Diagnostic Accuracy Study

Medicine®

OPEN

Study on the gene signature related to immune microenvironment on viral and nonviral infections of hepatocellular carcinoma

Hong Qian, MD[a], Shanglin Song, MD[a], Yunling Tian, MD[a], Lei Qi, MD[b,*], Haihong Gao, MD[c], Lingyan Yuan, PhD, MD[a,d,e]

Abstract

The pathogenesis of hepatocellular carcinoma (HCC) can be divided into viral infection (VIR) and nonviral (NVIR) infection. Two types of HCC performed different tumor immune microenvironment (TIME) which directly affected prognosis of HCC. This study aimed to identify an effective 2 types of HCC prognostic gene signature that related to immune TIME.

The differential expression genes (DEGs) were analyzed by Limma R package from the Cancer Genome Atlas. Immune related genes getting from IMMport database were matched to DEGs for testing prognosis. Prognostic index (PI) consisted of prognostic immune related genes was calculated in different types of HCC by COX regression and the correlation with the abundance of immune infiltrates, including 6 type cells, via gene modules. Tumor immune estimation resource database was applied to analyze TIME. Finally, the correlations between PI of DEGs and TIICs were analyzed by the Spearman method.

Results showed that PI consisted of 11 messenger RNAs in VIR and 12 messenger RNAs in NVIR groups. The PI related to HCC prognosis has different correlations with immune infiltrating cells in VIR and NVIR groups. The PI value of DEGs has significant correlations with neutrophils ($R=0.22$, P-value$=.029$) and dendritic ($R=0.21$, P-value$=.036$) infiltration levels in VIR group. However, in NVIR group, the result showed there were no significant correlations between PI and other 5 type cell infiltration levels (P-value$>.05$).

The 11-gene signature in VIR and 12-gene signature in NVIR group selected based on data from the Cancer Genome Atlas database had a different correlation with immune infiltrating cells of HCC patients.

Abbreviations: AUC = area under the curve, HCC = hepatocellular carcinoma, NVIR = nonviral infection, ROC = receiver operating characteristic, TCGA = the Cancer Genome Atlas, TIICs = tumor-infiltrating immune cells, VIR = viral infection.

Keywords: gene signature, hepatocellular carcinoma, prognosis, tumor microenvironment

1. Introduction

Hepatocellular carcinoma (HCC) is the second most common cause of cancer-related death world-wide, and its incidence is expected to rise in the future.[1] More than 50% of global incidence and mortality of HCC occurs in China. Chronic inflammation is associated with HCC risk factors including hepatitis B virus (HBV), hepatitis C virus (HCV), and metabolic disorders such as nonalcoholic fatty liver disease, promotes an immunosuppressive environment and T-cell exhaustion.[2] The progression of HCC in some cases follows the 3 steps of hepatitis, cirrhosis, HCC.[3] Unfortunately, HCC incidence is currently rising in almost all countries[4] while the survival has not been improved as greatly as for many other cancers. On 1 hand, only one-third of patients diagnosed with HCC are candidates for curative treatments such as surgical resection, radiofrequency ablation, and liver transplantation.[5] To make it worse, the 5-year recurrence or metastasis rate can reach 70% after surgical resection.[6] On the other hand, the majority of patients with HCC are poor candidates for curative treatments and the only available option is systemic therapy. Until recently, the only available systemic treatment was sorafenib. It often causes adverse events, and prolongs overall survival (OS) by only 3 months in patients with advanced HCC.[4]

HCC is considered as an immunogenic tumor that arises in chronically inflamed livers due to underlying chronic liver disease caused by viral and nonviral pathogenesis. HCV and HBV can also drive an immune-mediated inflammatory response which promotes neoplastic change, and the latter can also mediate its carcinogenic properties via direct oncogenic transformation following incorporation into host cell DNA.[7] In general, there are some differences in the treatment of viral HCC and nonviral

Editor: Patricia Severino.

HQ and SS contributed equally to this work.

The authors have no funding and conflicts of interest to disclose.

Supplemental Digital Content is available for this article.

The datasets generated during and/or analyzed during the current study are publicly available.

[a] The First Hospital of Lanzhou University, Lanzhou, [b] Xifeng District People's Hospital of Qingyang City, [c] Qingyang Huanxian People's Hospital, Qingyang, [d] The First Clinical Medical College of Lanzhou University, [e] Evidence Based Medicine Center, School of Basic Medical Science, Lanzhou University, Lanzhou, China.

[*] Correspondence: Lei Qi, Xifeng District People's Hospital of Qingyang City, Qingyang 745000, China (e-mail: 497441480@qq.com).

How to cite this article: Qian H, Song S, Tian Y, Qi L, Gao H, Yuan L. Study on the gene signature related to immune microenvironment on viral and nonviral infections of hepatocellular carcinoma. Medicine 2021;100:15(e25374).

Received: 2 February 2020 / Received in final form: 19 February 2021 / Accepted: 10 March 2021

http://dx.doi.org/10.1097/MD.0000000000025374

HCC. Moreover, the prognosis of the 2 types of HCC is different, and the difference of prognosis may be related to the immune microenvironment of HCC. However, there are few reports on this aspect. This work mainly discusses the differences of immune microenvironment between 2 types of HCC and the gene signature related to immune microenvironment. Therefore, we constructed prognostic index (PI) from gene expression and coefficient of Cox regression. Then, the correlation between PI and immune microenvironment of 2 types of HCC was analyzed. In viral infection (VIR) group, the prognostic markers were mainly associated with neutrophil, macrophage, and dendritic ($P < .05$). In the NVIR group, the prognostic gene signature were not significantly correlated with the 6 immune microenvironments ($P > .05$). Furthermore, once liver has developed as tumors, it can be associated with a rich immune cell infiltrate. Some analysis indicated that approximately 25% of HCCs have high inflammatory scores, with high or moderate levels of lymphocyte infiltration.[8] The tumor-infiltrating immune cells (TIICs) form a large component in solid tumors, in an attempt by the host to mediate an antitumor reaction. This promotes immune tolerance and has been shown to confer a worse prognosis.[9] Generally, HCC includes VIRs and nonviral environmental. These 2 types of HCC performed different gene expression and pathways.[9] Thus, HCC patients with viral and nonviral also performed different prognosis and biological process (BP). In this study, we proposed Cox regression and deconvolution method for verification gene signature and TIICs in 2 types of HCC. Therefore, identification special gene signature and TIICs in 2 types of HCC could provide meaningful evidence for improving clinical manager.

Therefore, prognostic biomarkers related to TIICs not only help clinicians to classify patients into different risk levels, but also identify patients who benefit most from the new immunotherapy. The results showed that 11 prognostic genes (*CXCL5*, *NR0B1*, *CHGA*, *STC2*, *MC1R*, *SEMA3G*, *FABP6*, *TDGF1*, *CMTM4*, *LTBP2*, and *R3HDML*) in VIR group and 12 prognostic genes (*CCR10*, *FABP7*, *TNFSF4*, *CSPG5*, *HTR3B*, *SPP1*, *PTHLH*, *MC1R*, *PDCD1*, *PCSK1*, *PROK1*, and *APOD*) in NVIR group. Of these genes, MC1R performed different biological function in 2 types HCC. Then, combination of gene expression and coefficient of Cox regression as PI to classify patients into high-risk and low-risk groups. Moreover, considering the relationship between PI and TIICs, the Spearman correlation method was applied to the analysis. This work may provide effective candidate genes for the prediction and treatment of different types of HCC patients.

2. Materials and methods

2.1. Design of workflow

For investigating the gene signature in HCC with VIR and NVIR in molecular level, a hypothesis was proposed, which assumes that gene signature may involve in TIME that against the HCC. In this study, the prognosis of gene signature related with immune was explored in HCC patients with into VIR and NVIR from the Cancer Genome Atlas (TCGA) database. The detail of the design workflow was listed in Figure 1.

