



Multi-omics Analysis of Microenvironment Characteristics and Immune Escape Mechanisms of Hepatocellular Carcinoma

Wenli Li^{1,2†}, Huimei Wang^{2,3†}, Zhanzhong Ma⁴, Jian Zhang⁴, Wen Ou-yang^{2,5}, Yan Qi⁶ and Jun Liu^{4*}

¹ Department of Reproductive Medicine Center, Yue Bei People's Hospital, Shaoguan, China, ² Morning Star Academic Cooperation, Shanghai, China, ³ State Key Laboratory of Medical Neurobiology, Department of Integrative Medicine and Neurobiology, School of Basic Medical Sciences, Institute of Brain Science, Shanghai Medical College, Fudan University, Shanghai, China, ⁴ Department of Clinical Laboratory, Yue Bei People's Hospital, Shaoguan, China, ⁵ The Second Clinical Medical College, Zhujiang Hospital, Southern Medical University, Guangzhou, China, ⁶ Yunnan Provincial Key Laboratory of Traditional Chinese Medicine Clinical Research, First Affiliated Hospital of Yunnan University of Traditional Chinese Medicine, Kunming, China

OPEN ACCESS

Edited by:

Benjamin Frey,
University of Erlangen
Nuremberg, Germany

Reviewed by:

Zong Sheng Guo,
University of Pittsburgh, United States
Kristian Unger,
Helmholtz Center Munich, Germany

*Correspondence:

Jun Liu
liuyu8566@126.com

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Cancer Immunity and Immunotherapy,
a section of the journal
Frontiers in Oncology

Received: 18 June 2019

Accepted: 23 September 2019

Published: 15 October 2019

Citation:

Li W, Wang H, Ma Z, Zhang J,
Ou-yang W, Qi Y and Liu J (2019)
Multi-omics Analysis of
Microenvironment Characteristics and
Immune Escape Mechanisms of
Hepatocellular Carcinoma.
Front. Oncol. 9:1019.
doi: 10.3389/fonc.2019.01019

The immune environment in primary tumor has a profound impact on immunotherapy. However, the clinical relevance of immune environment in hepatocellular carcinoma (HCC) is largely unknown. Here, the immune profile and its clinical response in HCC were investigated. The gene expression profiles of 569 HCCs from three cohorts (The Cancer Genome Atlas, TCGA, $n = 257$; Gene Expression Omnibus, GEO, $n = 170$; International Cancer Genome Consortium, ICGC, $n = 142$) were used in the current study. Five gene expression subtypes (C1–C5) responsible for global immune genes were identified in HCCs at stage I/II. It was found that subtype C4 was associated with upregulation and subtype C5 was associated with downregulation of immune profiles in most metagenes. Immune-correlation analysis of the five subtypes demonstrated that C3 and C4 had higher immune score and better prognostic outcome, as compared with other subtypes. Moreover, the mutation frequencies of TP53, CTNNB1, and AXIN1 had significant difference in the five subgroups. Further, the expression of PDCD1, CD274, PDCD1LG2, CTLA4, CD86, and CD80 was higher in subtype C4 in comparison with the other subtypes. The WGCNA of immune-related genes in the five subtypes revealed that blue and turquoise modules were positively correlated with subtype C4 and were associated with 12 common pathways in the KEGG database. These results were validated in external cohorts from the NCI (National Cancer Institute) cohort (GSE14520) and the ICGC (International Cancer Genome Consortium) cohort. In summary, one immune-enhanced subtype and one immune-decreased subtype having different immune and clinical characteristics may provide guidance for developing novel treatment strategies for immune system malfunction-related cancer.

Keywords: hepatocellular carcinoma, molecular subtype, bioinformatics analysis, cancer stage, immune escape, immune-enhanced subtype

INTRODUCTION

Hepatocellular carcinoma (HCC) is ranked as the second cause of cancer-related mortality and the fifth most common malignancy in patients worldwide (1). It was estimated in 2012 that around 700,000 people die of HCC every year worldwide (2). The incidence of HCC varies with geographical area, sex, age, and risk factor related to cancer development (3, 4). Despite advances in HCC treatment, such as liver transplantation, surgical resection, and radiofrequency ablation, the prognosis of HCC patients is still low, with a 5-year survival rate of <30% (5). Moreover, the tumor heterogeneity and microenvironment cells also play a critical role in treatment or analysis of tumor stage in HCC (6). Notably, the molecular mechanisms by which the microenvironment cells regulate the development of HCC have not been extensively explored.

HCCs are mostly caused by chronically inflamed liver and are considered as typical immunogenic cancers (7). Immune procedures play important roles in HCC carcinogenesis and progression. Immune suppressor cells, including regulatory T cells (Tregs), tumor-associated macrophages (TAMs), and myeloid-derived suppressive cells (MDSCs) in the HCC tumor microenvironment (8), could result in tumor immune evasion or immune escape by interfering with immune surveillance (9). Moreover, the immune checkpoint, which is known to modulate different stages and signaling procedures of the immune response, is one of the mechanisms of escaping anti-cancer immune surveillance. It has been revealed that immune treatment targeting coinhibitory receptors (i.e., CTLA4 and PD1) increase immune response by inhibiting the immunosuppressive mechanisms in several cancers, such as metastatic melanoma and lung cancer (10, 11), but immunotherapy has not been successfully explored in HCC for decades (12). In addition, the expression profile and clinical relevance of immune checkpoint molecules in HCC have not been studied extensively.

Previous studies explored the effects of tumor microenvironment in HCC. Survival outcome of HCC patients was investigated using immunohistochemistry and quantitative PCR techniques to elucidate the underlying immune gene expression profiles (13). The expression of 49 immune genes was detected in 68 HCC patients. However, the sample size may have reduced the reliability of the conclusions. Moreover, qPCR method detected only two tumor phenotypes (proliferation and apoptosis) and 49 immune genes that were associated with the survival outcome of HCC patients. There is a need to comprehensively detect all the tumor phenotype and global immune profiles using high-throughput techniques. Transcriptome analysis using next-generation sequencing and microarray profiling is a powerful method for systemically exploring the tumor microenvironment. Previous studies revealed the presence of microenvironment cells in tumor tissues via RNA sequencing and microarray analysis of their underlying expression profiles (14, 15). Using microarray analysis and qRT-PCR technique, a unique inflammation/immune response-associated signature of the liver microenvironment was found to be a predictor of venous metastases, recurrence, and prognosis of HCC (16). However, due to the cross-validation in training set and lack of external

validation, a larger HCC cohort is needed to further validate the conclusions.

The current research aimed to comprehensively explore the heterogeneous immune microenvironment phenotypes and their associated clinical relevance in HCC at stage I/II. The Cancer Genome Atlas (TCGA) cohort was used to successfully classify 257 HCC samples into five consensus molecular subtypes of tumors with potential immune escape mechanisms and genomic drivers underlying the gene expression profiles of global immune genes. Both immune-enhanced and immune-decreased subtypes were identified in HCC. Moreover, the five subtypes were validated using an external dataset from the NCI (National Cancer Institute) cohort and the ICGC cohort.

METHODS AND MATERIALS

Genomic Analysis of Immune Genes

Thirteen immune metagenes were obtained from an immune gene set in the TIMER database (17, 18) to reflect the types and functions of various immune genes. Median level of gene expression reflected the scores of the metagenes. In addition, scores of 10 types of immune-related cells were calculated using the MCP counter R package (19).

HCC Sample Datasets

The RNA-seq data, SNP data, clinical data, and immune-associated genes in HCC were retrospectively collected from the TCGA database (20) (<https://cancergenome.nih.gov/>) and the NCBI GEO database (21) (<https://www.ncbi.nlm.nih.gov/geo/profile/>). The gene expression profiles of clinical data from 257 HCCs were retrieved from the TCGA database with the following criteria: (a) at stage of I and II; (b) accompanied by detailed follow-up information; (c) accompanied by HCC gene expression profiles; (d) genes with expression levels >0 in each sample accounting for more than 30% of the genes identified in the immune gene set (22, 23). Disease-free survival (DFS) and progression-free survival (PFS) were analyzed for the TCGA cohort.

The external validation cohort included 445 HCCs that were collected by the NCI from the GSE14520 dataset (24, 25). A total of 170 HCCs with the desired gene expression profiles and at stage I and II were included in the analysis. The other external validation cohort was from ICGC, which included 142 HCCs (Table S1).

Data Preprocessing

The RNA-seq data were analyzed using the Illumina platform. The fragments per kilobase of gene per million fragments mapped with upper quartile normalization (FPKM-*uq*) and single-nucleotide polymorphism (SNP) from TCGA Data Portal were downloaded. Next, gene annotation was performed using the Ensemble database. The gene expression value was log₂-transformed for further exploration. The gene expression of the NCI cohort was calculated using Affymetrix HT Human Genome U133A Array and Affymetrix Human Genome U133A 2.0 Array. The expression data and related clinical data of the validation set were obtained from the Gene Expression Omnibus (GEO) (GSE14520). Probe annotations were downloaded from

the GEO database. The mRNA expression data and clinical information were downloaded from ICGC. Entrez Gene IDs were used for gene expression data analysis in the three cohorts. The corresponding scores of six types of immune-related HCC cells in each sample were downloaded from Timer database (<https://cistrome.shinyapps.io/timer/>). The immune score, the matrix scores, and identification of immune gene expression profiles were achieved using the R package estimate (19).

Identification of HCC Subtypes Based on the Immune Genes

The ConsensusClusterPlus package (26) was utilized to perform consistent clustering and screening of molecular subtypes based on immune gene expression profiles. The Euclidean distance is used to calculate the similarity distance between samples (27), and K-means is used for clustering (28). The clustering was performed using 100 iterations, with each iteration containing 80% of samples. The optimal cluster number was determined by cumulative distribution function (CDF) curves of the consensus score (29). SigClust analysis was applied for pairwise comparisons to test the significance of clustering among identified subtypes (30). The genes with high expression in some subtypes were identified using Kolmogorov–Smirnov test. Bonferroni correction was applied for multiple testing. The Benjamini–Hochberg method was used to calculate the false discovery rate (FDR), and genes with $FDR < 0.05$ were considered to be significantly upregulated genes. The top 100 upregulated genes in each subtype were selected and subjected to three-dimensional principal component analysis (PCA) to distinguish different molecular subtypes (31). PCA is a statistical method used to determine the main variables in a multidimensional dataset, which represent the differences among observations (32). Thus, several key clusters rather than all of the selected upregulated genes were utilized to classify the subtypes using PCA.

Gene Co-expression Network Analysis

The common pathways related to the six gene modules were determined using the WGCNA R package (33). The construction of the WGCNA network and module detection were conducted using an unsigned type of topological overlap matrix (TOM), a power β of 3, a minimal module size of 30, and a branch merge cutoff height of 0.25. KEGG enrichment analysis was performed using R package clusterProfiler with $FDR < 0.05$. The most significant correlated genes with WGCNA edge weight >0.15 were visualized using Cytoscape 3.7.1 (34).

