A Five-Gene Signature Associated With DNA Damage Repair Molecular Subtype Predict Overall Survival for Hepatocellular Carcinoma

Junyu Huo[1], Xinyi Fan[2], Bingxin Qi[3] and Peng Sun[4]*

[1]Liver Disease Center, The Affiliated Hospital of Qingdao University, Qingdao, China, [2]Department of Allergy, The Affiliated Hospital of Qingdao University, Qingdao, China, [3]School of Public Health, Qingdao University, Qingdao, China, [4]Department of Hepatobilary and Pancreatic Surgery, The Affiliated Hospital of Qingdao University, Qingdao, China

Background: DNA damage repair (DDR) is an important mechanism for the occurrence and development of hepatocellular carcinoma (HCC), but its impact on prognosis has not been fully understood.

Materials and methods: A total of 904 HCC patients were included in our study, TCGA ($n = 370$) and GSE14520 ($n = 239$) were merged into a large-sample training cohort ($n = 609$). The training cohort was clustered into C1 and C2 based on prognostic DDR-related genes, the differentially expressed genes (DEGs) between C1 and C2 were identified by the Wilcoxon signed-rank test referred to criteria ($|log2FC|\geq1$ and FDR< 0.05). The univariate Cox analysis was used to screen the prognostic-related DEGs, and Lasso penalized Cox regression analysis was used to construct the risk score. The patients were clarified into high- and low-risk groups based on the median risk score. ICGC ($n = 231$) and GSE116174 ($n = 64$) cohorts were used for external validation of the risk score's prognostic value.

Results: The Kaplan–Meier survival analysis showed that the high-risk group had a significantly reduced overall survival (OS) compared to the low-risk group in the three independent cohorts, and the time-dependent ROC curve showed that the five-gene (STMN1, PON1, PLOD2, MARCKSL1, and SPP1) risk score with a high accuracy in predicting OS. The patients with AFP >300 ng/ml, tumor poor differentiation (grade 3–4), micro and macro vascular tumor invasion, advanced stage (AJCC III-IV, BCLC stage B-C, and CLIP score >2) exhibited a higher risk score. Subgroup survival analysis found that the risk score was applicable to patients with different clinical characteristics. GO and KEGG functional enrichment analysis revealed that cell cycle, p53 signaling, TNF signaling-related pathways were upregulated in the high-risk group. The higher infiltration level of activated

---

Abbreviations: HCC, hepatocellular carcinoma; DDR, DNA damage repair; TCGA, the cancer genome atlas; ICGC, international cancer genome consortium; GEO, gene expression omnibus; PRGs, prognostic related genes; DEGs, differential expressed genes; GO, gene ontology; KEGG, kyoto encyclopedia of genes and genomes; KM, Kaplan-Meier; LASSO, least absolute shrinkage and selection operator; ROC, receiver operating characteristic; AUC, area under curve; OS, overall survival; NES, normalized enrichment score; GSEA, gene set enrichment analysis; ssGSEA, single sample gene set enrichment analysis; IPS, immunophenoscore; ICIs, immune checkpoint inhibitors.

CD4 T cell, CD56 bright natural killer cell, plasmacytoid dendritic cell, and type 2 T helper cells were found to lead an unfavorable impact on the OS of HCC patients, and these four kinds of immune cells exhibited a higher infiltration level in the high-risk group.

Conclusion: The five-gene risk score proposed in the research may provide new insights into the individualized evaluation of HCC prognosis.

