in the mutation profiles between the two groups. The TMB was higher in the low-risk group, though without significance, which overturned our assumption that the TMB should been higher in the high-risk group. Even, the TMB showed no significant relationship with overall survival and progression-free interval, different from previous studies [31, 32].

Prediction of ICI response

We explored the association between the IPS and the signature. The IPS, IPS-PD1/PD-L1/PD-L2, IPS-CTLA4, and IPS-PD1/PD-L1/PD-L2 + CTLA4 markedly increased in the 6-IRG signature low-risk group. And the expression of PD-L1 and PD-L2 was significantly higher in the low-risk group. It's suggested that the signature may represent the immunogenic tumor microenvironment of HCC. All above results support that the signature can predict which patients will benefit from ICI. Additionally, there were studies reporting that PD-L1 expression and TMB might play the similar roles [33–37], though the results were negative in several other studies [38–40]. In this study, significant differences of TMB were not observed between the risk groups. Therefore, we haven't fully understood the relevance of the signature and ICI response.

We established a firm immune related risk signature which can estimate the tumor immune environment and predict the survival and response to ICI of HCC patient. As is known to us, this is the first study of immune-related predictive model based on RNA-seq, which estimated the relative proportion of 22-type immune cells of HCC data from TCGA. This model was also used for exploration of the association among TMB,