2.2. Collection of genomic data and clinical data of HCC

Genomic data and clinical data and sample information for TCGA-liver cancer hepatocellular carcinoma samples were downloaded from TCGA (https://portal.gdc.cancer.gov/), which provides a normalized form of data (level 3). The corresponding genomic data and clinical data of HCC from public datasets were retrieved. In this study, we have employed transcriptome data of HCC. The genomic data was mainly downloaded messenger RNA (mRNA) data. The normalization data represents fragments per kilobase of exon per million reads mapped data of gene expression from TCGA that were log2 transformed. OS information of all TCGA datasets was directly downloaded from the UCSC Xena browser (GDC hub). Subsequently, 3 types of samples were removed:

(1) patients without OS data;
(2) patients with OS time less than 30 days;
(3) patients without mRNA expression data.

List of clinical information was listed in Table 1. This manuscript only deals with genomic data in public databases and does not use real patient tissues. Therefore, we state that it is not necessary for ethical review.

2.3. Differential expression genes (DEGs) between HCC primary tumors (HCC) and HCC tumor adjacent tissues (HTA)

The DEGs between HCC (n = 255) and HTA (n = 50) tissues were calculated. In addition, we divided primary tumors into VIR (n = 97) and NVIR (n = 158) according to clinical information for further study. For assay, the deregulation gene expression, "Limma" package of R software was employed to test the DEGs. The mRNAs with log2 fold change $|\log FC| > 1$ (FDR adjusted P-value $< .01$) were considered to be significantly differentially expressed mRNAs.

2.4. Infiltrating immune cells of HCC analysis in tumor immune estimation resource (TIMER) database

To characterize and quantify each immune cell subtype, TIMER use levels of 6 tumor infiltrating immune subsets are precalculated for 10,897 genes from 32 cancer types (http://cistrome.dfci.harvard.edu/TIMER/).[10] TIMER is a comprehensive resource for systematic analysis of immune infiltrates across diverse cancer types (https://cistrome.shinyapps.io/timer/). TIMER applies a deconvolution previously published statistical method[8] to infer the abundance of TIICs from gene expression profiles. We analyzed DEGs expression in different types of HCC and the correlation with the abundance of immune infiltrates, including B cells, $CD4^+$ T cells, $CD8^+$ T cells, neutrophils, macrophages, and dendritic cells, via gene modules.

2.5. Immune genes analysis from public database

The genes related to immune were downloaded from the public database IMMPORT (https://www.immport.org/home). This database contains 456 cytokines genes related to immune which are reported in the literature. The database not only includes transcription factors in different types of cancer in TCGA database, but also includes all the known markers of immune cells. These genes were matched to prognostic genes.

2.6. Univariable cox proportional hazards regression

For analysis of the prognosis of immune genes in HCC, follow-up data and gene expression were applied to calculate coefficient of each gene. Package survival of R was employed to conduct the

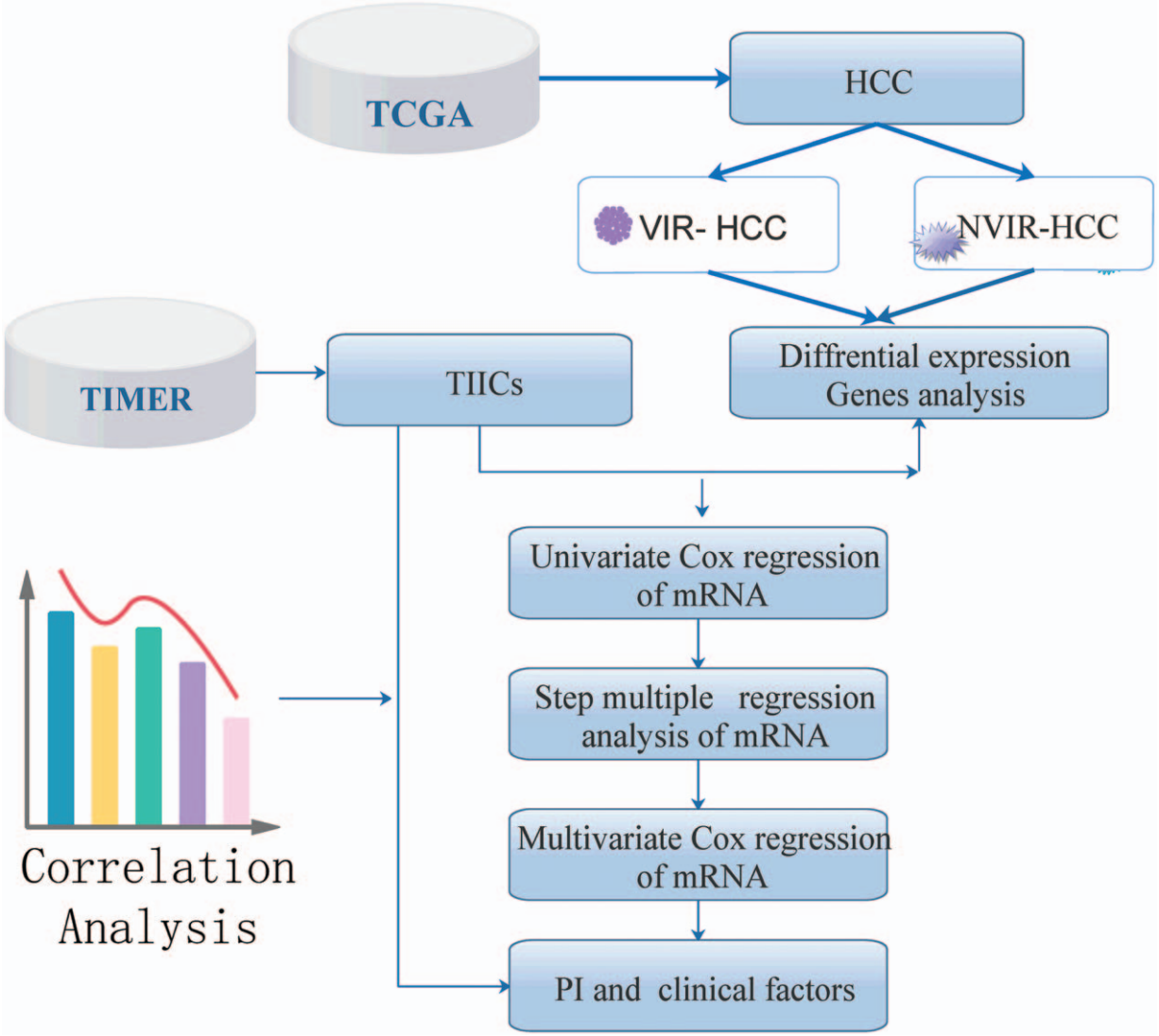

Figure 1. Flow diagram for data analysis.

univariable Cox proportional hazards regression. This was applied to every mRNA selected from the former step to analyze the independent effect. The hazard ratio (HR) of each RNA was calculated using the following equation:

$$HR = e^{\beta} \quad (1)$$

where β represented the coefficient from Cox regression. The 2-sided Wald test was used to test the statistical significance of every predictor. Genes with P-value $< .05$ were selected as valuable predictors of HCC prognosis.

Table 1

List of clinical data of selected samples.

Types of HCC	No. of cases	Male/ female	Median overall survival	Death/ rate (%)
HCC with VIR	97	77/20	693	24/24.7%
HCC with nVIR	158	89/69	601	80/50.6%

HCC = hepatocellular carcinoma, nVIR = nonviral infection, VIR = viral infection.

PI analysis in HCC between VIR and NVIR

Based on the included genes, a risk score named PI was calculated for every patient as an integrated inductor for survival analysis. PI was calculated as follows.

$$PI = \sum_{i=1}^{m} \beta_i \times E_i \quad (2)$$

where β_i represented the coefficient of the involved gene i, and E_i represented the expression level of the corresponding gene. To obtain the best PI cutoff value to divide HCC patients into a high-risk or low-risk survival group, the time-dependent receiver operating characteristic (ROC) curve analysis and Youden index (J) were applied. Next, Kaplan–Meier (K-M) survival curves were created to evaluate the predictive value of PI.

2.7. Analysis of clinical factors and their joint effect with PI

In previous studies, several clinical characteristics have been reported to be related to the OS of HCC patients. Therefore, we explored the role of several clinical factors of HCC patients in

TCGA via univariable Cox proportional hazards regression. Next, multiple Cox regression analyses were conducted to explore the joint effect of these clinical factors, and the previously calculated PI. This allowed for comparison of the prognostic value of the PI to that of each clinical factor. In this study, 7 clinical factors which included age, gender, race, tumor lymph node and metastasis (TNM) stage, T stage, N stage, and M stage involved in analysis. Among the above clinical variables, only age was divided into 2 groups according to the median. Other variables were analyzed according to the actual situation.