Validation of Five Immune-Related Subtypes

To validate five immune-related subtypes identified from the TCGA cohort, the genes in the co-expression gene modules (blue, brown) that are closely related to the C3 and C4 subtypes were selected, and the correlation between the genes and the modules was calculated, and then the cancer samples in NCI cohort were classified based on the featured genes with the correlation coefficient >0.8 . Moreover, the gene expression profiles extracted from the validation set were used to classify the samples by the

Support Vector Machine (SVM). To further validate the five immune-related subtypes, the normalized data of 445 samples were downloaded from the GSE14520 dataset and 142 samples from the ICGC. A total of 312 samples at stages I and II and containing the gene expression profiles of the featured genes were extracted and classified using SVM.

Statistical Analysis

The relationship between clinical variables and subtypes was analyzed by chi-square test or Fisher's exact test. Multiple testing was corrected by Benjamini–Hochberg's FDR. Kaplan–Meier curves and log-rank test were used to compare the 10-year DFS and PFS rates of the five immune subtypes. All tests were two sided, and $P < 0.05$ was considered to be statistically significant. Student's *t*-test was used to compare the immune scores and expression values of checkpoint genes among the HCC subtypes. The FDR correction was performed to decrease false-positive rates in multiple tests. All statistical tests were two-sided. All statistical analyses were performed using R software (version 3.5.3, <http://www.R-project.org>).

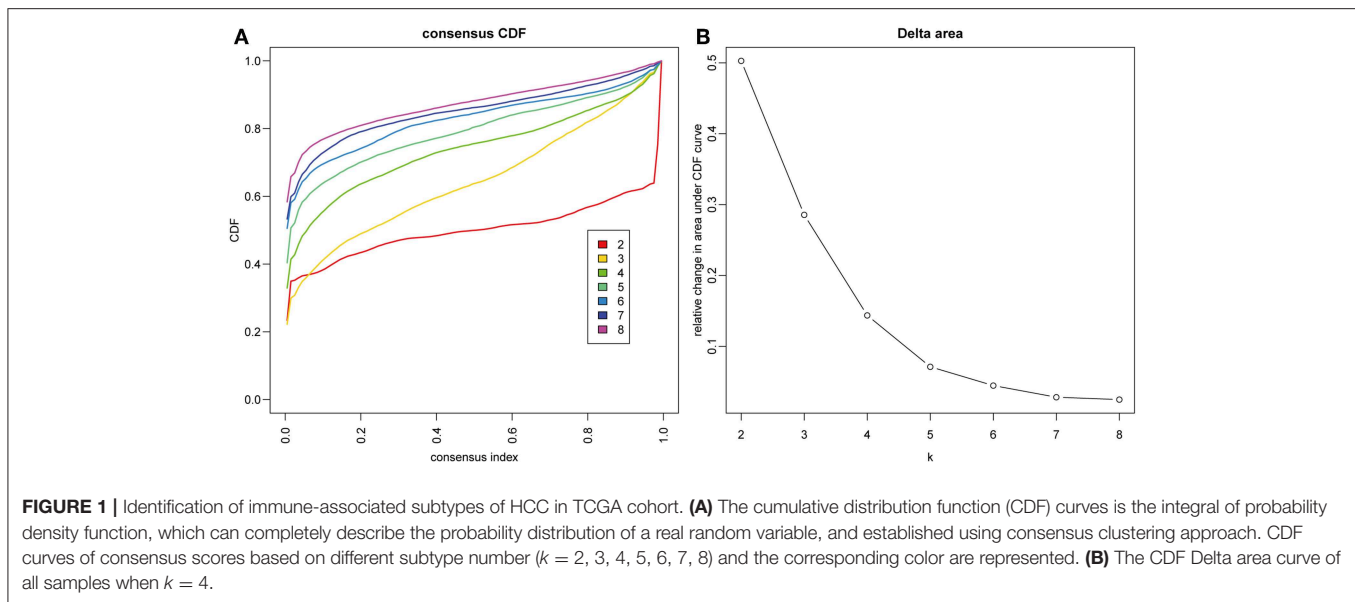
RESULTS

Identification of HCC Subtypes Based on the Immune Genes

The gene expression profiles of 778 immune-associated genes were used to investigate the HCC subtypes from the TCGA cohort. All tumor samples were divided into k ($k = 2, 3, 4, 5, 6, 7, 8$) different subtypes using Consensus Cluster Plus. The optimal division was reached when $k = 5$ based on the CDF curves of the consensus score (Figure 1A). Moreover, SigClust analysis showed that the consensus clusters ($k = 5$) were significant in all the pairwise comparisons (Figure 1B). There was no significant difference in expression distribution of C1 vs. C5, C4 vs. C5, and C2 vs. C3 ($P < 0.05$). However, significant edge effects were detected in C1 vs. C2, C1 vs. C3, C2 vs. C4, and C3 vs. C5. Thus, the five clusters of samples were separated and the 257 HCC tumor samples extracted from the TCGA cohort were classified into five molecular subtypes underlying the whole immune gene expression profile (Figure 2A).

Five Subtypes Were Characterized in Immune Microenvironment

The upregulated immune-related genes in each molecular subtype in comparison with other subtypes were analyzed using Kolmogorov–Smirnov test ($FDR < 0.05$). Among the 778 immune-related genes, 230 genes in subtype C1, 64 genes in subtype C2, 118 genes in subtype C3, 125 genes in subtype C4, and 72 genes in subtype C5 were significantly upregulated. More importantly, 54 genes overlapped between subtypes C3 and C4, and 54 genes overlapped between subtypes C1 and C4. However, only a few overlapped genes were identified in the other pairs of subtypes (Figure 2B). The top 100 upregulated genes from each subtype were selected for three-dimensional PCA and the top two principal components were extracted and visualized using a scatter plot (Figure 2C). PCA results demonstrated that these genes were clearly classified into five subtypes. To further



identify the gene expression pattern of each subtype, the selected top 100 genes in each subtype were examined with a heatmap (**Figure 2D**). The heatmap results showed a distinct expression pattern in the immune upregulated gene profiles of each subtype.

Clinical Characteristics of the Five Subtypes

To investigate the clinical relevance of tumor microenvironment, clinical factors, including age, gender, tumor, node, metastasis (TNM) staging, and stage in the five subtypes, were analyzed. There was a significant difference in age distribution when the age threshold was set as 60 (chi-square distribution test, $P = 0.00019$) (**Figure 3A**). The average age in subtype C2 was relatively lower, while it was relatively higher in subtype C5 as compared to other subtypes (**Figure 3A**). Moreover, the stage relevance in the five subtypes was further analyzed (**Figure 3B**). The proportion of stage I in subtype C2 and the proportion of stage II in subtype C3 were significantly lower (chi-square distribution test, $P < 0.001$). Furthermore, the grade distribution in the different subtypes was estimated (**Figure 3C**), and it was revealed that the proportion of G1 in subtype C3 and the proportion of G3 and G4 was significantly higher in comparison with other subtypes ($P < 0.001$).

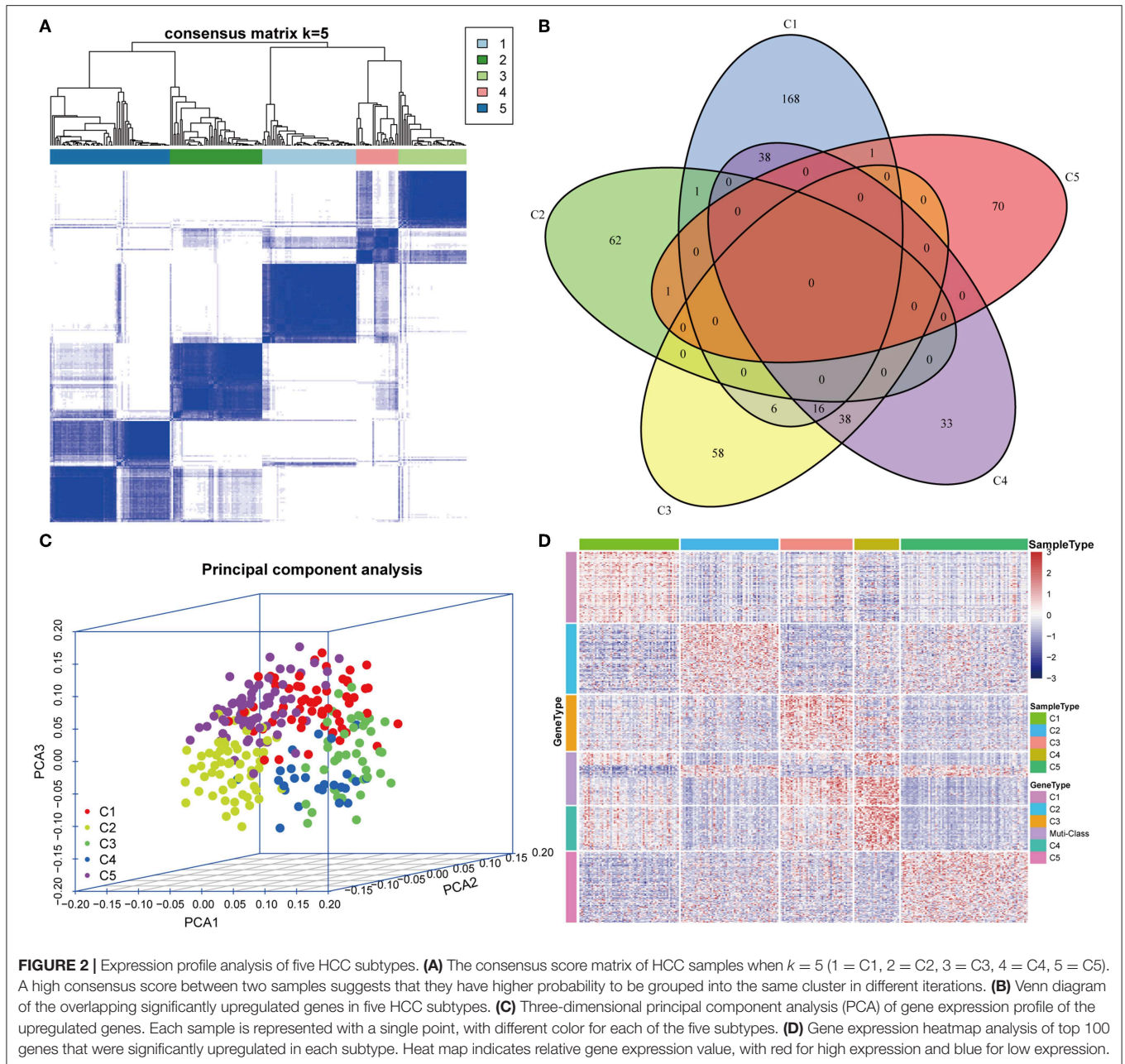
Of the TNM staging, only one type could not be compared further in the different subtypes in terms of the node and metastasis. The tumor staging in the five subtypes is shown in **Figure 3D**, and the proportion of T2 in subtype C2 was significantly higher in comparison with the other subtypes ($P < 0.001$). In addition, the relationship between subtypes and gender was analyzed in **Figure 3E**. The result demonstrated that the proportion of females in subtype C1 and the proportion of males in subtype C2 were significantly higher in comparison with other subtypes ($P < 0.001$). The molecular subtypes of HBV/HCV were further explored for comprehensive genomic analysis of HCCs that were reported in the five subtypes (**Figures 3G,H**).