Keywords: hepatocellular carcinoma, DNA damage repair, prognostic, signature, risk score

BACKGROUND

Somatic cells are often affected by a variety of *in vivo* and *in vitro* factors, such as ionizing radiation (IR), ultraviolet (UV) and other physical factors (Soutoglou and Misteli, 2008), alkylating agents, nitrosamines and other chemical factors (Gidron et al., 2006), virus infection and other biological factors (Lewis et al., 2013), these factors will cause a variety of DNA damage, such as single-strand and double-strand gaps (SSB, DSB), base mutation, base oxidative damage and so on (Jackson and Bartek, 2009). If the damage was repaired correctly and the cells survive, and if there were incorrect repairs such as deletions and insertions in the repair process, the accumulation of these abnormal bases may lead to cancer (Alhmoud et al., 2020). If the damage was serious and could not be repaired, the cell would initiate the apoptosis process and induce cell death, which was also considered to be the last barrier to prevent the malignant transformation of the cell (Roos and Kaina, 2006). Therefore, signal transduction, damage repair, and apoptosis induction form a complex network which was closely related and influence each other, which enables cells as a whole to respond to DNA damage.

Hepatocellular carcinoma (HCC) is a primary liver cancer derived from hepatocytes, accounting for 85~90% of all primary liver cancers (Ozakyol, 2017). It is the fifth most common cancer in men and the seventh most common cancer in women worldwide; it is estimated that there are about 782,000 new cases every year, causing 600,000 deaths every year (Ozakyol, 2017). Its high mortality and short survival time lead to a serious global health burden (Sayiner et al., 2019). The occurrence of HCC is a complex process of interaction between genetic and environmental factors (Aravalli et al., 2008), and its mechanism has not been fully elucidated. Important risk factors include environmental toxins such as chronic viral hepatitis, liver cirrhosis, and aflatoxin, lifestyle factors such as non-alcoholic fatty liver disease (NAFLD), drinking, smoking, and diet (Gomaa et al., 2008). Various HCC related risk factors could cause DNA damage (Schütte et al., 2009). If the damaged DNA was not repaired in time and correctly, it could lead to gene mutation and genomic instability, which is gradually considered to be the common feature of human HCC (Farazi and DePinho, 2006). The disorder of the DNA damage repair process was related to the susceptibility to liver cancer, and this process was often enhanced in HCC cells, resulting in the unsatisfactory effect of anticancer treatment against HCC cells (Yang et al., 2014).

DNA damage repair (DDR) is not only an important mechanism for the occurrence and development of HCC but also an important reason for the poor effect of chemotherapy and other treatments. However, the carcinogenic mechanism of DDR in HCC remains to be investigated, and the improvement and innovation of individualized evaluation of HCC prognosis have a broad prospect, which is worthy of further exploration.

MATERIALS AND METHODS

Data Acquisition

Four independent HCC cohorts with prognosis information were included in our research: TCGA-LIHC, $n = 370$; GSE14520, $n = 239$; ICGC-LIRI-JP, $n = 231$; GSE116174, $n = 64$. Their gene expression and clinical data were obtained from three public databases: The Cancer Genome Atlas (TCGA, https://portal.gdc.cancer.gov/), the International Cancer Genomics Consortium (ICGC, https://icgc.org/), and Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/). The clinicopathological information for all cohorts is shown in **Table 1**. The TCGA and GSE14520 were merged into a large-sample training cohort ($n = 609$). The usage rules of the TCGA, ICGC, and GEO database were fully complied with during data collection. The gene expression format in the three RNA-seq cohorts were normalized to transcripts per million kilobase (TPM) values based on R package "limma", and the ComBat function of the R "SVA" package was used to eliminate the batch effect in different datasets (Huo et al., 2021e; Huo et al., 2021f). As the data utilized in our research were acquired from public databases, approval from the local ethics committee was not needed.

DNA Damage Repair Related Genes Cluster Analysis

We extracted 223 DNA damage repair (DDR) pathways (including homologous recombination (HR), mismatch repair (MMR), base excision repair (BER), nucleotide excision repair (NER), and nonho-mologous end-joining (NHEJ) related genes from the molecular signatures database (MSigDB, http://www.gsea-msigdb.org/gsea). Based on the prognostic related genes (PRGs) with *p*-value < 0.05 screened out by the univariate Cox regression analysis, the training cohort were conducted cluster analysis by R-package "ConsensusClusterPlus". The Kaplan-Meier method with a two-sided log-rank test was employed to compare the overall survival (OS) difference between different clusters. The gene sets enrichment analysis (GSEA) was applied to assess the DDR pathway activities for different clusters. The differentially expressed genes (DEGs)

TABLE 1 | The clinicopathological information for all cohorts.