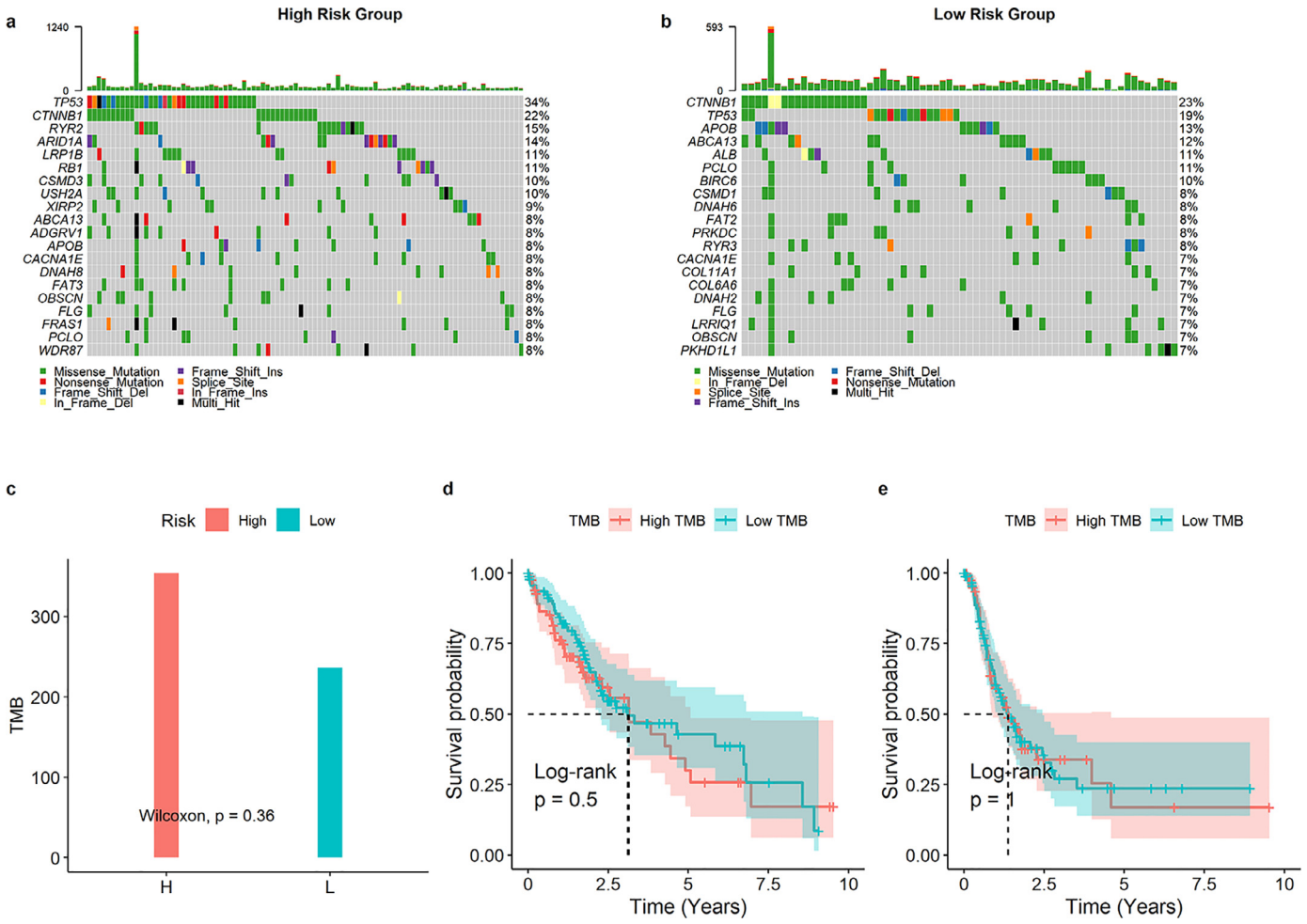

Fig. 9. The mutation profile and TMB among high-risk and low-risk groups. (a) Mutation profile of high-risk groups; (b) Mutation profile of low-risk groups; (c) The relationship between the immune related risk signature and TMB; (d) The association of TMB and overall survival and progression-free interval in TCGA HCC dataset.

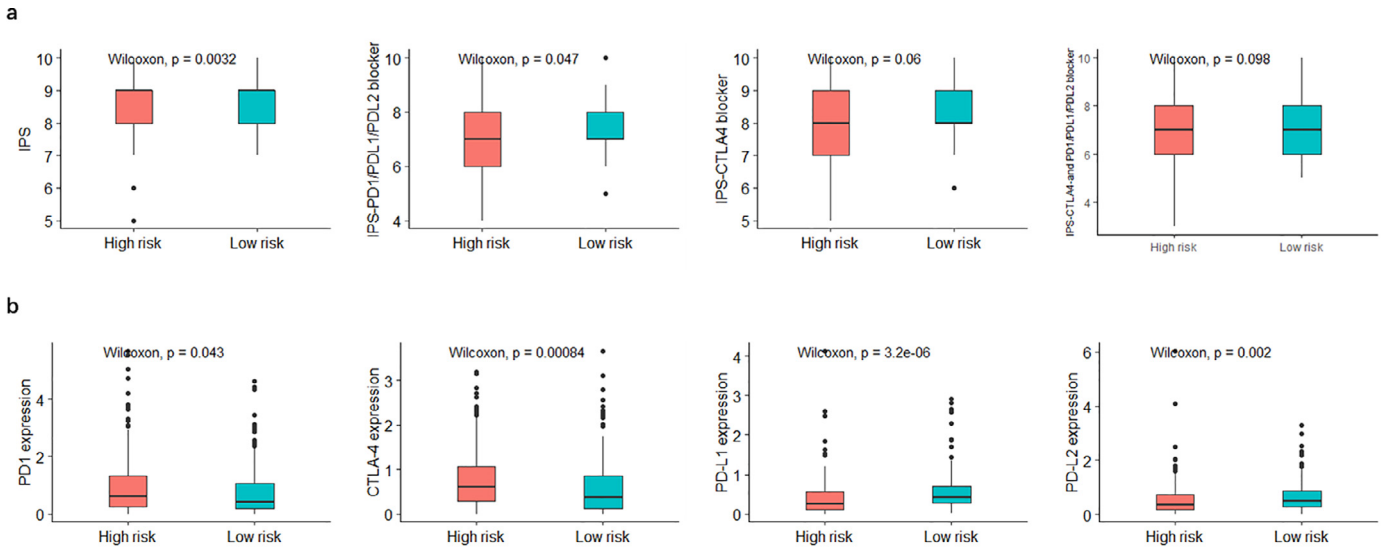

Fig. 10. IPS and immunotherapy gene expression analysis. (a) The association between IPS and the immune related risk signature of HCC patients. (b) The gene expression of PD1, CTLA4, PD-L1, and PD-L2 in high-risk and low risk groups.

abundance of the immune cells, and prognosis. The signature was able to predict the survival and progression of HCC. In spite of the inspiring results, there remains problems. First above all, the signature was established on limited data from retrospective studies. Secondly, the proportion of several races was small, which made it unsure the signature would function accurately on all patients of different races. More data of patients of different races is necessary in further research.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Ying Xu: Project administration, Supervision, Writing - original draft. **Zheng Wang:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Supervision, Visualization. **Fufeng Li:** Funding acquisition, Writing - review & editing.

Acknowledgments

No grants or other financial support was received for this work.

Data Availability statement

The data used to support the findings of this study are included within the article.

Ethical approval

All analyses were based on previous published studies. Thus, no ethical approval and patient consent are required.

Fund Support

Shanghai Sailing Program (20yf1449500); National Natural Science Foundation of China (No. 82004162); China Postdoctoral Science Foundation (No. 2020M681367).