According to previous studies (35), we classified the five subtypes into three iclusters (**Figure 3F**). The proportion of icluster1 in subtype C2 and the proportion of icluster3 in subtype C5 were significantly higher in comparison with the other subtypes ($P < 0.001$). However, there was no significant difference in the HBV/HCV proportions among all the subtypes.

Tumor Immunogenicity of HCC

The potential immune escape mechanisms of HCC in the five subtypes were further explored. The scores of 13 types of immune metagenes, tumor immune component (matrix, immunity, tumor purity), 6 types of immune infiltrating cells, and 10 types of immune cell-related MCP counter were collected. Most of the metagenes were overexpressed in subtype C4 and underexpressed in subtype C5 (**Figures 4A,B**). The comprehensive immune component score was significantly higher in subtype C4 and was significantly lower in subtype C5 compared with other subtypes (**Figures 4C,D**). Notably, the matrix score in subtype C3 was relatively higher when compared with other subtypes (**Figures 4C,D**). Of the 10 types of immune cell-related MCP counter, the scores of T cells, and CD8 cells were significantly higher in subtype C4 and were significantly lower in subtype C5 in comparison with the other groups (**Figures 4E,F**). Of the six types of immune-infiltrating cells, the scores of B cell, CB8 cell, neutrophil, dendritic, and macrophage were significantly higher in subtype C4 and were significantly lower in subtype C5 (**Figures 4G,H**). In summary, most of the immune signatures were upregulated in subtype C4 and downregulated in subtype C5 in comparison with the other subtypes, which suggested that subtype C4 and C5 had enhanced immune profile and a decreased immune profile (**Figure S1**).

The expression profiles of eight immune checkpoint genes, which play a key role in immune modulation, were further examined. As shown in **Figure 8A**, PDCD1, CD274, CTLA4, CD86, and CD80 were upregulated significantly in subtype C4 in comparison with the other subtypes, while CD276 was

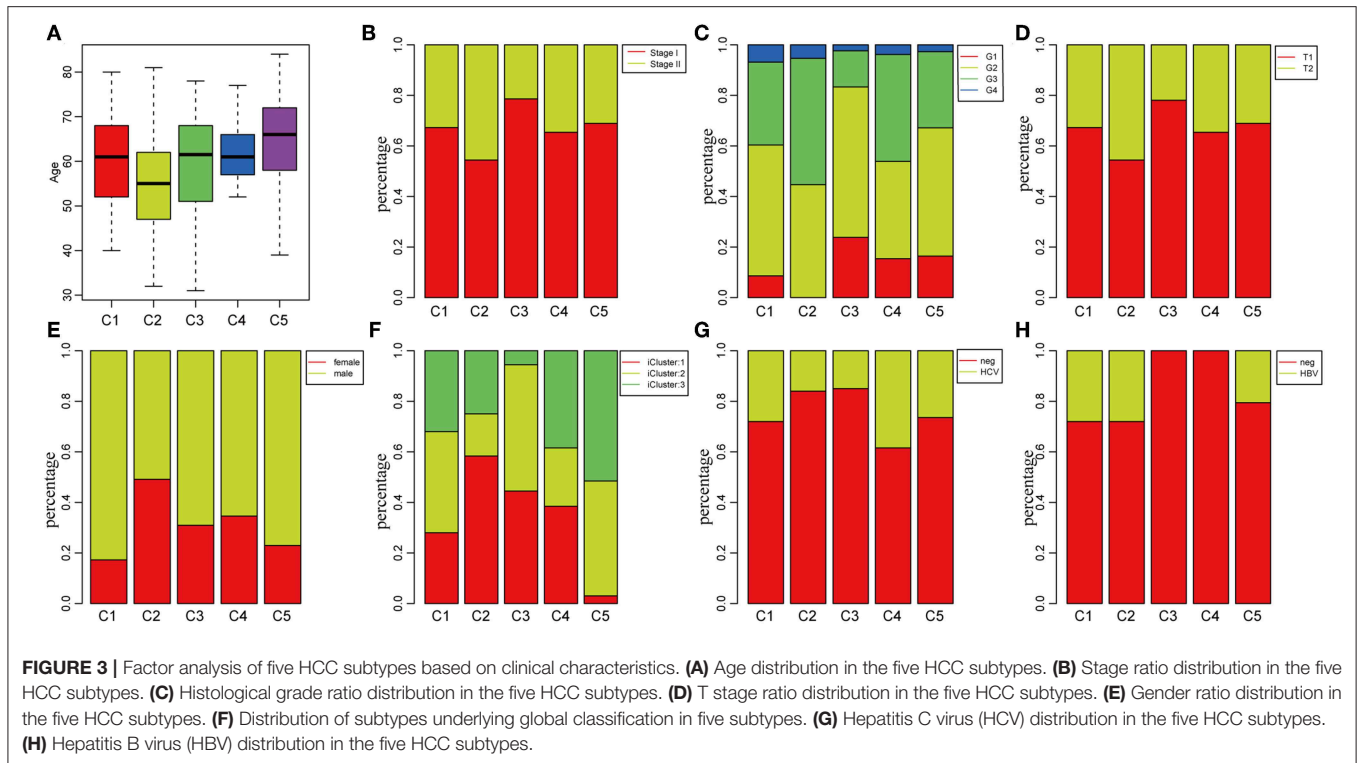


significantly upregulated in subtype C2 in comparison with the other subtypes. Subsequently, the expression values of eight immune checkpoint genes in the five subtypes (**Figure 6B**) were analyzed. There was a significant difference in the expression value of these checkpoint genes except in VTCN1.

Prognostic Significance of the Five Molecular Subtypes

The poor prognosis in HCC is mainly caused by high incidence of tumor progression and disease recurrence (36). Based on the differently expressed immune profile, the association between clinical outcome of HCC patients and the five subtypes was subsequently investigated. In the TCGA cohort, Kaplan–Meier

curves suggested significantly different DFS (log-rank test, DFS, $P = 0.0486$, **Figure 5A**) and PFS (log-rank test, $P = 0.04426$, **Figure 5B**) of the HCC patients in the five subtypes. Patients in subtype C5 had the worst outcome, and HCC patients in subtype C3 had the best outcome among all the five subtypes underlying both DFS and PFS. Notably, subtype C4 had the highest immune score but did not show the best outcome, which could have been as a result of the relatively small sample size. The samples in subtype C3 and C4, which had relatively high immune scores, were combined and the clinical outcome of HCC patients in the four groups were further investigated. More significant difference in DFS (log-rank test, $P = 0.04318$) and PFS (log-rank test, $P = 0.03802$) are shown in **Figures 5C,D**,



suggesting that high immune scores in early liver cancer could be a protective factor in early HCC. To explore the relationship between clinical outcome and immune score, DFS and PFS in subtypes C3/C4 with high immune scores and subtype C5 with a low immune score were compared and it was found that subtypes C3/C4, two immune-enhanced molecular subtypes, showed better clinical outcome (log-rank test, DFS, $P = 0.0573$; PFS, $P = 0.00444$, **Figures 5E,F**).

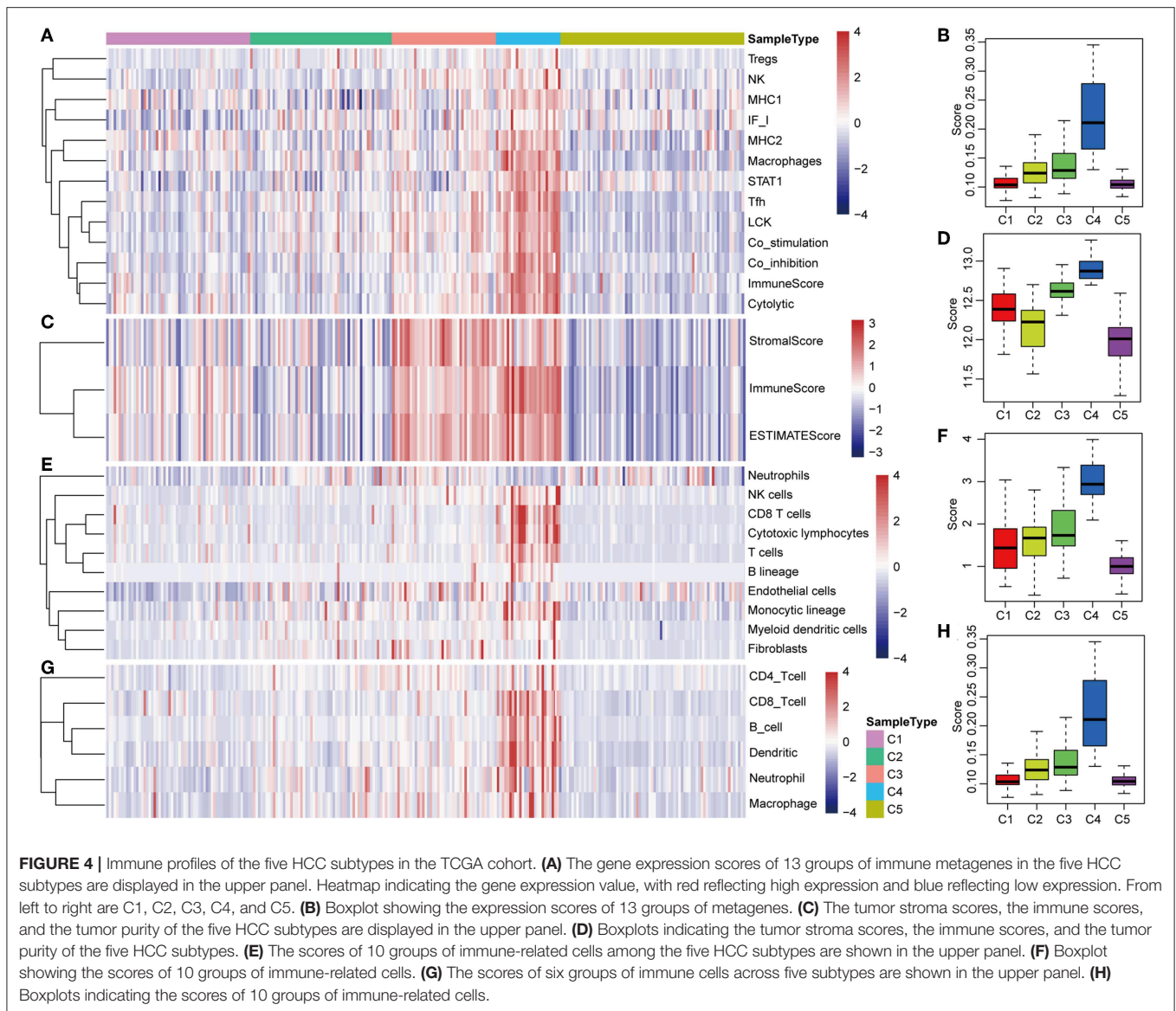
Frequencies of Mutant Genes of the Five Molecular Subtypes