	TCGA	GSE14520	ICGC	GSE116174
Survival status				
alive	240	143	189	37
dead	130	96	42	27
Gender				
male	249	189	170	6
female	121	28	61	58
Age				
≤65	232	198	89	55
>65	138	19	142	9
HBV				
Positive				47
Negative				17
Alcohol				
Yes				13
None				51
Grade				
G1	55			
G2	177			
G3	121			
G4	12			
AJCC TNM stage				
I&II	256	168	141	53
III&IV	90	49	90	11
priorMalignancy				
None			201	
Yes			30	
AFP				
≤300 ng/ml	197	120		
>300 ng/ml	62	97		
vascular_tumor_cell_type				
None	206			35
Micro and macro	108			29
new_tumor_event_after_initial_treatment				
None	162			
Yes	168			
BCLC stage				
0-A		165		
B-C		52		
CLIP_Score				
<2		169		
≥2		48		
ALT				
≤50U/L		127		
>50U/L		90		

between different clusters were identified by R package "limma" referred to criteria (|log2FC|≥1 and FDR <0.05) (Huo et al., 2021d).

Development and Validation of a Risks Score Predicting OS of HCC

The DEGs with p-value < 0.001 obtained from the univariate Cox regression analysis were considered to be the prognostic related genes (PRGs) in the training cohort ($n = 609$). Next, least absolute shrinkage and selection operator (LASSO) regression with 10-fold cross-validation was performed, and 1,000 cycles were run via the R software package "glmnet". For each cycle, random stimulation was set to 1,000 times, and the penalty parameter (λ) was decided by the minimum partial likelihood deviance (Huo et al., 2021c). The genes with nonzero regression

coefficients obtained from lasso regression analysis were included in the multivariate Cox regression analysis (Huo et al., 2021h). The risk score was established by the expression level of each gene multiple its corresponding regression coefficients derived from multivariate Cox regression analysis of each gene (Huo et al., 2021g). After each patient in the training cohort ($n = 609$) got their own risk score, we arranged them in a sequence from low to high, and took the median value to divide them into a high-risk group and low-risk group (Huo et al., 2021a; Huo et al., 2021b), the Kaplan–Meier survival analysis was implemented to compare the OS of the two groups. The time-dependent receiver operating characteristic (ROC) analysis was used to evaluate the accuracy of the risk score in predicting OS of HCC. The univariate and multivariate Cox regression analysis were used to assess the independent prognostic value of the risk score. Internal validation was conducted in TCGA and GSE14520 cohorts, external validation was conducted in ICGC and GSE116174 cohorts, clinical subgrouping validation was used to test the risk score's universal applicability.