2.8. Correlation analysis between TIICs (VIR and NVIR) and genes signature in HCC

The gene signature was analyzed in TIICs (VIR and NVIR). The "Spearman" method was employed to test correlation between TIME score and PI. The P-value $< .05$ and $|\text{coefficient}| > 0.25$ was considered as significant correlation.

2.9. Gene ontology (GO) and Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathways enrichment for targets

To identify the function of gene signature in VIR and NVIR, GO and KEGG analysis was performed by R software. For analysis gene signature function, which included molecular function (MF), BP, cellular component. GO and KEGG pathway enrichment analysis was carried out by Cluster Profiler[11] package to expound promising signaling pathways correlated with the immune. The significant enrichment was considered as P-value $< .05$.

2.10. Statistical analysis

Univariable and multiple Cox regression were employed to R software. Survival curves were generated by the K-M plots. The results of Kaplan- Meier plots were displayed with HR and P-values a log-rank test. P-value < 0.05 was considered to be significant. Hazard ratio (HR) and 95% confidence interval (CI) were calculated to identify low risk ($HR < 1$) or high-risk factors ($HR > 1$). And the median value of PI was considered as cut-off value for high-risk and low-risk group. In this study, The ROC curve analysis was conducted by the survival ROC package, and the area under the curve (AUC) was calculated to measure the predictive accuracy of this prognostic signature for time-dependent cancer death.

3. Results

Gene signatures of HCC with VIR and NVIR were matched with known immune genes from IMMport database (https://www.immport.org/). The matched genes constructed PI model for predicting prognosis of 2 types of HCC. And the PI model that consisted of gene expression and Cox coefficient was tested correlation between PI and tumor immune microenvironment. HCC with VIR and NVIR showed different correlation in tumor immune microenvironment.

3.1. DEGs analysis in tumor and normal tissue

Our study sample comprised 255 patients. Patient characteristics are shown in Table 2. For identification of mRNAs that were significantly associated with disease-free survival and OS, the clinical factors of 255 patients were analyzed.

In this study, we initially collected expression levels of 19,115 mRNAs of 255 HCC patients from TCGA. The results showed that the prognosis of patients in VIR and NVIR groups of HCC subgroups was very different (Fig. 2A). The OS of patients in VIR group was significantly better than that in NVIR group. Then, for mRNAs, differential expression analysis was performed and 3131 predictors were involved in the model (Supplementary Table 1, http://links.lww.com/MD2/A14). Of these DEGs, all were upregulation. The volcano plot was showed in Figure 2B. This operation filtered out the statistically significant signatures, which reduced the complexity of the model.

3.2. Immune related gene signature and prognostic analysis

Here, we applied the original TIMER gene signature file which defines 6 immune cell subtypes and analyzed datasets from VIR and NVIR. Publicly available gene expression profiles from HCC primary tumors (HCC, $n = 255$). All tumor samples were analyzed for immune cell profiles by TIMER to estimate the abundance of 6 (TIICs) subsets (B cells, CD4 T cells, CD8 T cells, macrophages, neutrophils, and dendritic cells). From all the samples analyzed, we have selected 97/158 VIR/NVIR tissues, respectively. Immune cell profile was calculated for each sample and mean values for each tissue type (VIR and NVIR) were calculated. One-way analysis of variance was applied to analyze the differences between VIR and NVIR tissues.

The extent of immune cell infiltration into tumors has important prognostic value in HCC and other cancers.[12,13] Tumor-infiltrating lymphocytes are an independent predictor of sentinel lymph node status and survival in cancers.[14] Therefore, through overlapping the DEGs and immune infiltration levels, 212 DEGs were selected to further study (Supplementary Table 2, http://links.lww.com/MD2/A15). For testing prognosis of each gene in TIICs, univariate Cox regression was employed to analyze TIICs in HCC with VIR and NVIR (Supplementary Table 3, http://links.lww.com/MD2/A16 and Table 4, http://links.lww.com/MD2/A17). After univariable Cox proportional hazard regression, only 20 mRNAs were included in the model in VIR (Table 3) and 34 mRNAs were included in the model in NVIR (Table 4).

For testing prognosis of each gene in immune cell infiltration, multiple stepwise regression was employed to analyze TIICs in VIR and NVIR groups. In VIR group, a total of 11 genes were involved in model for which the HRs and 95% CIs are listed in Figure 3A. Among them, 6 genes (*CXCL5*, *NR0B1*, *CHGA*, *STC2*, *MC1R*, and *SEMA3G*) were considered as high-risk factors because the HRs were larger than 1. On the contrary, 5 genes (*FABP6*, *TDGF1*, *CMTM4*, *LTBP2*, and *R3HDML*) were protective or low-risk factors. A description of 12 genes of NVIR is provided in Figure 3B. Among them, 7 genes (*CCR10*, *FABP7*, *TNFSF4*, *CSPG5*, *HTR3B*, *SPP1*, and *PTHLH*) were considered as high-risk factors because the HRs were larger than 1. The other 5 genes (*MC1R*, *PDCD1*, *PCSK1*, *PROK1*, and *APOD*) were protective or low-risk factors.

In VIR group, for every HCC patient, K-M curves (Fig. 4A and B) show that the P-value of the log-rank test was 2.01×10^{-10}, which implied that the PI could significantly separate patients in high-risk and low-risk groups. The time-dependent ROC curve was created, as shown in Figure 4C and D, and AUC of the ROC curve in VIR group showed that 3 years AUC = 0.937, 1 years AUC = 0.949. In NVIR group, AUC of the ROC curve in VIR

Table 2

Univariate analyses of variables influencing survival of 255 patients with HCC from TCGA database.

Factors	VIR (n = 97) N (%)	HR (95% CI)	*P*-value Univariate analysis	NVIR (n = 158) N (%)	HR (95%CI)	*P*-value Univariate analysis
Age		1.343 (0.853–2.114)	.202		0.696 (0.536–0.904)	.601
<=55	50			<=64 73		
>55	47			>64 85		
Gender		1.062 (0.395–2.851)	.905		0.887 (0.566–1.390)	.601
Male	77			89		
Female	20			69		
Race		0.744 (0.330–1.675)	.475		0.925 (0.591–1.446)	.732
American Indian or Alaska native	1					
Asian	69			45		
Black or African American	8			4		
White	19			102		
Nonclassfication				7		
Stage		1.504 (1.113–2.034)	.007		1.337 (1.133–1.577)	.0005*
Stage I	60			54		
Stage II	21			31		
Stage III	3			58		
Stage IV	1			2		
Nonreport	1			13		
T		1.747 (1.206–2.532)	.003*		1.548 (1.240–1.932)	<.0001*
T1	60			58		
T2	23			36		
T3	10			55		
T4	2			8		
Nonreport	1			1		
N	1.126 (0.563–2.252)	0.736			1.149 (0.909–1.451)	.245
N0	78			101		
N1	17			3		
N2						
NX	2			54		
M	1.545 (0.953–2.503)	0.077			1.144 (0.901–1.453)	.269
M0	79			57		
M1	1			2		
MX	17			36		

CI = confidence interval, HR = hazard ratio, NVIR = nonviral infection, VIR = viral infection.
* Median survival age.