Previous studies have reported that TP53 (37), CTNNB1 (38), and AXIN1 (39–41) mutations are closely related to the development of HCC. The genomic mutations of these three genes in the five subtypes were subsequently investigated. The results showed that there was significant difference in sample proportions of TP53, CTNNB1, and AXIN1 mutations as well as non-mutations among the five subtypes (**Figures 6A–C**). The proportion of TP53 mutations in the subtype C3 was significantly lower than that of the other subgroups; the proportion of CTNNB1 mutations in subtypes C2 and C3 was significantly lower than that of the other subtypes; and the proportion of AXIN1 mutations in the subtypes C2 was significantly higher than that of the other subtypes. Notably, there were no mutations in subtype C3. Moreover, there was a significant difference in the frequencies of the mutant genes in the five subtypes (**Figure 6D**). Similarly, the mutant frequencies in the subtype C3 were significantly lower compared with that of the other subtypes ($P = 0.02$). We further analyzed the relationship between the expression of 8 immune checkpoints in these five

subtypes. As shown in **Figure 7A**, PDCD1, CD274, PDCD1LG2, CTLA4, CD86, and CD80 were significantly higher in C4 than in other types, and CD276 was expressed higher in C2 than in other groups. Further analysis of the gene expression distribution of the eight immune checkpoints is shown in **Figure 7B**. It can be seen that there is a significant difference in the expression distribution among the five types of samples except VTCN1.

Gene Co-expression Network Analysis

To further explore the potential markers associated with the immune microenvironment of HCC, the data on the expression profiles of a total of 492 immune-related differentially expressed genes were obtained, and the distance between different transcripts was calculated using the Pearson correlation coefficient. The scale-free co-expression network is that logarithm $\log(k)$ of the node with the connection degree k is negatively correlated with the logarithm $\log(P(k))$ of the probability of the node, and the correlation coefficient is >0.8 . To construct a scale-free network, the value of β was set as 3 (**Figures 8A,B**). The expression matrix was then converted into an adjacency matrix, and the adjacency matrix was converted into a topological matrix. Based on TOM, the average-linkage hierarchical clustering method was used to cluster the genes according to the criteria of the hybrid dynamic cut tree, and the minimum number of genes was set at 30 per module. After applying the dynamic shear method to determine the gene module, the eigengenes of each module were calculated and the cluster analysis for each module was performed. The closer modules were then merged into new modules, and set height =



0.25, deepSplit = 2, and minModuleSize = 30. Finally, a total of seven modules with all immune-related differentially expressed genes were identified using WGCNA (Figure 8C). Notably, the gray module was a collection of genes that could not be aggregated into other modules. The transcript statistics of each module are shown in Table 1. A total of 371 transcripts were divided into five co-expression modules. The correlation between the eigengenes of six modules and five subtypes (Figure 8D) was calculated. The blue module was positively correlated with subtypes C3 and C4, and was negatively correlated with subtype C5; the yellow module was positively correlated with subtype C3, and was negatively correlated with subtype C5; the turquoise module was positively correlated with subtypes C1, C3, and C4, and was negatively correlated with subtype C5; the green module was positively correlated with subtype C2, and was negatively correlated with subtype C3; the brown and red

modules were positively correlated with subtype C1, and was negatively correlated with subtype C2.

To investigate the biological functions of the five modules, KEGG enrichment analysis was performed. The brown module was mainly enriched in the B cell receptor signaling pathway; the yellow module was mainly enriched in eight pathways closely associated with cancer, such as EGFR tyrosine kinase inhibitor resistance, small cell lung cancer, focal adhesion, and other pathways; the blue module was mainly enriched in 24 pathways, including inflammation-related pathways, such as *Staphylococcus aureus* infection, and intestinal immune network for IgA production; the red module was mainly enriched in cell adhesion molecules (CAMs), transcriptional misregulation in cancer, and inflammatory mediator regulation of TRP channels; while the turquoise module was mainly enriched in 21 pathways, such as T cell receptor signaling pathway, CAMs, Th1, and Th2

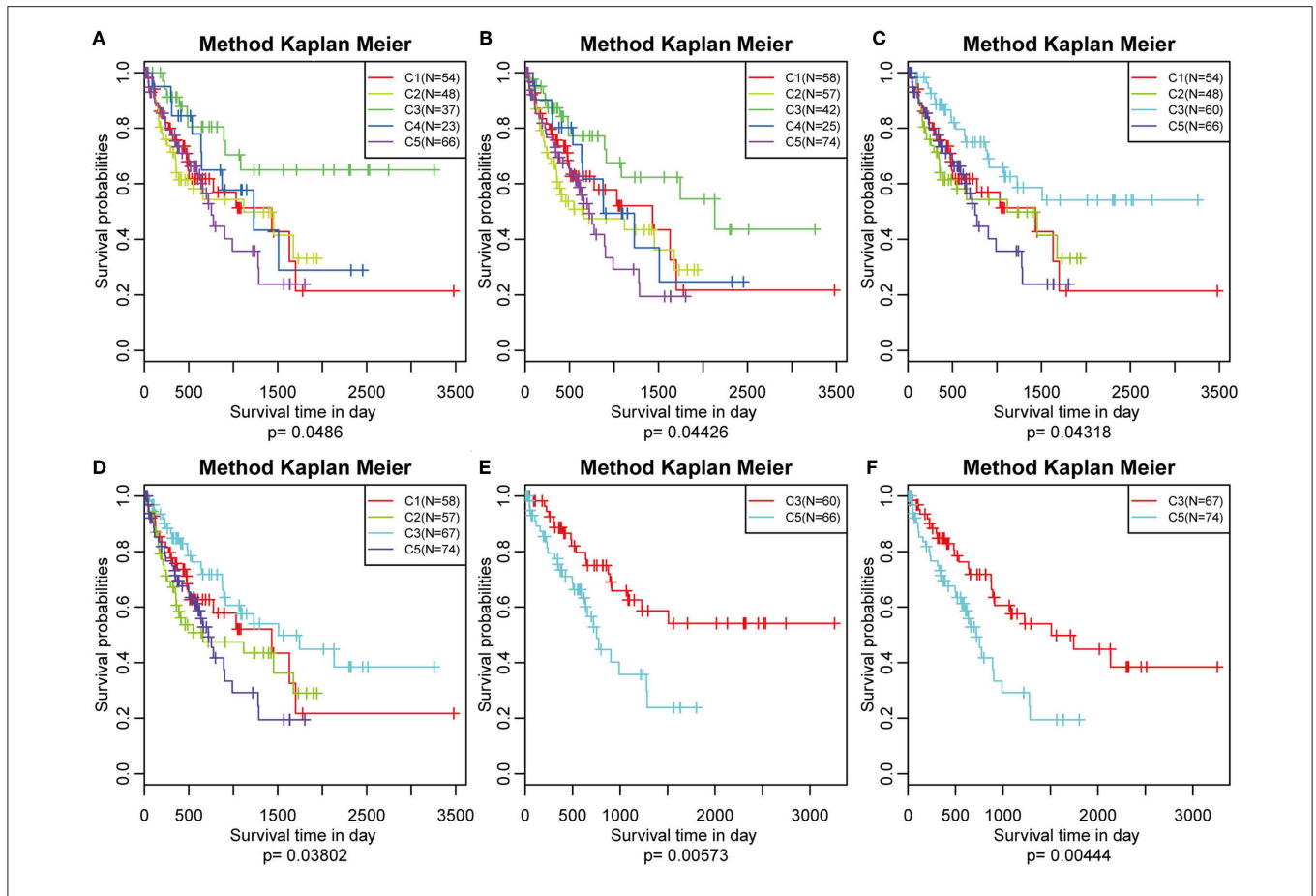


FIGURE 5 | Survival analysis of the five HCC subtypes. Kaplan–Meier curves showing the distinct outcomes of HCC patients in the five molecular subtypes. The *P*-value was calculated using the log-rank test, by comparing the overall five subtypes together or comparing C1, C2, C3, and C5 or subtype C3 vs. subtype C5 only. **(A)** KM curve of disease-free survival (DFS) prognosis of five subtypes. **(B)** KM curve of non-progressive survival prognosis of five subtypes. **(C)** The prognosis difference KM curve of the non-DFS of merged C3, C4, and other subtypes. **(D)** The prognosis difference KM curve of the non-progression-free survival (PFS) of merged C3, C4, and others subtypes. **(E)** The prognosis difference KM curve of the non-DFS of merged C3, C4, and C5. **(F)** The prognosis difference KM curve of the non-PFS of merged C3, C4, and C5.