Reference

[1] F Bray, J Ferlay, I Soerjomataram, RL Siegel, LA Torre, A Jemal, Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA Cancer J. Clin. 68 (6) (2018) 394–424.
[2] NP Malek, S Schmidt, P Huber, MP Manns, TF Greten, The diagnosis and treatment of hepatocellular carcinoma, Dtsch Arztebl Int. 111 (7) (2014) 101–106.
[3] J Long, A Wang, Y Bai, J Lin, X Yang, D Wang, et al., Development and validation of a TP53-associated immune prognostic model for hepatocellular carcinoma, EBioMedicine 42 (2019) 363–374.
[4] Q Song, J Shang, Z Yang, L Zhang, C Zhang, J Chen, et al., Identification of an immune signature predicting prognosis risk of patients in lung adenocarcinoma, J. Transl. Med. 17 (1) (2019) 70.
[5] S Shen, G Wang, R Zhang, Y Zhao, H Yu, Y Wei, et al., Development and validation of an immune gene-set based Prognostic signature in ovarian cancer, EBioMedicine 40 (2019) 318–326.
[6] J Wu, Y Zhao, J Zhang, Q Wu, W Wang, Development and validation of an immune-related gene pairs signature in colorectal cancer, Oncoimmunology 8 (7) (2019) 1596715.
[7] HW Wang, TH Hsieh, SY Huang, GY Chau, CY Tung, CW Su, et al., Forfeited hepatogenesis program and increased embryonic stem cell traits in young hepatocellular carcinoma (HCC) comparing to elderly HCC, BMC Genomics 14 (2013) 736.
[8] K Schulze, S Imbeaud, E Letouze, LB Alexandrov, J Calderaro, S Rebouissou, et al., Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets, Nat. Genet. 47 (5) (2015) 505–511.
[9] H Wang, X Huo, XR Yang, J He, L Cheng, N Wang, et al., STAT3-mediated upregulation of lncRNA HOXD-AS1 as a ceRNA facilitates liver cancer metastasis by regulating SOX4, Mol. Cancer 16 (1) (2017) 136.
[10] E Clough, T. Barrett, The gene expression omnibus database, Methods Mol. Biol. 1418 (2016) 93–110.
[11] S Bhattacharya, S Andorf, L Gomes, P Dunn, H Schaefer, J Pontius, et al., ImmPort: disseminating data to the public for the future of immunology, Immunol. Res. 58 (2-3) (2014) 234–239.
[12] ME Ritchie, B Phipson, D Wu, Y Hu, CW Law, W Shi, et al., limma powers differential expression analyses for RNA-sequencing and microarray studies, Nucleic Acids Res. 43 (7) (2015) e47.
[13] da W Huang, BT Sherman, RA Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources, Nat. Protoc. 4 (1) (2009) 44–57.
[14] Z Wang, Q Song, Z Yang, J Chen, J Shang, W Ju, Construction of immune-related risk signature for renal papillary cell carcinoma, Cancer Med. 8 (1) (2019) 289–304.
[15] T Hastie, B. Efron, lars: least angle regression, Lasso and Forward Stagewise (2013).
[16] HY Chen, SL Yu, CH Chen, GC Chang, CY Chen, A Yuan, et al., A five-gene signature and clinical outcome in non-small-cell lung cancer, N. Engl. J. Med. 356 (1) (2007) 11–20.
[17] M Lorent, M Giral, Y Foucher, Net time-dependent ROC curves: a solution for evaluating the accuracy of a marker to predict disease-related mortality, Stat. Med. 33 (14) (2014) 2379–2389.
[18] B Holleczek, H. Brenner, Model based period analysis of absolute and relative survival with R: data preparation, model fitting and derivation of survival estimates, Comput. Methods Programs Biomed. 110 (2) (2013) 192–202.
[19] AM Newman, CL Liu, MR Green, AJ Gentles, W Feng, Y Xu, et al., Robust enumeration of cell subsets from tissue expression profiles, Nat. Methods 12 (5) (2015) 453–457.
[20] A Mayakonda, DC Lin, Y Assenov, C Plass, HP Koeffler, Maftools: efficient and comprehensive analysis of somatic variants in cancer, Genome Res. 28 (11) (2018) 1747–1756.
[21] DR Robinson, YM Wu, RJ Lonigro, P Vats, E Cobain, J Everett, et al., Integrative clinical genomics of metastatic cancer, Nature 548 (7667) (2017) 297–303.
[22] P Charoentong, F Finotello, M Angelova, C Mayer, M Efremova, D Rieder, et al., Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade, Cell Rep. 18 (1) (2017) 248–262.