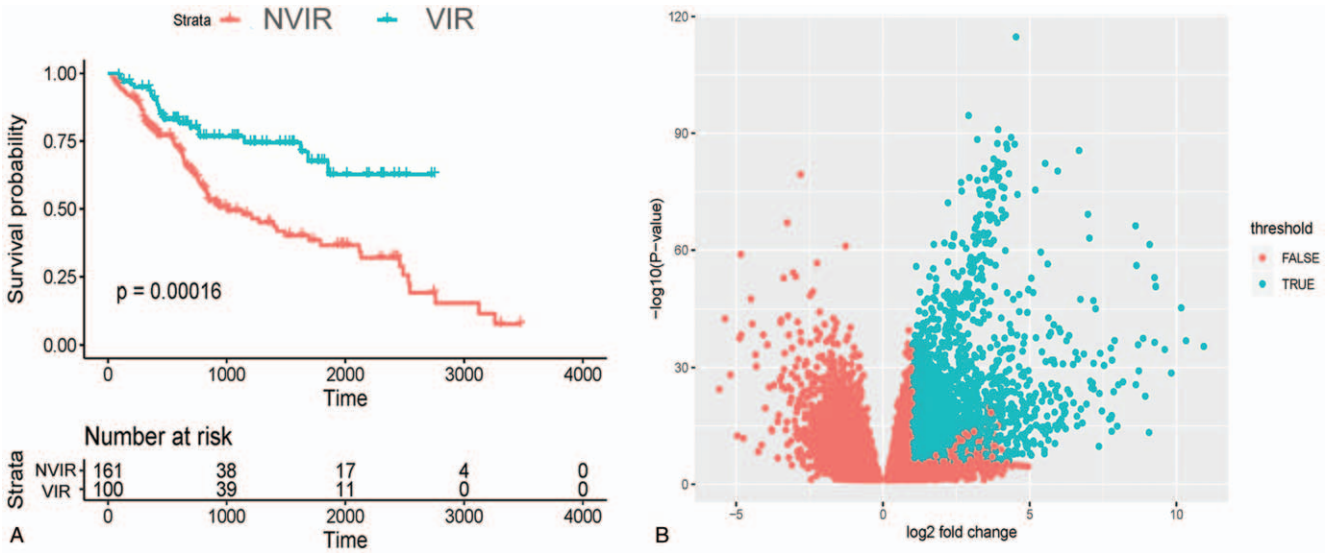

Figure 2. (A) K-M curves showing overall survivals of high-risk and low-risk patients. (B) Volcano plot represented DEGs. DEGs = differential expression genes, K-M = Kaplan–Meier.

Table 3

The DEGs with immune cell infiltration in VIR of HCC.

Genes	HR	Low	Up	Coef	*P*-value
CXCL5	1.129	1.013	1.258	0.121	.028
S100A2	1.279	1.018	1.607	0.246	.034
COLEC12	1.353	1.072	1.709	0.303	.011
FABP6	1.094	1.007	1.189	0.090	.033
FABP7	1.131	1.005	1.272	0.123	.041
R3HDML	0.891	0.809	0.982	−0.115	.020
LYZ	0.563	0.318	0.998	−0.574	.050
PDGFRB	0.459	0.234	0.901	−0.777	.024
ACTA1	0.747	0.592	0.944	−0.291	.014
SEMA3G	0.500	0.314	0.797	−0.693	.004
CHGA	1.300	1.125	1.503	0.263	.000
CMTM4	0.573	0.345	0.952	−0.557	.031
EGF	1.137	1.038	1.246	0.129	.006
LTBP2	0.580	0.405	0.831	−0.544	.003
RLN3	1.107	1.008	1.216	0.102	.034
STC2	1.870	1.220	2.866	0.626	.004
TDGF1	0.715	0.588	0.868	−0.336	.001
MC1R	2.629	1.402	4.985	0.966	.043
NR0B1	1.148	1.052	1.252	0.138	.002
TNFRSF25	0.521	0.299	0.909	−0.652	.022

DEGs = differential expression genes, HCC = hepatocellular carcinoma, VIR = viral infection.

Table 4

The DEGs with immune cell infiltration in NVIR of HCC.

Genes	HR	Low	Up	Coef	*P*-value
CCR10	1.238	1.022	1.500	0.214	.029
DEFB132	0.942	0.899	0.987	−0.059	.012
FABP6	1.075	1.021	1.131	0.072	.006
FABP7	1.076	1.014	1.143	0.074	.016
FABP4	0.781	0.693	0.880	−0.245	<.001
RBP7	0.707	0.557	0.898	−0.347	.005
APOD	0.744	0.622	0.889	−0.296	.001
IKBKE	1.348	1.032	1.762	0.299	.029
NOS2	0.813	0.692	0.954	−0.207	.011
TPM2	0.597	0.360	0.990	−0.516	.045
SPINK5	0.890	0.814	0.973	−0.116	.011
ABCC4	1.184	1.000	1.403	0.169	.050
PDGFRB	0.610	0.421	0.883	−0.495	.009
PDCD1	0.901	0.821	0.990	−0.104	.030
PCSK1	0.903	0.822	0.991	−0.103	.031
BIRC5	1.578	1.207	2.063	0.456	.001
TNFSF4	1.370	1.059	1.772	0.315	.016
CKLF	1.869	1.004	3.479	0.626	.048
ADM2	1.361	1.085	1.707	0.308	.008
CSPG5	1.215	1.022	1.445	0.195	.027
LTBP2	0.807	0.658	0.989	−0.215	.039
PMCH	1.059	1.003	1.117	0.057	.038
PROK1	0.914	0.864	0.966	−0.090	.002
PTHLH	1.133	1.016	1.264	0.125	.024
SPP1	1.392	1.107	1.751	0.331	.005
AVPR2	0.873	0.800	0.952	−0.136	.002
GHRHR	1.065	1.006	1.127	0.063	.030
HTR3B	1.071	1.011	1.135	0.069	.019
IL3RA	0.523	0.341	0.801	−0.649	.003
MC1R	0.703	0.555	0.890	−0.353	.003
MTNR1B	1.084	1.029	1.141	0.080	.002
NR0B1	1.055	1.008	1.104	0.053	.023
TACR1	0.921	0.855	0.991	−0.083	.028
TNFRSF19	0.882	0.784	0.992	−0.126	.037

DEGs = differential expression genes, HCC = hepatocellular carcinoma, NVIR = nonviral infection.

group showed that 3 years AUC = 0.878, 1 years AUC = 0.875. A PI as a risk score was calculated and distribution of PI showed in Figure 4E and F. The heatmap was created based on high-risk and low-risk HCC patients in Figure 4G and H. Therefore, it can be concluded that the PI calculated via DEGs had a good predictive value of prognosis for VIR group and NVIR group of HCC patients in the TCGA database.

3.3. Analysis of clinical factors and PI

In this study, 3 common clinical factors recorded in the TCGA were explored, including age, gender, TNM stage. We next introduced the PI and conducted multiple Cox regression analyses. The data showed that the PI and TNM stage continued to be influencing factors of the prognosis in VIR and NVIR groups of HCC patients from TCGA, with $P < .001$. In VIR group, the PI was the most significant with the smallest $P < .001$ and the largest HR of 1.088 (95% CI = 2.068–4.259), which

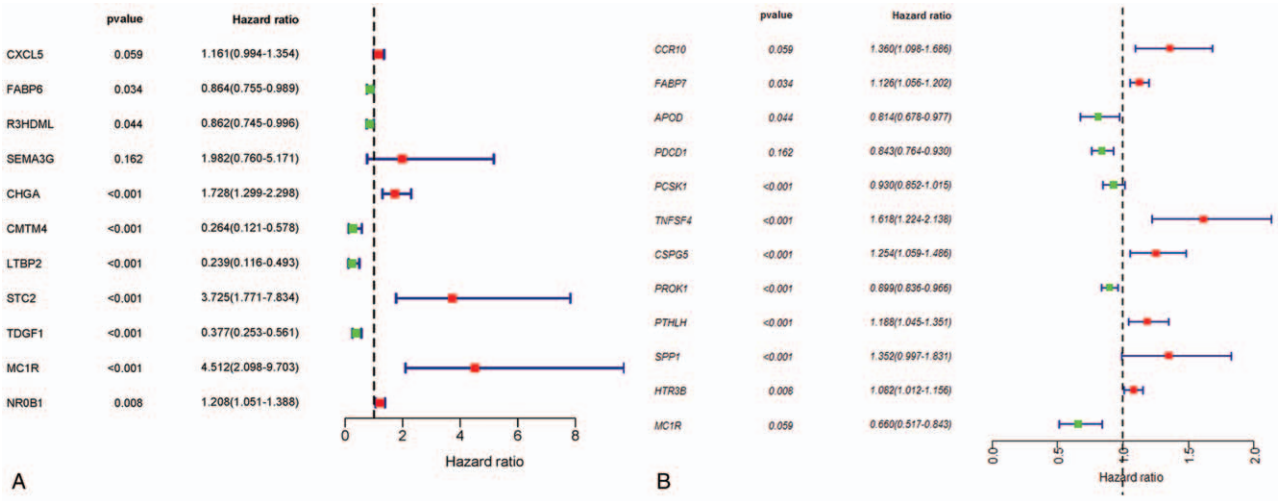

Figure 3. The Forest plot of each prognostic gene biomarkers. (A) Forest plot of HRs by applying gene biomarkers to classify VIR group patients. (B) Forest plot of HRs by applying gene biomarkers to classify NVIR group patients. HRs = hazard ratios, NVIR = nonviral infection, VIR = viral infection.