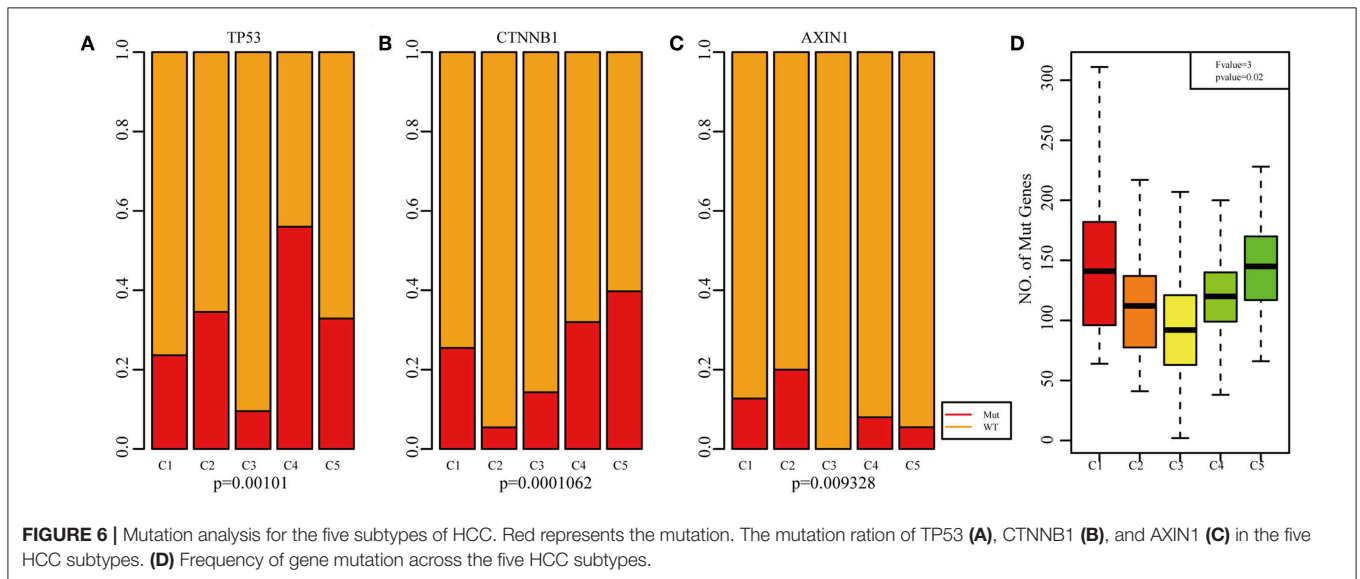
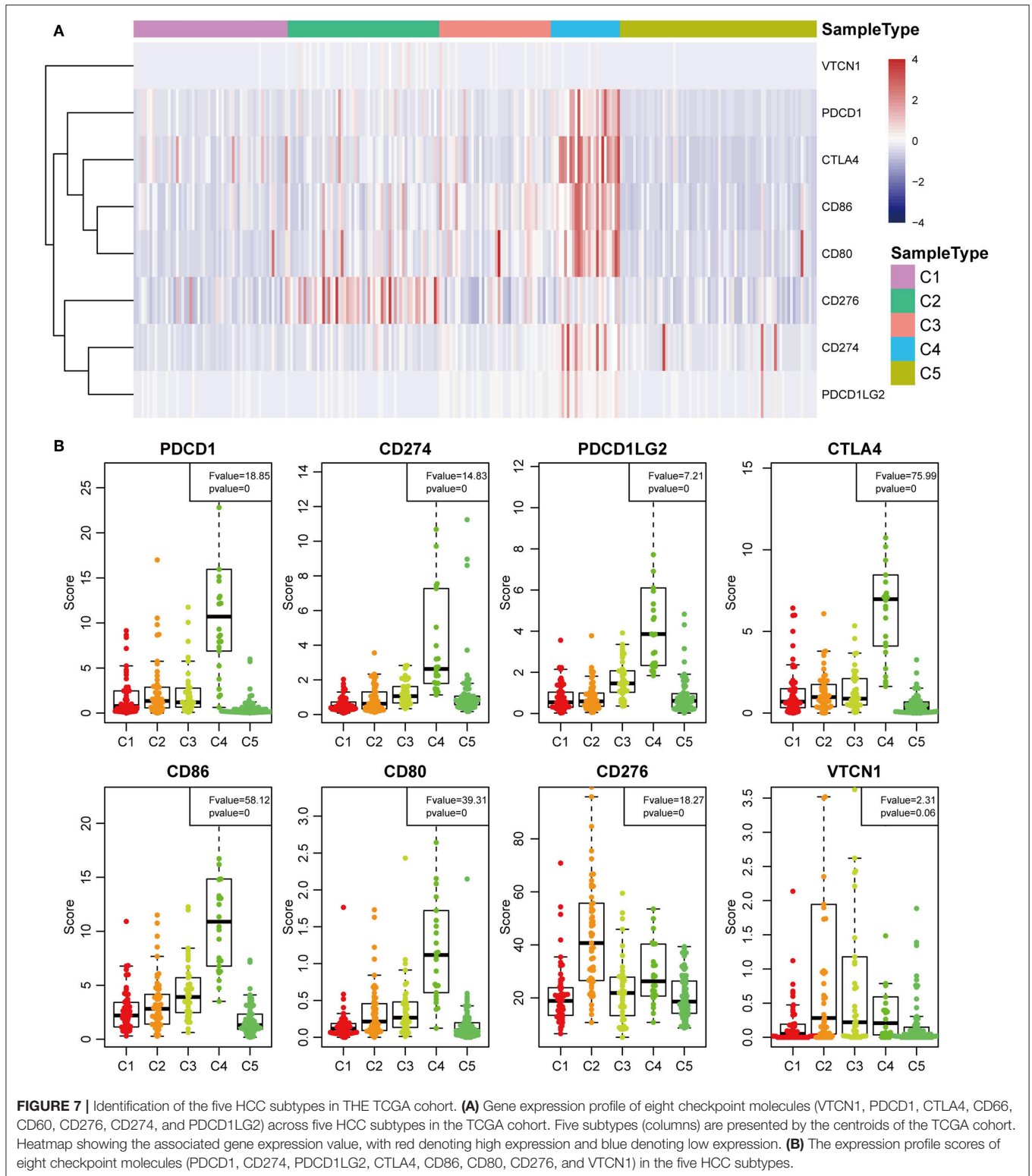


FIGURE 6 | Mutation analysis for the five subtypes of HCC. Red represents the mutation. The mutation ratio of TP53 **(A)**, CTNNB1 **(B)**, and AXIN1 **(C)** in the five HCC subtypes. **(D)** Frequency of gene mutation across the five HCC subtypes.



cell differentiation, cytokine–cytokine receptor interaction and other immune-related pathways.

Subsequently, the relationship network of enriched pathways in these modules was visualized. As shown in **Figure 9**, it was

found that these modules were enriched in 44 pathways. There were 12 common pathways enriched in blue and turquoise modules. Few intersections of the enriched pathways were enriched in other modules, suggesting that the genes in the blue

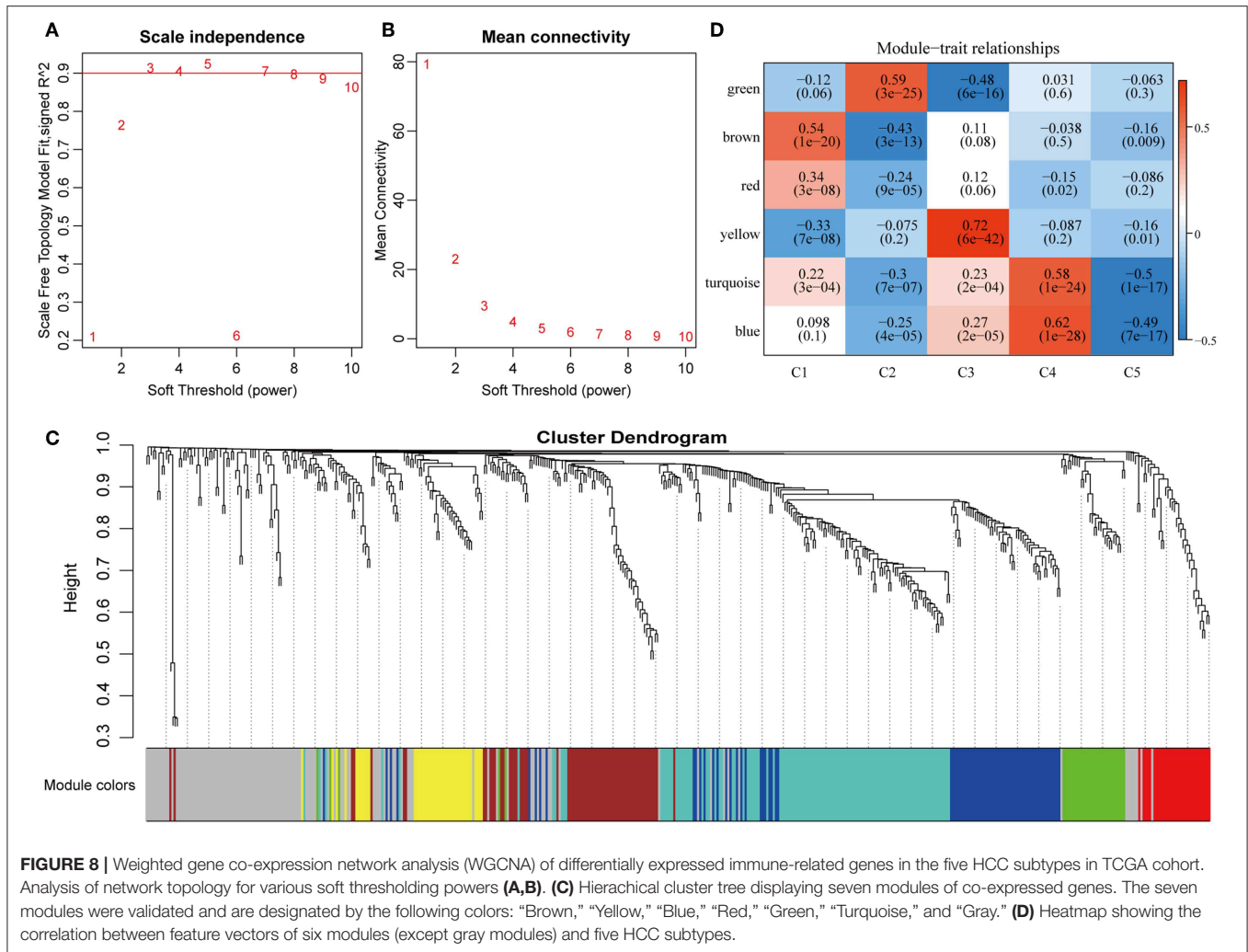


TABLE 1 | Transcript data for six modules.