[23] C Yang, X Ma, G Guan, H Liu, Y Yang, Q Niu, et al., MicroRNA-766 promotes cancer progression by targeting NR3C2 in hepatocellular carcinoma, Faseb J. 33 (1) (2019) 1456–1467.
[24] S Yang, P He, J Wang, A Schetter, W Tang, N Funamizu, et al., A novel MIF signaling pathway drives the malignant character of pancreatic cancer by targeting NR3C2, Cancer Res. 76 (13) (2016) 3838–3850.
[25] D Elias, HJ. Ditzel, Fyn is an important molecule in cancer pathogenesis and drug resistance, Pharmacol Res. 100 (2015) 250–254.
[26] I Kotsantis, P Economopoulou, A Psyrri, E Maratou, D Pectasides, H Gogas, et al., Prognostic significance of IGF-1 signalling pathway in patients with advanced non-small cell lung cancer, Anticancer Res. 39 (8) (2019) 4185–4190.
[27] J Tommelein, E De Vlieghere, L Verset, E Melsens, J Leenders, B Descamps, et al., Radiotherapy-activated cancer-associated fibroblasts promote tumor progression through paracrine IGF1R activation, Cancer Res. 78 (3) (2018) 659–670.
[28] J Nishida, K Miyazono, S Ehata, Decreased TGFBR3/betaglycan expression enhances the metastatic abilities of renal cell carcinoma cells through TGF-beta-dependent and -independent mechanisms, Oncogene 37 (16) (2018) 2197–2212.
[29] CA Maestri, R Nisihara, HW Mendes, J Jensenius, S Thiel, I Messias-Reason, et al., MASP-1 and MASP-2 serum levels are associated with worse prognostic in cervical cancer progression, Front. Immunol. 9 (2018) 2742.
[30] M Xu, X Hu, M Zhang, Y Ge, What is the impact of BIRC5 gene polymorphisms on urinary cancer susceptibility? Evidence from 9348 subjects, Gene (2019) 144268.
[31] AM Goodman, S Kato, L Bazhenova, SP Patel, GM Frampton, V Miller, et al., Tumor mutational burden as an independent predictor of response to immunotherapy in diverse cancers, Mol. Cancer Ther. 16 (11) (2017) 2598–2608.
[32] F Innocenti, FS Ou, X Qu, TJ Zemla, D Niedzwiecki, R Tam, et al., Mutational analysis of patients with colorectal cancer in CALGB/SWOG 80405 identifies new roles of microsatellite instability and tumor mutational burden for patient outcome, J. Clin. Oncol. 37 (14) (2019) 1217–1227.
[33] DB Johnson, GM Frampton, MJ Rioth, E Yusko, Y Xu, X Guo, et al., Targeted next generation sequencing identifies markers of response to PD-1 blockade, Cancer Immunol. Res. 4 (11) (2016) 959–967.
[34] G Singal, PG Miller, V Agarwala, G Li, G Kaushik, D Backenroth, et al., Association of patient characteristics and tumor genomics with clinical outcomes among patients with non-small cell lung cancer using a clinicogenomic database, Jama 321 (14) (2019) 1391–1399.
[35] AM Goodman, S Kato, R Chattopadhyay, R Okamura, IM Saunders, M Montesion, et al., Phenotypic and genomic determinants of immunotherapy response associated with squamousness, Cancer Immunol. Res. 7 (6) (2019) 866–873.
[36] RE Vilain, AM Menzies, JS Wilmott, H Kakavand, J Madore, A Guminski, et al., Dynamic changes in PD-L1 expression and immune infiltrates early during treatment predict response to PD-1 blockade in melanoma, Clin. Cancer Res. 23 (17) (2017) 5024–5033.
[37] MD Hellmann, T Nathanson, H Rizvi, BC Creelan, F Sanchez-Vega, A Ahuja, et al., Genomic features of response to combination immunotherapy in patients with advanced non-small-cell lung cancer, Cancer Cell 33 (5) (2018) 843–852.
[38] Y Owada-Ozaki, S Muto, H Takagi, T Inoue, Y Watanabe, M Fukuhara, et al., Prognostic impact of tumor mutation burden in patients with completely resected non-small cell lung cancer: brief report, J. Thorac. Oncol. 13 (8) (2018) 1217–1221.
[39] W Hugo, JM Zaretsky, L Sun, C Song, BH Moreno, S Hu-Lieskovan, et al., Genomic and transcriptomic features of response to anti-PD-1 therapy in metastatic melanoma, Cell 165 (1) (2016) 35–44.
[40] GT Gibney, LM Weiner, MB Atkins, Predictive biomarkers for checkpoint inhibitor-based immunotherapy, Lancet Oncol. 17 (12) (2016) e542–ee51.