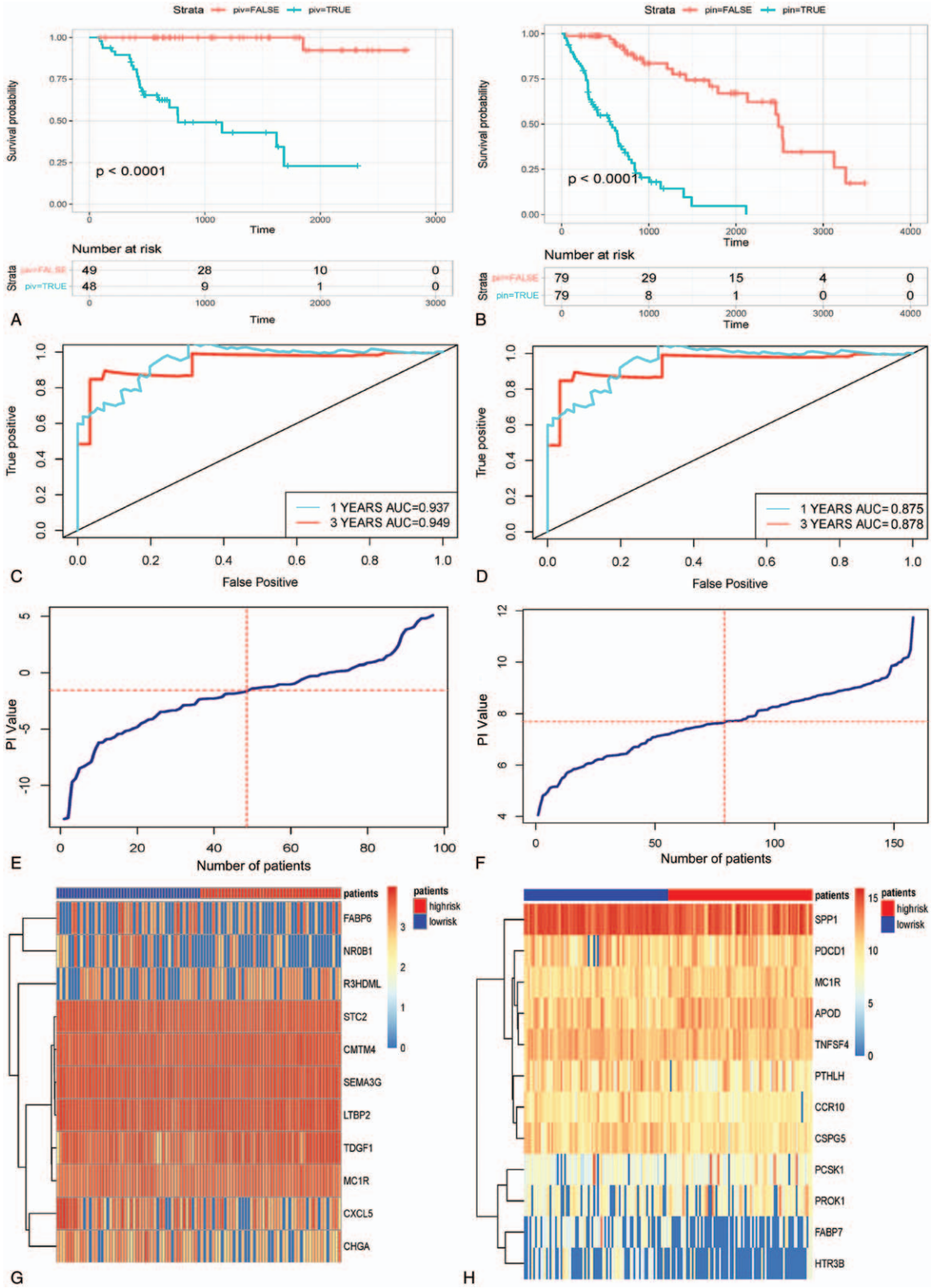

Figure 4. The gene signature in VIR and NVIR groups. (A) K-M curves showing overall survivals of high-risk and low-risk in VIR group patients. (B) K-M curves showing overall survivals of high-risk and low-risk in NVIR group patients. (C) The ROC curves in VIR group patients. (D) The ROC curves in VIR group patients in NVIR group patients. (E) Time-dependent ROC curve based on PI in VIR groups patients. (F) Time-dependent ROC curve based on PI in VIR groups patients in NVIR group patients. (G) The heatmap of the 9 mRNAs of high-risk and low-risk patients in VIR groups patients. (H) The heatmap of the 12 mRNAs of high-risk and low-risk patients in NVIR groups patients. K-M = Kaplan–Meier, mRNAs = messenger RNAs, NVIR = nonviral infection, PI = prognosis index, ROC = receiver operating characteristic, VIR = viral infection.

Table 5

Multivariable Cox regression of clinical factors and prognosis index.

Factor	HR	(95% CI)		*P*-value
VIR				
PI	1.088	2.068	4.259	<.001
Age >38	1.528	0.463	45.848	.192
Sex	0.171	0.390	3.612	.764
T stage	0.483	0.833	3.157	.155
N stage	0.450	0.406	6.062	.515
M stage	−0.286	0.198	2.854	.675
NVIR				
PI	0.964	1.947	3.533	<.001
Age >64	−0.856	0.237	0.763	.004
Sex	0.077	0.678	1.722	.746
T stage	0.875	1.463	3.931	.001
N stage	0.255	0.755	2.207	.352
M stage	0.275	0.747	2.319	.341

CI = confidence interval, HR = hazard ratio, NVIR = nonviral infection, PI = prognostic index, VIR = viral infection

implied that the PI was superior to conventional prognostic markers, including TNM stage or pathologic stage with respect to predicting HCC patient outcome (Table 5).

We investigated whether PI value of DEGs was correlated with immune infiltration levels in different types of HCC. We assessed the correlations between PI value of DEGs and immune marker genes of different TIICs, included barcode B cells, CD4+T cells, CD8+ T cells, neutrophils, macrophages, and dendritic in VIR and NVIR groups of HCC. The results showed that PI value of DEGs has significant correlations with neutrophils, macrophages, and dendritic infiltration levels in VIR group (Fig. 5). However, in NVIR group, macrophages expression infiltration levels invalid so exclude, the result showed there were no significant correlations between PI and other 5 type cell infiltration levels. And it showed no significant in NVIR group (*P*-value > .05) (Fig. 6).

3.4. GO and pathway enrichment

After filtering gene signature and the results of GO enrichment showed that the most significant BP was response to endogenous