Modules	Genes
Blue	77
Brown	66
Green	34
Red	31
Turquoise	121
Yellow	42

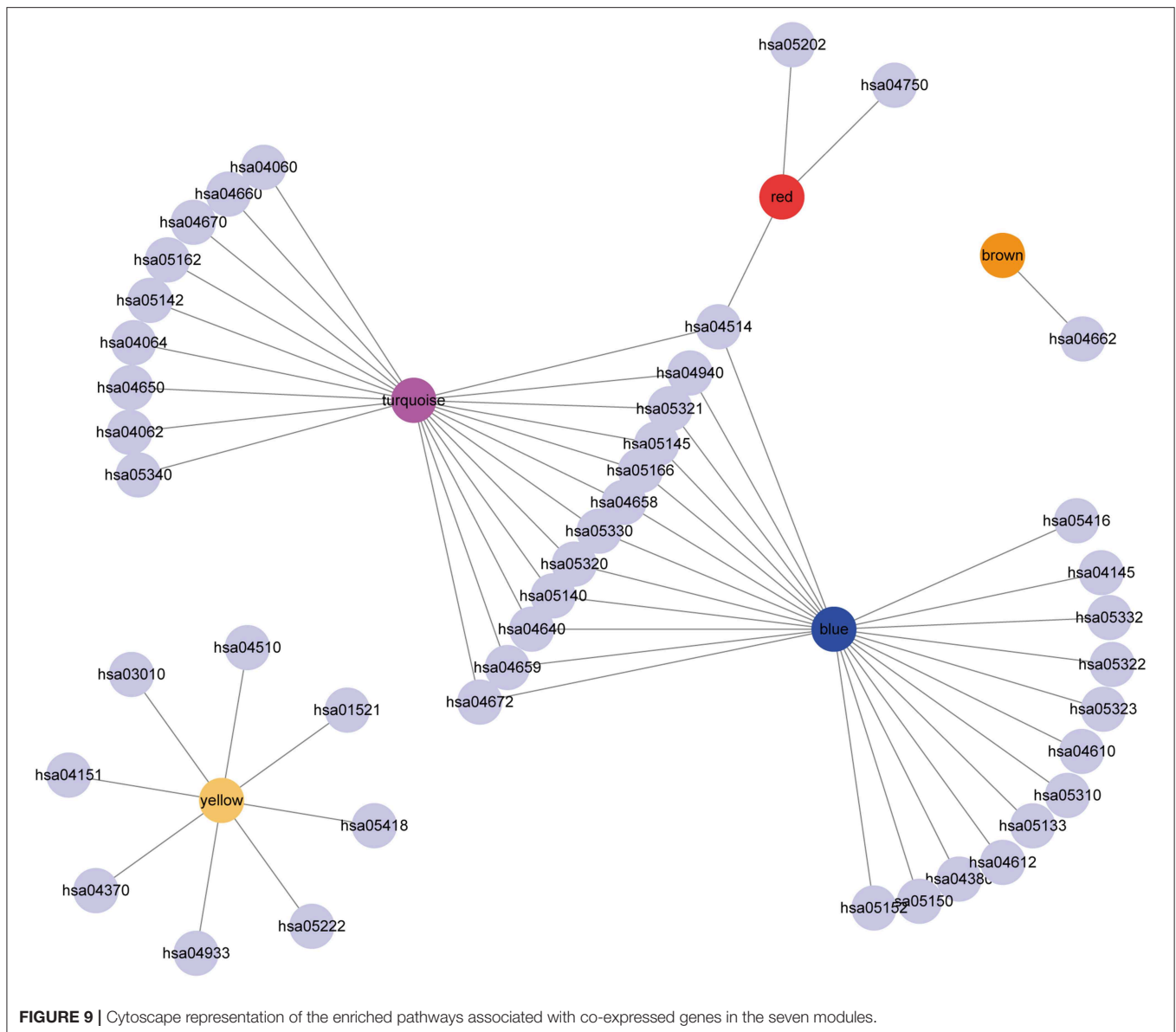
and turquoise modules may share similar regulatory processes in the five subtypes.

External Validation of the Five Subtypes

A total of 73 featured genes with a correlation coefficient >0.8 were obtained from the co-expression gene modules (blue, brown). The expression profile of the featured genes was further extracted to serve as a training set and employed SVM to classify the samples. The classification accuracy rate of SVM was 91.1%. To further validate the five subtypes, 170 samples were classified

using SVM. Thirty-nine samples were predicted in subtype C1, 40 samples were predicted in subtype C2, 18 samples were predicted in subtype C3, 29 samples were predicted in subtype C4, and 44 samples were predicted in subtype C5.

The expression distribution of 13 immune metagenes in the five subtypes was subsequently analyzed as was shown in **Figures 10A–D**. Most of the metagenes were highly expressed in subtype C4, which was consistent with the result in training set. Moreover, the immune scores were investigated as was shown in **Figure 10B**. It was found that the immune score in subtype C4 group was significantly higher as compared with the other subtypes, and the matrix score of subtype C3 was significantly higher in comparison with the other subtypes, which was also consistent with the training set. Further analysis of the sample immune scores is shown in **Figures 10C,D**; the results showed that the immune score of subtype C4 was significantly higher in comparison with the other subtypes, and the matrix score of subtype C3 was significantly higher than that of the other subtypes, which was consistent with the training set. Moreover, the distribution of 10 immune-related cells in the five subtypes of samples was analyzed as shown in **Figures 10E,F**. Similar with the training set, it was found that the proportion of most

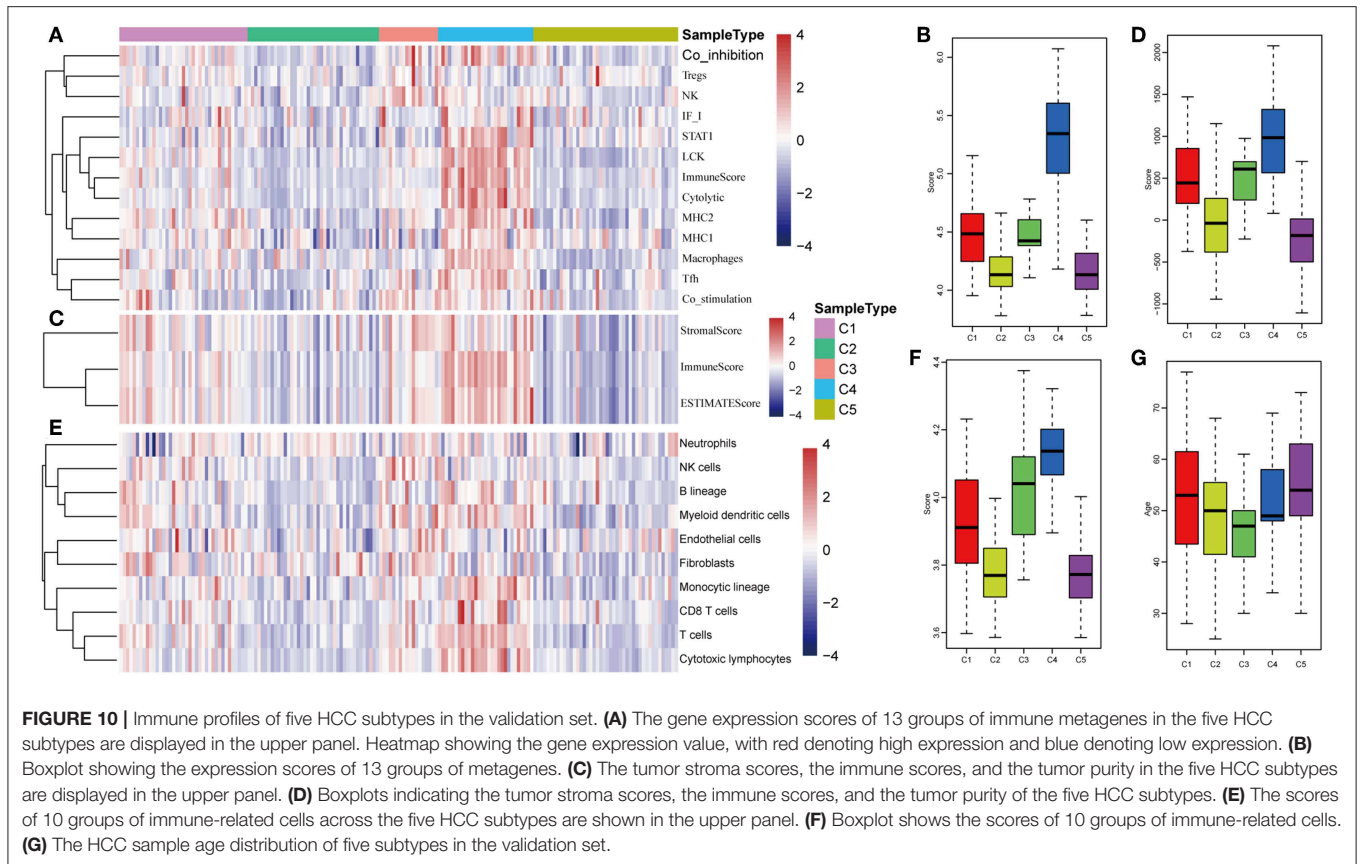


of the cells was higher in subtype C4 as compared with the other subtypes. Finally, the age distribution of the five subtypes was analyzed and the post-distribution difference is shown in **Figure 10G**. It was found that the age distribution of the five subtypes was also consistent with that of the training set. Similar with these results, the expression profiles of immune metagenes in the five HCC subtypes were further validated in the ICGC database in **Figures 11A–F**. These data suggested that there were immune-enhanced subtypes and immune-decreased subtypes in early HCC, and there was a significant difference between the two subtypes.

DISCUSSION

HCC is an aggressive malignancy, which is still the third leading cause of tumor-related deaths. Although there have

been advances in treatment strategies, no effective molecular targeted therapy has been successfully validated. The intricate microenvironment, sustained by the production of growth factors with parenchyma, as well as the infection of hepatitis viruses, promotes the occurrence and development of HCC (42). The current study aimed at systematically analyzing the heterogeneous HCC microenvironment subtypes underlying global immune genes in stage I/II and related clinical significance using multi-omics data extracted from the TCGA cohort. Five molecular subtypes in immune microenvironment were found to exhibit significantly different clinical characteristics, immune escape mechanisms, genomic alterations, signaling, and outcomes. Subtype C4 was found to be an immune-enhanced subtype while subtype C5 was found to be an immune-decreased subtype in the immune microenvironment of HCC. The TP53, CTNNB1, and AXIN1 mutations were closely associated