Exploration of the Molecular Mechanism Underlying the Prognostic Signature

The DEGs between the high- and low-risk groups were identified by the "limma" R package (fdr <0.05), the R package "clusterProfiler" was employed for the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) functional enrichment analysis of the DEGs.
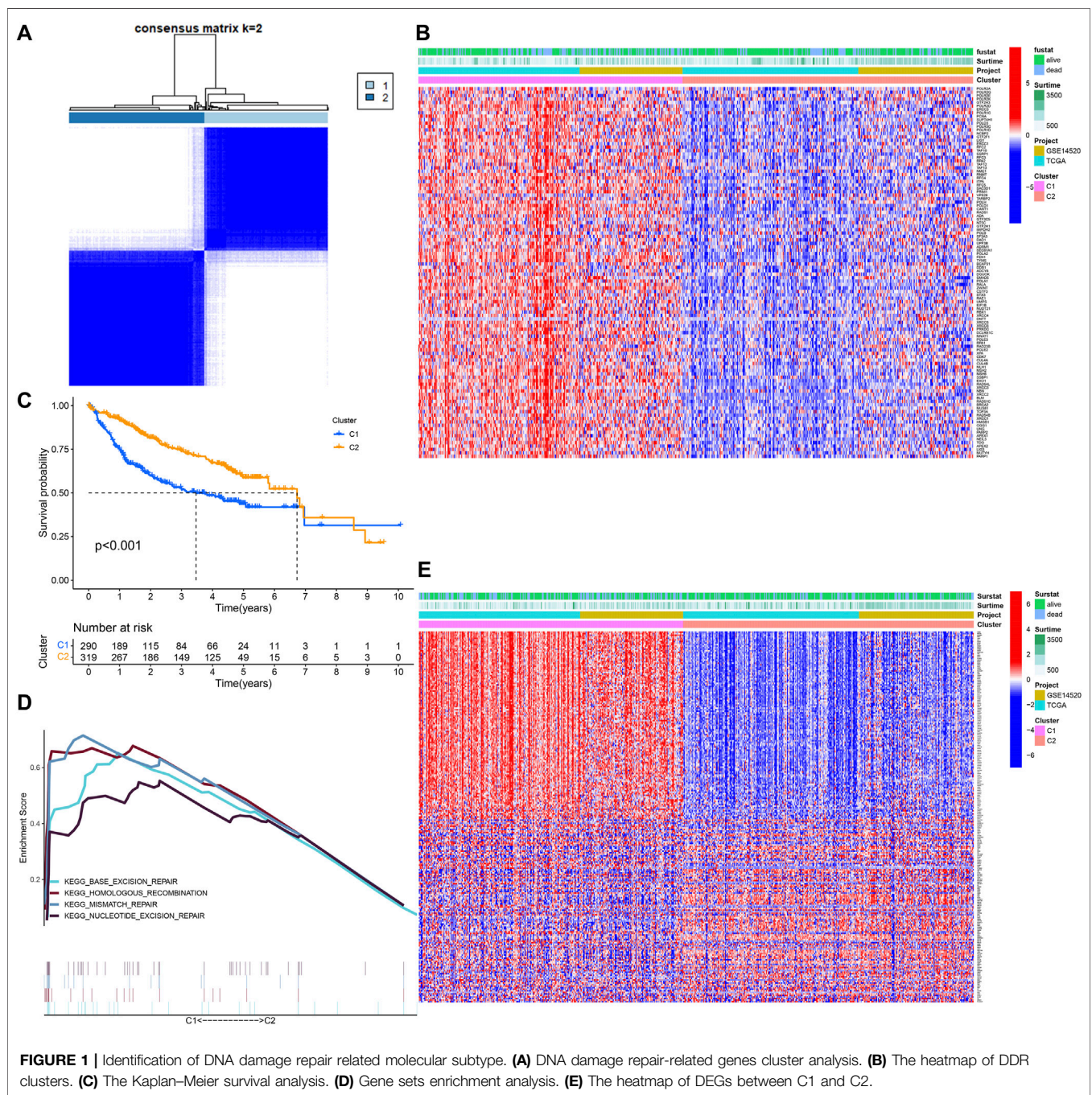
Quantification of 23 Types of Immune Cells Infiltration Using ssGSEA Algorithm

The single sample gene set enrichment analysis (ssGSEA) algorithm was employed to calculate the normalized enrichment score (NES) for the quantification of 23 types of immune cells infiltration abundance (Barbie et al., 2009). The immune cells infiltration differences between the high- and low-risk groups were compared by the independent-sample t-tests, $p < 0.05$ was considered to be of statistical significance.