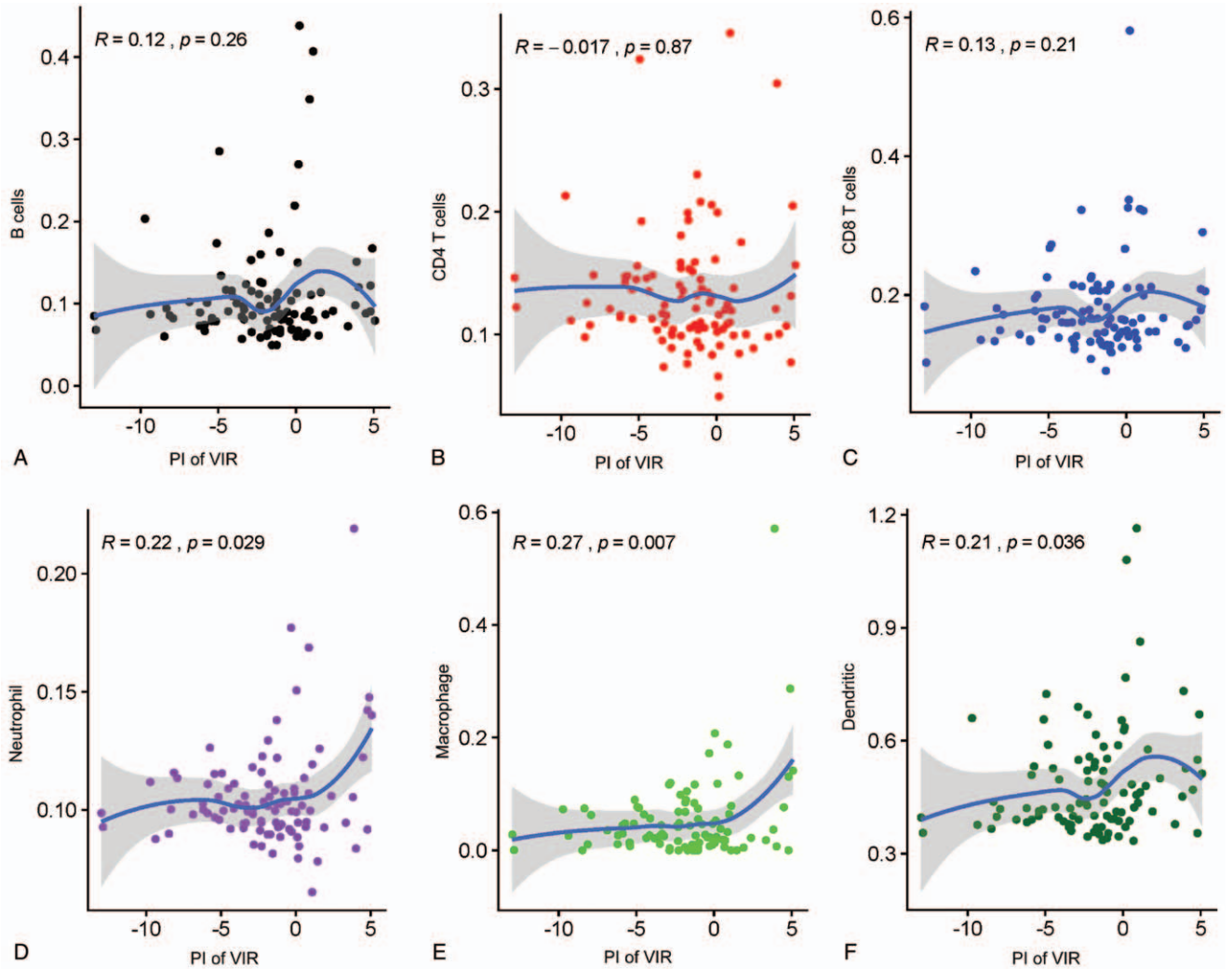

Figure 5. Correlation analysis between PI values and immune marker sets in VIR. (A) Correlation between PI of VIR and B cells (*P*-value = .26). (B) Correlation between PI of VIR and CD4 T cells (*P*-value = .87). (C) Correlation between PI of VIR and CD8 T cells (*P*-value = .21). (D) Correlation between PI of VIR and neutrophil (*P*-value < .05). (E) Correlation between PI of VIR and macrophage (*P*-value < .05). (F) Correlation between PI of VIR and dendritic (*P*-value < .05). PI = prognosis index, VIR = viral infection.

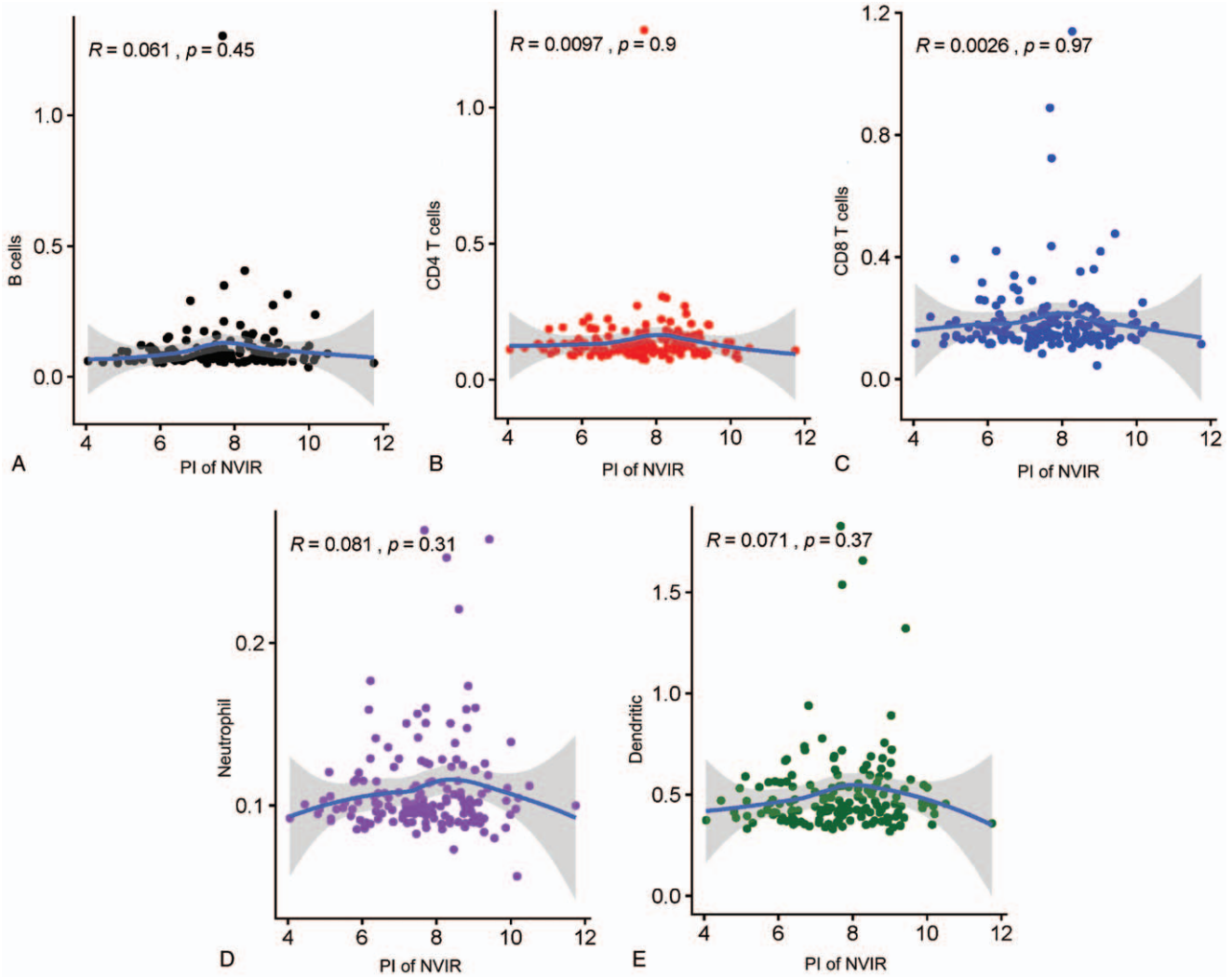

Figure 6. Correlation analysis between PI values and immune marker sets in NVIR. (A) Correlation between PI of NVIR and B cells (*P*-value = .45). (B) Correlation between PI of NVIR and CD4 T cells (*P*-value = .90). (C) Correlation between PI of NVIR and CD8 T cells (*P*-value = .97). (D) Correlation between PI of NVIR and neutrophil (*P*-value = .31). (E) Correlation between PI of NVIR and macrophage (*P*-value = .37). NVIR = nonviral infection, PI = prognosis index.

stimulus and cellular response to chemical stimulus in VIR groups (Fig. 7A). In MF, receptor binding was only 1 enrichment function (Fig. 7B). In cellular component of VIR group, gene signature mainly involved in extracellular region (Fig. 7C). In NVIR group, gene signature involved in many BPs. The results showed that cell proliferation and negative regulation of multicellular organismal process were mainly biological enrichment (Fig. 8A). In MF enrichment of NVIR group, gene signature mainly involved in receptor binding function (Fig. 8B). The gene signature of NVIR group mainly located cell surface, plasma membrane (Fig. 8C). The results of molecular functional analysis showed that they were both related to molecular functional in VIR group and VIR group.