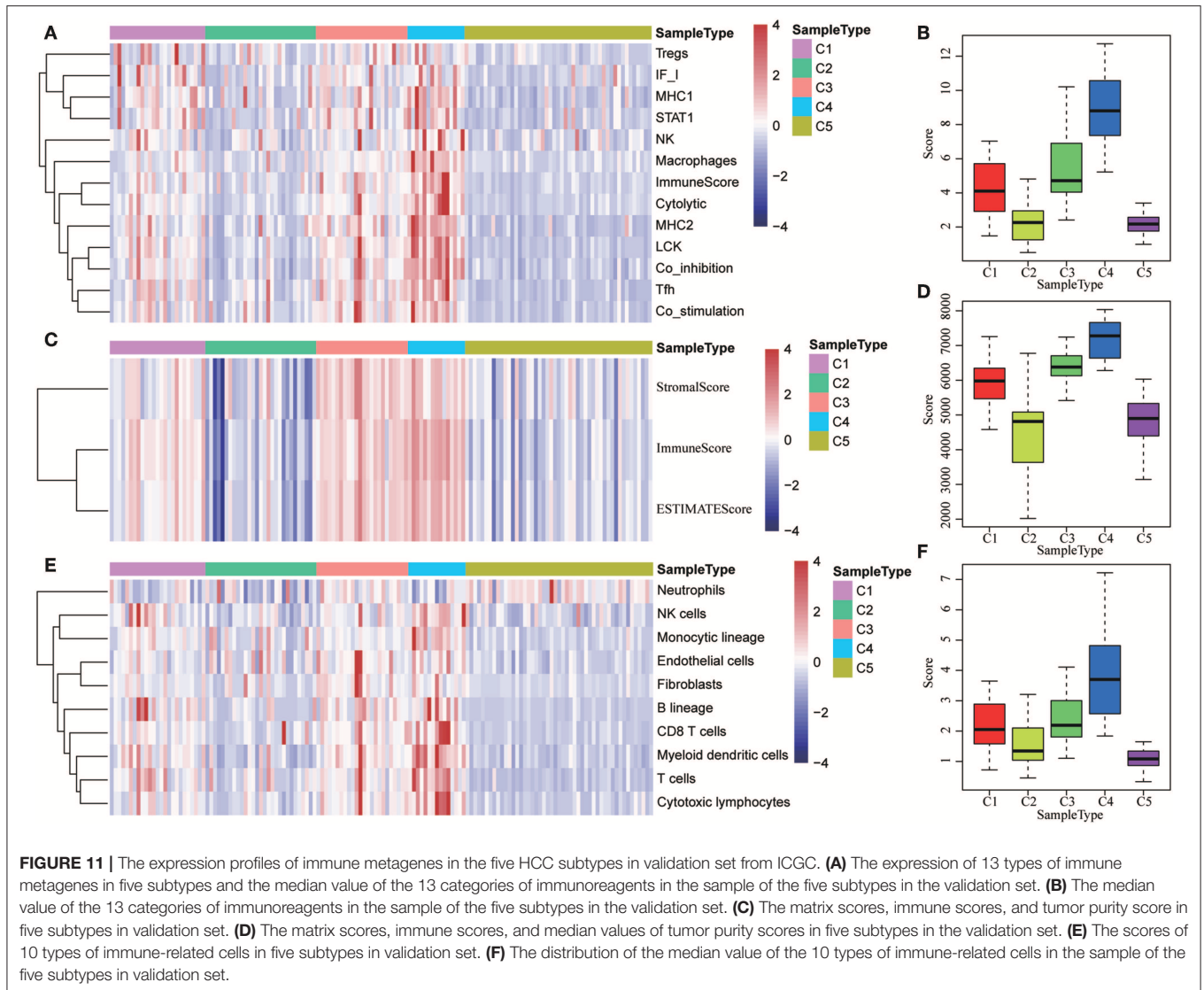


with immune-enhanced molecular subtypes. Finally, these five subtypes were validated in an external cohort of NCI.

Many previous studies have revealed some prognostic subtypes and histological subtypes of HCCs underlying the genome-wide profiles (43–45). The current study mainly aimed to investigate the global immune profiles for a more comprehensive analysis of immune landscape in HCC. A total of five molecular subtypes were identified. The results revealed that the immune profile of subtype C4 was significantly higher in comparison with that of the other molecular subtypes, and the immune profile of subtype C5 was significantly lower than that of the other subtypes. Moreover, subtype C4 showed increased immune cell infiltration score (including tumor matrix, immunity, purity), which was positively related to the expression signatures of multiple types of immune-related cells. These results demonstrated that subtype C2 was correlated with an enhanced immune status in the HCC immune microenvironment. On the contrary, subtype C5 showed lower expression of immune profiles and had the lowest immune cell infiltration score among the five subtypes, which was negatively related to expression signatures of the selected immune-related cells. Consistent with the findings of the training set, the validation set also revealed that most metagenes were highly expressed and showed higher immune scores in subtype C4 than that of the other four subtypes. Thus, it was hypothesized that both immune-enhanced subtype and immune-decreased subtype existed in the HCC immune microenvironment and were

significantly different in terms of expression profile of metagenes, the immune components score, the immune infiltration score, and the MCP counter of immune cells.

This study further investigated the potential immune escape molecular mechanisms of HCC. It is well-known that there are two main aspects of intrinsic immune escape, including immunogenicity and the expression of immune checkpoint molecules (46). Mutations of TP53, CTNNB1, and AXIN1 gene in HCC were validated by whole genome, exome, and transcriptome sequencing (40, 41, 47). However, the roles played by these gene mutations in the molecular mechanism of immune microenvironment of HCC have not been studied to date. This study revealed that the five molecular subtypes were significantly different in terms of the mutations and frequencies of TP53, CTNNB1, and ACIN1. Previous studies reported that immune checkpoints are often activated in HCC with high immune response accompanied with upregulated gene expression (48). The most common immune checkpoint receptors are BTLA, VISTA, PD-1, CTLA-4, LAG-3, TIM-3, and OX40. The current study revealed significant overexpression of checkpoint genes, including PDCD1, CD274, PDCD1LG2, CTLA4, CD86, and CD80 in subtype C4 than in the other subtypes. These results suggest that the immune-enhanced subtype C4 may be closely associated with intrinsic immune escape of HCC, which may provide new insights into immunotherapy of HCC using immune checkpoint blockers.



Valerie and colleagues reported that the survival outcome of HCC patients is positively correlated with higher expression of a group of innate immune-related and inflammatory genes, including NK-associated molecules and macrophage (13). Moreover, many other targets have been considered as prognostic indicators for HCC, including immunoregulatory enzyme indoleamine 2,3-dioxygenase (IDO) (49), epithelial neutrophil-activating peptide-78 (CXCL5) (50), CXCR6 (51), and so on. It was found in the current study that these five molecular subtypes were consistently associated with different survival outcomes in HCC patients. Subtype C5 had the worst survival outcome while subtype C3 had the best outcome among all the five subtypes. Notably, subtypes C3 and C4, which had higher immune scores, showed better survival outcome compared with subtype C5, which suggests that the subtype with high immune score may play a protective role in the early stage of HCC. Thus, this study may provide immune signature for survival prediction in the early stage of HCC.

The current research also explored the potential targets and pathways of the five subtypes in the HCC immune microenvironment using WGCNA analysis. The immune-related genes were enriched in six different modules, and the results revealed that inflammatory pathways were mainly enriched in the blue module, and the immune-related pathways were mostly enriched in the turquoise module. The blue and turquoise modules shared the most common pathways of all pairwise comparisons, which suggests that the genes in these two modules may play similar roles in the HCC immune microenvironment. Consistent with the current results, previous studies reported that the tumor immune microenvironment plays key roles in regulating the process of hepatocarcinogenesis, tumor invasion, and metastasis (52). Innate immunity plays a critical role in modulating HCC tumor occurrence and development because the liver is an organ with predominant innate immunity (53). It has also been hypothesized that the adaptive immune response may be induced in the progression of HCC (54). Therefore,

clarifying the molecular mechanisms based on the immune microenvironment of HCC may facilitate identification novel therapeutic and chemopreventive targets for HCC.

There are some limitations in the current study. Firstly, to comprehensively reflect the factors and effects influencing HCC microenvironment phenotypes, more clinical characteristics and demographic features of HCC patients should be included in subgroup analysis. Secondly, the sample size in each subtype was relatively small in training as well as validation set, and only the NCI and ICGC cohorts were used for external validation, which may have resulted in one-sided results and a high false-positive rate. It will be helpful to perform cross-validation in internal validation and increase the sample size for external validation in future studies on the immune microenvironment of HCC. Finally, more experimental evidence for immunogenomic analysis is needed to validate the roles of mutation genes, checkpoint genes, and the enriched pathways involved in immune microenvironment.

CONCLUSION

The current study suggests that microenvironment phenotypes of HCC could be classified into five molecular subtypes with potential immune escape mechanisms in HCC. These subtypes are distinct in immunity characteristics, immune checkpoint molecules, and patient outcomes. Moreover, specific functional pathways may drive the formation of microenvironment phenotypes. These results may provide guidance for developing novel strategies of immunotherapy in HCC.

REFERENCES

- Takayasu K, Aarii S, Sakamoto M, Matsuyama Y, Kudo M, Kaneko S, et al. Impact of resection and ablation for single hypovascular hepatocellular carcinoma ≤ 2 cm analysed with propensity score weighting. *Liver Int.* (2018) 38:484–93. doi: 10.1111/liv.13670
- Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer.* (2015) 136:E359–86. doi: 10.1002/ijc.29210
- Fattovich G, Stroffolini T, Zagni I, Donato F. Hepatocellular carcinoma in cirrhosis: incidence and risk factors. *Gastroenterology.* (2004) 127(5 Suppl 1):S35–50. doi: 10.1053/j.gastro.2004.09.014
- Sherman M. Hepatocellular carcinoma: epidemiology, risk factors, and screening. *Semin Liver Dis.* (2005) 25:143–54. doi: 10.1055/s-2005-871194
- Wang J, Ha J, Lopez A, Bhuket T, Liu, B, Wong RJ. Medicaid and uninsured hepatocellular carcinoma patients have more advanced tumor stage and are less likely to receive treatment. *J Clin Gastroenterol.* (2018) 52:437–43. doi: 10.1097/MCG.0000000000000859
- Ayuso C, Rimola J, Vilana R, Burrell M, Darnell A, Garcia-Criado A, et al. Diagnosis and staging of hepatocellular carcinoma (HCC): current guidelines. *Eur J Radiol.* (2018) 101:72–81. doi: 10.1016/j.ejrad.2018.01.025
- Pardee AD, Butterfield LH. Immunotherapy of hepatocellular carcinoma: unique challenges and clinical opportunities. *Oncoimmunology.* (2012) 1:48–55. doi: 10.4161/onci.1.1.18344
- Hernandez-Gea V, Toffanin S, Friedman SL, Llovet JM. Role of the microenvironment in the pathogenesis and treatment of hepatocellular carcinoma. *Gastroenterology.* (2013) 144:512–27. doi: 10.1053/j.gastro.2013.01.002

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: TCGA-LIHC, GSE14520, and ICGC-LIRI-JP.

AUTHOR CONTRIBUTIONS

JL and WL designed experiments and interpreted data. JL, HW, and ZM conducted bioinformatic and statistical analyses. HW, JZ, WO, and YQ wrote the paper. All authors have read and approved the manuscript for publication.

FUNDING

This work was supported by Grants from Special Science and Technology Innovation Strategy in Guangdong (201803011), the National Natural Science Foundation of China (Grant No. 81473369), and the Key R&D programs in Shandong (2016CYJS08A01-1).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2019.01019/full#supplementary-material>

Figure S1 | Boxplots indicating the infiltrating immune scores of all types of immune metagenes among the five subtypes.

Table S1 | Patients' information in the TCGA, GSE14520, and ICGC cohorts.