RESULTS

Identification of DEGs Between Different DDR Clusters

We conducted the univariate Cox regression analysis on the DDR-related genes and found that 108 genes were associated with the OS of HCC (**Supplementary Material**-1). The training cohort was clustered into C1 and C2 based on more than 108 genes (**Figures 1A,B**). The OS of C1 was obviously lower than that of C2 (**Figure 1C**), and the GSEA showed that the C1 was an active DDR activity enhanced subtype (**Figure 1D**). A total of 239 DEGs were identified between C1 and C2, 166 were upregulated in C1 and 74 were upregulated in C2 (**Figure 1E**). The Chi-square test showed that there were significant differences between the two subtypes in

FIGURE 1 | Identification of DNA damage repair related molecular subtype. **(A)** DNA damage repair-related genes cluster analysis. **(B)** The heatmap of DDR clusters. **(C)** The Kaplan–Meier survival analysis. **(D)** Gene sets enrichment analysis. **(E)** The heatmap of DEGs between C1 and C2.

survival status ($p = 0.002$), but there was no difference between TCGA and GSE14520 ($p = 0.959$) (**Supplementary Figure S1**).

A Five-Gene Risk Score Constructed in the Training Cohort

In all, 193 of 239 DEGs were considered to have a significant impact on the prognosis of HCC via univariate Cox regression analysis ($p < 0.001$) (**Supplementary Material**-2). Through the penalty parameter (λ) was decided by the minimum partial

likelihood deviance, 11 genes with nonzero lasso regression coefficient were retained (**Figure 2A**), a five-gene risk score were formatted by gene expression level and corresponding multivariate Cox regression coefficients: $STMN1 * 0.008619 - PON1 * 0.002848 + PLOD2 * 0.03729 + MARCKSL1 * 0.006922 + SPP1 * 0.001817$ (**Figure 2B**). The patients were divided into high- and low-risk groups referred to the median risk score (0.986). The OS of the high-risk patients was significantly reduced compared to the low-risk patients ($p < 0.001$, **Figure 2C**). The area under curve (AUC) values of 1-year, 3-
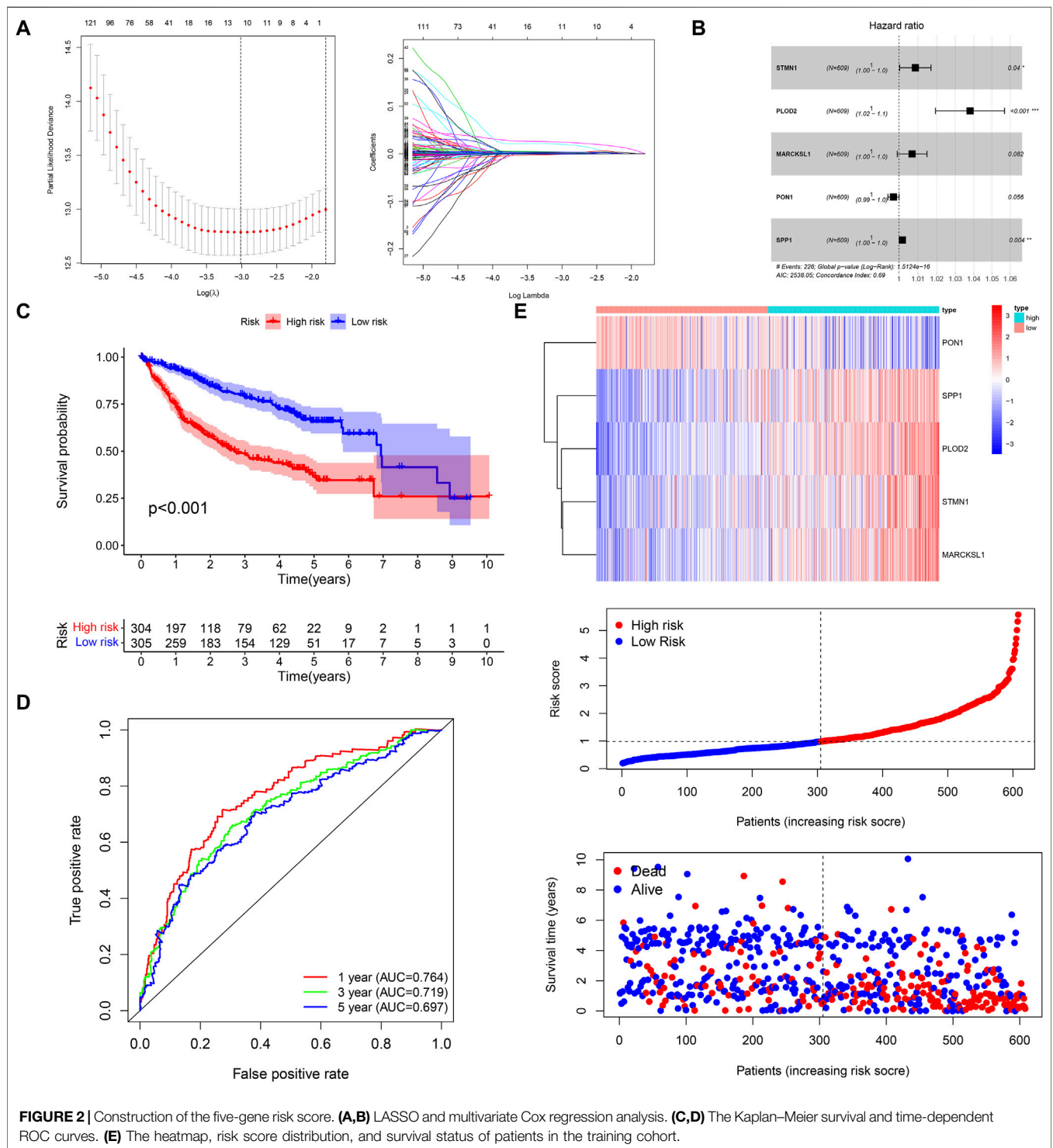
FIGURE 2 | Construction of the five-gene risk score. (A,B) LASSO and multivariate Cox regression analysis. (C,D) The Kaplan–Meier survival and time-dependent ROC curves. (E) The heatmap, risk score distribution, and survival status of patients in the training cohort.