4. Discussion

As the liver has dual vasculature, the nutrients and pathogen-derived molecules (i.e., lipopolysaccharides) from the portal vein and oxygenated blood from the systemic circulation via the hepatic artery are received by liver. This unique macroscopic anatomy leads to enormous exposure to gut-borne pathogens as well as exogenous nonpathogenic molecules. As such, the immunologic composition of the liver contains the largest concentration of immune effectors in the body.[15] HCC is a challenging environment for the immune system. Our study was aimed to analyze the DEGs on differential types of HCC, with special focus on their relationship with TIICs.

In this study, we found the different prognostic gene signature in VIR group and NVIR group, except *MC1R*. This gene is a G-protein coupled receptor with high affinity for alpha-melanocyte stimulating hormone which has been reported only in melanoma and not in other cancers.[16]*MC1R* gene showed different function which performed different BP. Not only that, the correlation between the different gene and TIICs was very different. In VIR group, there were 11-gene signatures, and the related TIICs include neutrophils, macrophages, and dendritic. In NVIR group, 12-gene signature was no correlation with TIICs. Moreover, the role of tumor-associated neutrophils (TANs) in HCC progression and their effect on the microenvironment remain undefined.

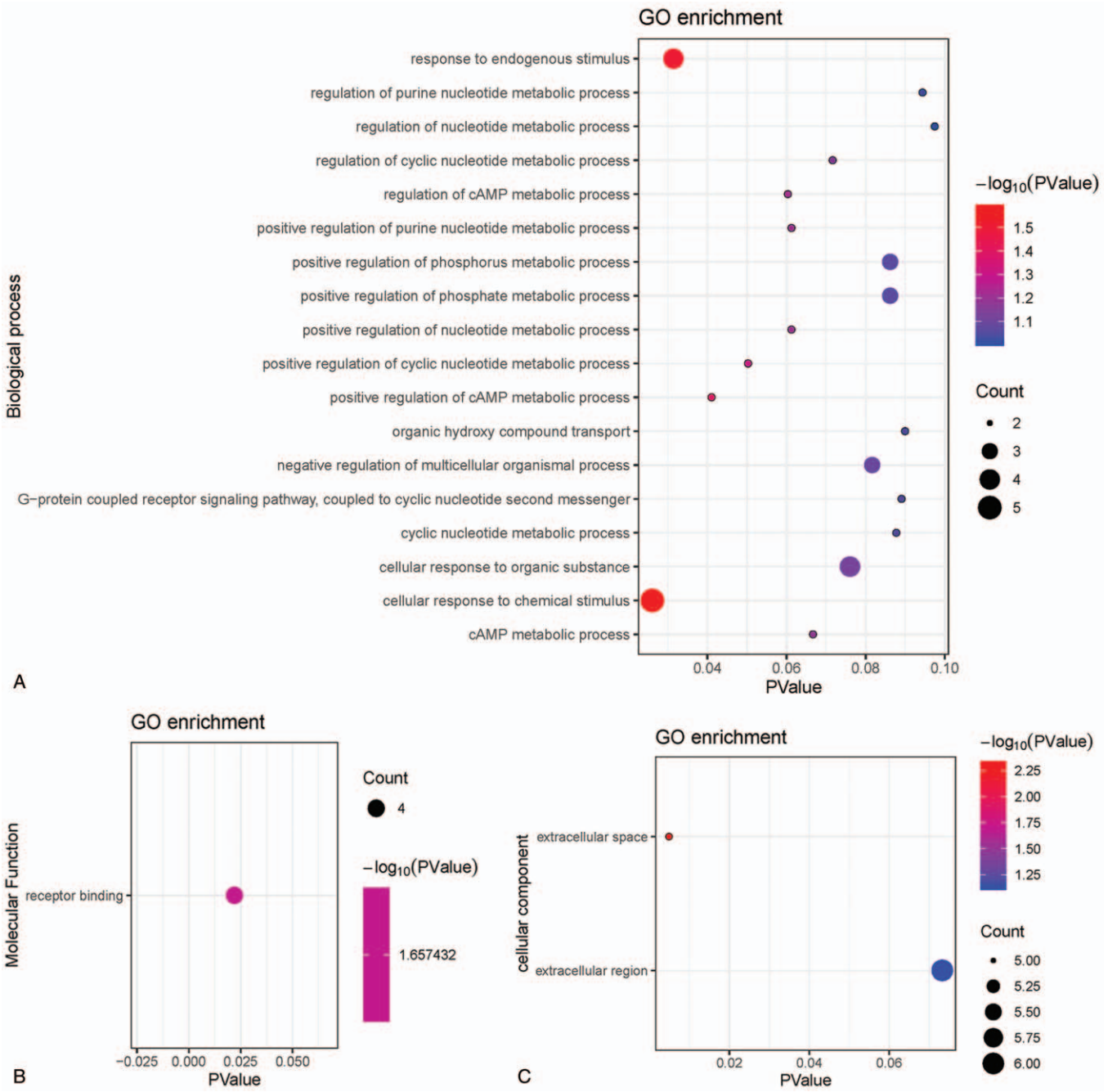

Figure 7. GO enrichment of gene signature in VIR. (A) Biological process of gene signature in VIR. (B) Molecular function of gene signature in VIR. (C) Molecular function of gene signature in VIR. VIR = viral infection.

Recently, there has been emerging interest to study the role of neutrophils in cancer. In the tumor microenvironment, TANs have been proposed to support tumor development by promoting cellular transformation, tumor progression, and antitumor immunity.[17] TAN infiltration is prognostic in several human cancers,[18] including HCC. TANs recruit macrophages and Treg cells to HCCs to promote their growth, progression, and resistance to sorafenib, and identified C-C motif chemokine 2 and C-C motif chemokine 17, which are secreted exclusively by TANs, as tumor promoting and prognostic factors in HCC.[19] Previous studies have shown that cytotoxic CD8+T cells, derived from livers that are chronically infected with HBV/HCV, have an exhausted phenotype and are less adept at delivering a cytotoxic immune response capable of controlling infection.[20–22] In HBV murine models, treatment with an anti-Programmed cell death protein 1 or an anti-Programmed cell death 1 ligand 1 antibody leads to resurgence of T-cell-mediated immunity and clearance of HBV in vivo.[2,4] Macrophages and Treg cells both have the ability to promote tumor progression.[23] The study showed a crucial role for TANs in recruiting macrophages and Treg cells both in vitro and in vivo, depletion of which significantly attenuated the protumor effect of TANs. Thus, TAN-induced tumor promotion is dependent on the recruitment of macrophages and Treg cells. Moreover, a further reduction in tumor