- Zamarron BF, Chen W. Dual roles of immune cells and their factors in cancer development and progression. *Int J Biol Sci.* (2011) 7:651–8. doi: 10.7150/ijbs.7.651
- Wolchok JD. PD-1 Blockers. *Cell.* (2015) 162:937. doi: 10.1016/j.cell.2015.07.045
- Chikuma S. CTLA-4, an essential immune-checkpoint for T-cell activation. *Curr Top Microbiol Immunol.* (2017) 410:99–126. doi: 10.1007/82_2017_61
- Makarova-Rusher OV, Medina-Echeverez J, Duffy AG, Greten TF. The yin and yang of evasion and immune activation in HCC. *J Hepatol.* (2015) 62:1420–9. doi: 10.1016/j.jhep.2015.02.038
- Chew V, Tow C, Teo M, Wong HL, Chan J, Gehring A, et al. Inflammatory tumour microenvironment is associated with superior survival in hepatocellular carcinoma patients. *J Hepatol.* (2010) 52:370–9. doi: 10.1016/j.jhep.2009.07.013
- Xiao Y, Ma D, Zhao S, Suo C, Shi J, Xue MZ, et al. Multi-omics profiling reveals distinct microenvironment characterization and suggests immune escape mechanisms of triple-negative breast cancer. *Clin Cancer Res.* (2019) 25:5002–14. doi: 10.1158/1078-0432.CCR-18-3524
- Cao B, Wang Q, Zhang H, Zhu G, Lang J. Two immune-enhanced molecular subtypes differ in inflammation, checkpoint signaling and outcome of advanced head and neck squamous cell carcinoma. *Oncoimmunology.* (2018) 7:e1392427. doi: 10.1080/2162402X.2017.1392427
- Budhu A, Forgues M, Ye QH, Jia HL, He P, Zanetti KA, et al. Prediction of venous metastases, recurrence, and prognosis in hepatocellular carcinoma based on a unique immune response signature of the liver microenvironment. *Cancer Cell.* (2006) 10:99–111. doi: 10.1016/j.ccr.2006.06.016
- Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, et al. TIMER: a web server for comprehensive analysis of tumor-infiltrating immune cells. *Cancer Res.* (2017) 77:e108–10. doi: 10.1158/0008-5472.CAN-17-0307

18. Li B, Severson E, Pignon JC, Zhao H, Li T, Novak J, et al. Comprehensive analyses of tumor immunity: implications for cancer immunotherapy. *Genome Biol.* (2016) 17:174. doi: 10.1186/s13059-016-1028-7
19. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol.* (2016) 17:218. doi: 10.1186/s13059-016-1070-5
20. Tomczak K, Czerwinski P, Wiznerowicz M. The cancer genome atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol.* (2015) 19:A68–77. doi: 10.5114/wo.2014.47136
21. Barrett T, Troup DB, Wilhite SE, Ledoux P, Rudnev D, Evangelista C, et al. NCBI GEO: mining tens of millions of expression profiles—database and tools update. *Nucleic Acids Res.* (2007) 35:D760–5. doi: 10.1093/nar/gkl887
22. Charoentong P, Finotello F, Angelova M, Mayer C, Efranova M, Rieder D, et al. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep.* (2017) 18:248–62. doi: 10.1016/j.celrep.2016.12.019
23. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell.* (2015) 160:48–61. doi: 10.1016/j.cell.2014.12.033
24. Roessler S, Jia HL, Budhu A, Forgues M, Ye QH, Lee JS, et al. A unique metastasis gene signature enables prediction of tumor relapse in early-stage hepatocellular carcinoma patients. *Cancer Res.* (2010) 70:10202–12. doi: 10.1158/0008-5472.CAN-10-2607
25. Roessler S, Long EL, Budhu A, Chen Y, Zhao X, Ji J, et al. Integrative genomic identification of genes on 8p associated with hepatocellular carcinoma progression and patient survival. *Gastroenterology.* (2012) 142:957–66.e12. doi: 10.1053/j.gastro.2011.12.039
26. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics.* (2010) 26:1572–3. doi: 10.1093/bioinformatics/btq170
27. Kim T, Chen IR, Lin Y, Wang AY, Yang JYH, Yang P. Impact of similarity metrics on single-cell RNA-seq data clustering. *Brief Bioinform.* (2018). doi: 10.1093/bib/bby076. [Epub ahead of print].
28. Kiselev VY, Andrews TS, Hemberg M. Challenges in unsupervised clustering of single-cell RNA-seq data. *Nat Rev Genet.* (2019) 20:273–82. doi: 10.1038/s41576-018-0088-9
29. Yang J, Hou Z, Wang C, Wang H, Zhang H. Gene expression profiles reveal key genes for early diagnosis and treatment of adamantinomatous craniopharyngioma. *Cancer Gene Ther.* (2018) 25:227–39. doi: 10.1038/s41417-018-0015-4
30. Huang H, Liu Y, Yuan M, Marron JS. Statistical significance of clustering using soft thresholding. *J Comput Graph Stat.* (2015) 24:975–93. doi: 10.1080/10618600.2014.948179
31. Chen F, Zhang Y, Parra E, Rodriguez J, Behrens C, Akbani R, et al. Multiplatform-based molecular subtypes of non-small-cell lung cancer. *Oncogene.* (2017) 36:1384–93. doi: 10.1038/ncr.2016.303
32. Raychaudhuri S, Stuart JM, Altman RB. Principal components analysis to summarize microarray experiments: application to sporulation time series. *Pac Symp Biocomput.* (2000) 2000:455–66. doi: 10.1142/9789814447331_0043
33. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics.* (2008) 9:559. doi: 10.1186/1471-2105-9-559
34. Kohl M, Wiese S, Warscheid B. Cytoscape: software for visualization and analysis of biological networks. *Methods Mol Biol.* (2011) 696:291–303. doi: 10.1007/978-1-60761-987-1_18
35. Cancer Genome Atlas Research Network. Electronic address: wheeler@bcm.edu, Cancer Genome Atlas Research Network. Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma. *Cell.* (2017) 169:1327–41.e23. doi: 10.1016/j.cell.2017.05.046
36. Fu J, Zhang Z, Zhou L, Qi Z, Xing S, Lv J, et al. Impairment of CD4+ cytotoxic T cells predicts poor survival and high recurrence rates in patients with hepatocellular carcinoma. *Hepatology.* (2013) 58:139–49. doi: 10.1002/hep.26054
37. Kan Z, Zheng H, Liu X, Li S, Barber TD, Gong Z, et al. Whole-genome sequencing identifies recurrent mutations in hepatocellular carcinoma. *Genome Res.* (2013) 23:1422–33. doi: 10.1101/gr.154492.113
38. Cleary SP, Jeck WR, Zhao X, Chen K, Selitsky SR, Savich GL, et al. Identification of driver genes in hepatocellular carcinoma by exome sequencing. *Hepatology.* (2013) 58:1693–702. doi: 10.1002/hep.26540
39. Guichard C, Amaddeo G, Imbeaud S, Ladeiro Y, Pelletier L, Maad IB, et al. Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. *Nat Genet.* (2012) 44:694–8. doi: 10.1038/ng.2256
40. Ahn SM, Jang SJ, Shim JH, Kim D, Hong SM, Sung CO, et al. Genomic portrait of resectable hepatocellular carcinomas: implications of RB1 and FGF19 aberrations for patient stratification. *Hepatology.* (2014) 60:1972–82. doi: 10.1002/hep.27198
41. Jhunjunwala S, Jiang Z, Stawiski EW, Gnad F, Liu J, Mayba O, et al. Diverse modes of genomic alteration in hepatocellular carcinoma. *Genome Biol.* (2014) 15:436. doi: 10.1186/PREACCEPT-1790146428112469
42. Tahmasebi Birgani M, Carloni V. Tumor microenvironment, a paradigm in hepatocellular carcinoma progression and therapy. *Int J Mol Sci.* (2017) 18:E405. doi: 10.3390/ijms18020405
43. Lee JS, Heo J, Libbrecht L, Chu IS, Kaposi-Novak P, Calvisi DF, et al. A novel prognostic subtype of human hepatocellular carcinoma derived from hepatic progenitor cells. *Nat Med.* (2006) 12:410–6. doi: 10.1038/nm1377
44. Yamashita T, Forgues M, Wang W, Kim JW, Ye Q, Jia H, et al. EpCAM and alpha-fetoprotein expression defines novel prognostic subtypes of hepatocellular carcinoma. *Cancer Res.* (2008) 68:1451–61. doi: 10.1158/0008-5472.CAN-07-6013
45. Calderaro J, Couchy G, Imbeaud S, Amaddeo G, Letouze E, Blanc JF, et al. Histological subtypes of hepatocellular carcinoma are related to gene mutations and molecular tumour classification. *J Hepatol.* (2017) 67:727–38. doi: 10.1016/j.jhep.2017.05.014
46. Schreiber RD, Old LJ, Smyth MJ. Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion. *Science.* (2011) 331:1565–70. doi: 10.1126/science.1203486
47. Shiraiishi Y, Fujimoto A, Furuta M, Tanaka H, Chiba K, Boroevich KA, et al. Integrated analysis of whole genome and transcriptome sequencing reveals diverse transcriptomic aberrations driven by somatic genomic changes in liver cancers. *PLoS ONE.* (2014) 9:e114263. doi: 10.1371/journal.pone.0114263
48. Hato T, Goyal L, Greten TF, Duda DG, Zhu AX. Immune checkpoint blockade in hepatocellular carcinoma: current progress and future directions. *Hepatology.* (2014) 60:1776–82. doi: 10.1002/hep.27246
49. Pan K, Wang H, Chen MS, Zhang HK, Weng DS, Zhou J, et al. Expression and prognosis role of indoleamine 2,3-dioxygenase in hepatocellular carcinoma. *J Cancer Res Clin Oncol.* (2008) 134:1247–53. doi: 10.1007/s00432-008-0395-1
50. Zhou SL, Dai Z, Zhou ZJ, Wang XY, Yang GH, Wang Z, et al. Overexpression of CXCL5 mediates neutrophil infiltration and indicates poor prognosis for hepatocellular carcinoma. *Hepatology.* (2012) 56:2242–54. doi: 10.1002/hep.25907
51. Gao Q, Zhao YJ, Wang XY, Qiu SJ, Shi YH, Sun J, et al. CXCR6 upregulation contributes to a proinflammatory tumor microenvironment that drives metastasis and poor patient outcomes in hepatocellular carcinoma. *Cancer Res.* (2012) 72:3546–56. doi: 10.1158/0008-5472.CAN-11-4032
52. Solinas G, Germano G, Mantovani A, Allavena P. Tumor-associated macrophages (TAM) as major players of the cancer-related inflammation. *J Leukoc Biol.* (2009) 86:1065–73. doi: 10.1189/jlb.0609385
53. Doherty DG, O'Farrelly C. Innate and adaptive lymphoid cells in the human liver. *Immunol Rev.* (2000) 174:5–20. doi: 10.1034/j.1600-0528.2002.017416.x
54. Hoshida Y, Villanueva A, Kobayashi M, Peix J, Chiang DY, Camargo A, et al. Gene expression in fixed tissues and outcome in hepatocellular carcinoma. *N Engl J Med.* (2008) 359:1995–2004. doi: 10.1056/NEJMoa0804525

Conflict of Interest: 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.

Copyright © 2019 Li, Wang, Ma, Zhang, Ou-yang, Qi and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.