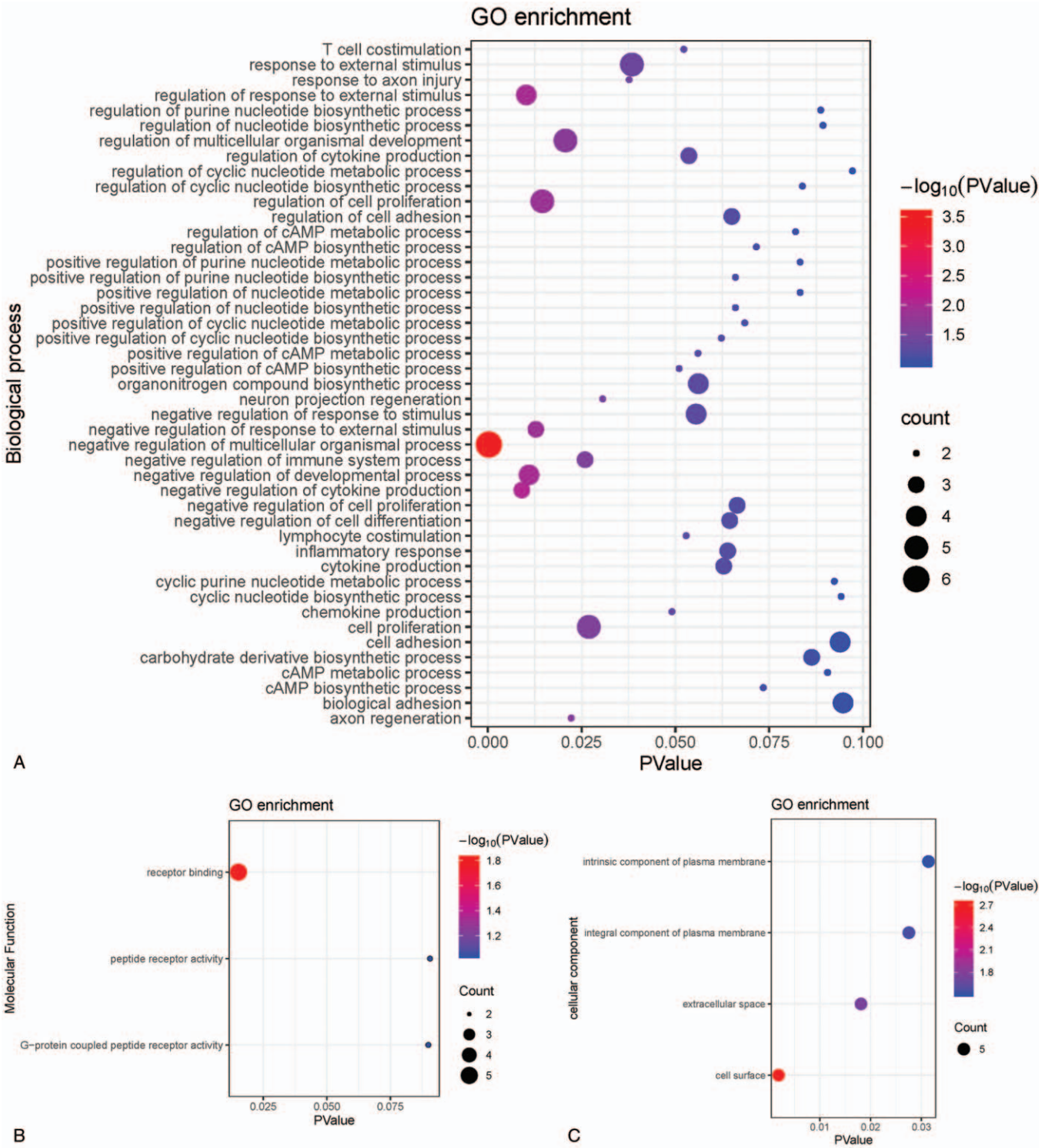

Figure 8. GO enrichment of gene signature in NVIR. (A) Biological process of gene signature in NVIR. (B) Molecular function of gene signature in NVIR. (C) Molecular function of gene signature in NVIR. NVIR = nonviral infection.

volume after depletion of TANs, suggesting that TANs also promote HCC progression through direct interaction with HCC cells.[19] HCC with a moderate mutational burden (mean somatic mutation rate of approximately 1.3 mutations per mega base) and rare examples of hypermutation may, therefore, be less sensitive to checkpoint blockade than other tumor types.[24] This may explain why the survival rate of patients in the VIR group was significantly better than that in the NVIR group.

Therefore, our study not only confirmed that gene signature in VIR group and NVIR group can be used as independent biomarkers for prognosis of liver cancer, but also provided ideas for precise treatment decision-making of HCC patients.

Acknowledgments

The authors thank the Chinese Academy of Sciences for its computing equipment. All the authors thank Qiong Yang and Jizheng Duan, the administrators of supercomputing, for their help.

Author contributions

Conceptualization: Hong Qian, Shanglin Song.
Formal analysis: Yunling Tian, Haihong Gao, Lingyan Yuan.
Investigation: Hong Qian, Shanglin Song, Yunling Tian, Lei Qi.
Methodology: Hong Qian, Lei Qi.
Project administration: Hong Qian.
Software: Haihong Gao.
Supervision: Yunling Tian, Lei Qi.
Validation: Lingyan Yuan.
Visualization: Lingyan Yuan.
Writing – original draft: Hong Qian.
Writing – review & editing: Hong Qian, Yunling Tian.

References

[1] Park C-K, Oh I-J, Kim Y-C. What is currently the best for adenocarcinoma without driver mutation? Tuberc Respir Dis 2018;81:258–9.
[2] Harding JJ, El Dika I, Abou-Alfa GK. Immunotherapy in hepatocellular carcinoma: primed to make a difference? Cancer 2016;122:367–77.
[3] Hagag NA, Ali YBM, Elsharawy AA, et al. Clinical impact of circulated miR-1291 in plasma of patients with liver cirrhosis (LC) and hepatocellular carcinoma (HCC): implication on glypican-3 expression. J Gastrointest Cancer 2019;51:234–41.
[4] Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2015;136:E359–86.
[5] Galle PR, Forner A, Llovet JM, et al. EASL clinical practice guidelines: management of hepatocellular carcinoma. J Hepatol 2018;69:182–236.
[6] Gao Q, Wang XY, Qiu SJ, et al. Overexpression of PD-L1 significantly associates with tumor aggressiveness and postoperative recurrence in human hepatocellular carcinoma. Clin Cancer Res 2009;15:971–9.
[7] Hernandez-Gea V, Toffanin S, Friedman SL, et al. Role of the microenvironment in the pathogenesis and treatment of hepatocellular carcinoma. Gastroenterology 2013;144:512–27.
[8] Li B, Severson E, Pignon JC, et al. Comprehensive analyses of tumor immunity: implications for cancer immunotherapy. Genome Biol 2016;17:174.
[9] Chang CS, Huang SM, Lin HH, et al. Different expression of apoptotic proteins between HBV-infected and non-HBV-infected hepatocellular carcinoma. Hepatogastroenterology 2007;54:2061–8.
[10] Li T, Fan J, Wang B, et al. TIMER: a web server for comprehensive analysis of tumor-infiltrating immune cells. Cancer Res 2017;77:e108–10.
[11] Aran D, Sirota M, Butte AJ. Systematic pan-cancer analysis of tumour purity. Nat Commun 2015;6:8971.
[12] Hato T, Goyal L, Greten TF, et al. Immune checkpoint blockade in hepatocellular carcinoma: current progress and future directions. Hepatology 2014;60:1776–82.
[13] Mlecnik B, Bindea G, Angell HK, et al. Integrative analyses of colorectal cancer show immunoscore is a stronger predictor of patient survival than microsatellite instability. Immunity 2016;44:698–711.
[14] Ohtani H. Focus on TILs: prognostic significance of tumor infiltrating lymphocytes in human colorectal cancer. Cancer Immun 2007;7:4.
[15] Jenne CN, Kubes P. Immune surveillance by the liver. Nat Immunol 2013;14:996–1006.
[16] Nasti TH, Timares L. MC1R, eumelanin and pheomelanin: their role in determining the susceptibility to skin cancer. Photochem Photobiol 2015;91:188–200.
[17] Galdiero MR, Garlanda C, Jaillon S, et al. Tumor associated macrophages and neutrophils in tumor progression. J Cell Physiol 2013;228:1404–12.
[18] Donskov F. Immunomonitoring and prognostic relevance of neutrophils in clinical trials. Semin Cancer Biol 2013;23:200–7.
[19] Zhou SL, Dai Z, Zhou ZJ, et al. Overexpression of CXCL5 mediates neutrophil infiltration and indicates poor prognosis for hepatocellular carcinoma. Hepatology 2012;56:2242–54.
[20] Owusu Sekyere S, Suneetha PV, Kraft AR, et al. A heterogeneous hierarchy of co-regulatory receptors regulates exhaustion of HCV-specific CD8 T cells in patients with chronic hepatitis C. J Hepatol 2015;62:31–40.
[21] Reignat S, Webster GJ, Brown D, et al. Escaping high viral load exhaustion: CD8 cells with altered tetramer binding in chronic hepatitis B virus infection. J Exp Med 2002;195:1089–101.
[22] Fisicaro P, Valdatta C, Massari M, et al. Antiviral intrahepatic T-cell responses can be restored by blocking programmed death-1 pathway in chronic hepatitis B. Gastroenterology 2010;138:682–93.e1-4.
[23] Czepielewski RS, Porto BN, Rizzo LB, et al. Gastrin-releasing peptide receptor (GRPR) mediates chemotaxis in neutrophils. Proc Natl Acad Sci U S A 2012;109:547–52.
[24] Schulze K, Imbeaud S, Letouze E, et al. Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. Nat Genet 2015;47:505–11.