year, and 5-year OS predicted by risk score were 0.764, 0.719, and 0.697 respectively (**Figure 2D**). The patients with lower risk scores were found to have a longer survival time and higher survival rate (**Figure 2E**). By comparing AUC, the 1, 3, and

5 years OS prediction efficiency of the risk score was better than that of single gene and clinical factors (**Supplementary Figure S2**). We also detected the expression of signature genes in human normal tissues and HCC tissues with the help of HPA database
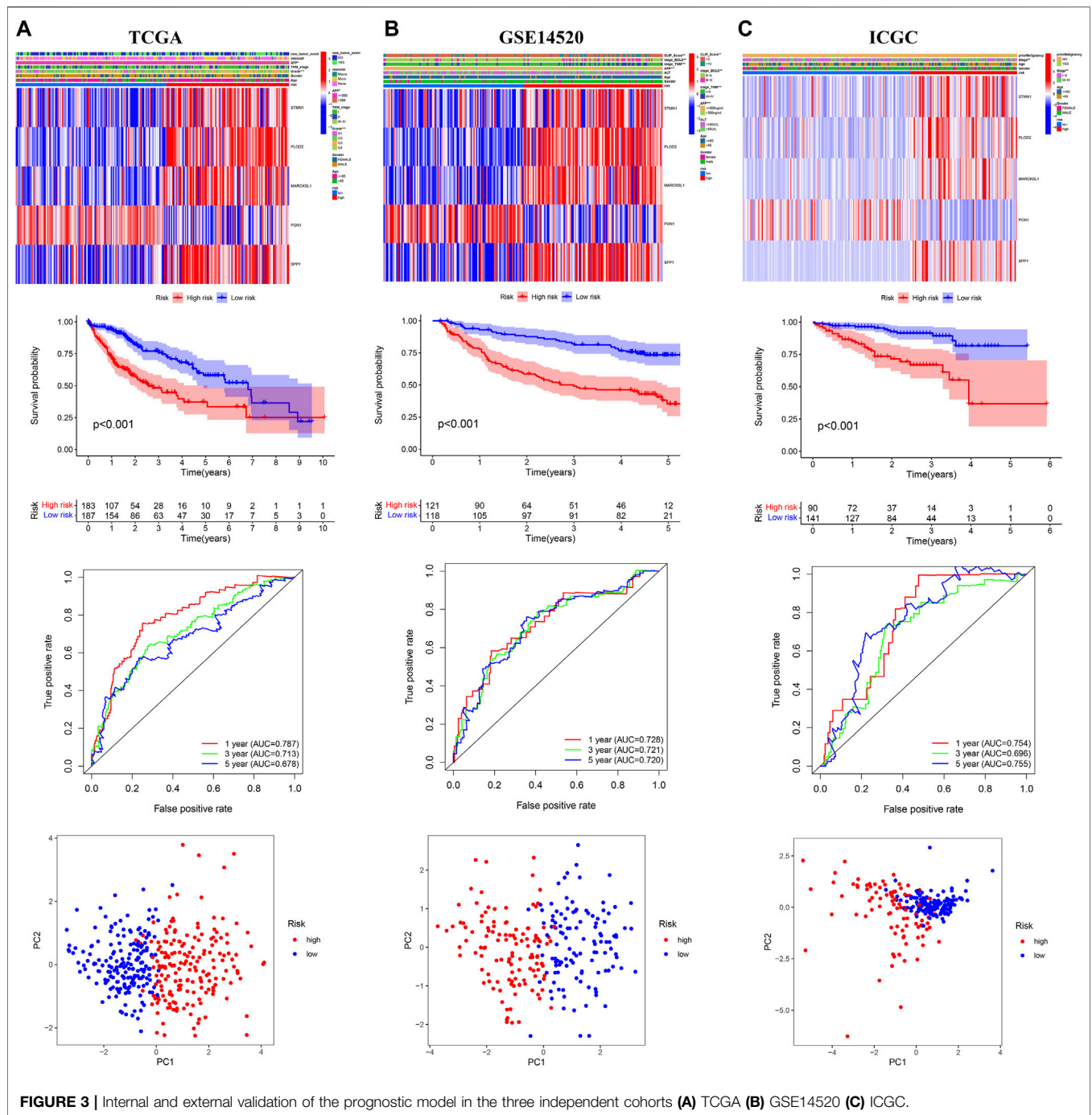
FIGURE 3 | Internal and external validation of the prognostic model in the three independent cohorts **(A)** TCGA **(B)** GSE14520 **(C)** ICGC.

(Human Protein Atlas, https://www.proteinatlas.org/) (**Supplementary Figure S3**).

Internal and External Validation of the Prognostic Signature

The OS of high-risk patients were significantly reduced relative to the low-risk patients in internal (TCGA, $n = 370$; GSE14520, $n = 239$) and external (ICGC, $n = 231$; GSE116174, $n = 64$) validation cohorts

(**Figures 3A–C; Supplementary Figure S4A**). In the TCGA cohort, the AUC values of 1-year, 3-year, and 5-year OS predicted by risk score were 0.787, 0.713, and 0.678 respectively (**Figure 3A**); In the GSE14520 cohort, the AUC values of 1-year, 3-year, and 5-year OS predicted by risk score were 0.728, 0.721, and 0.720 respectively (**Figure 3B**); In the ICGC cohort, the AUC values of 1-year, 3-year and 5-year OS predicted by risk score were 0.754, 0.696, and 0.755 respectively (**Figure 3C**). In the GSE116174 cohort, the AUC values of 1-year, 3-year, and 5-year OS predicted by risk score were 0.751, 0.610,

FIGURE 4 | Clinical correlation analysis in the three independent cohorts.
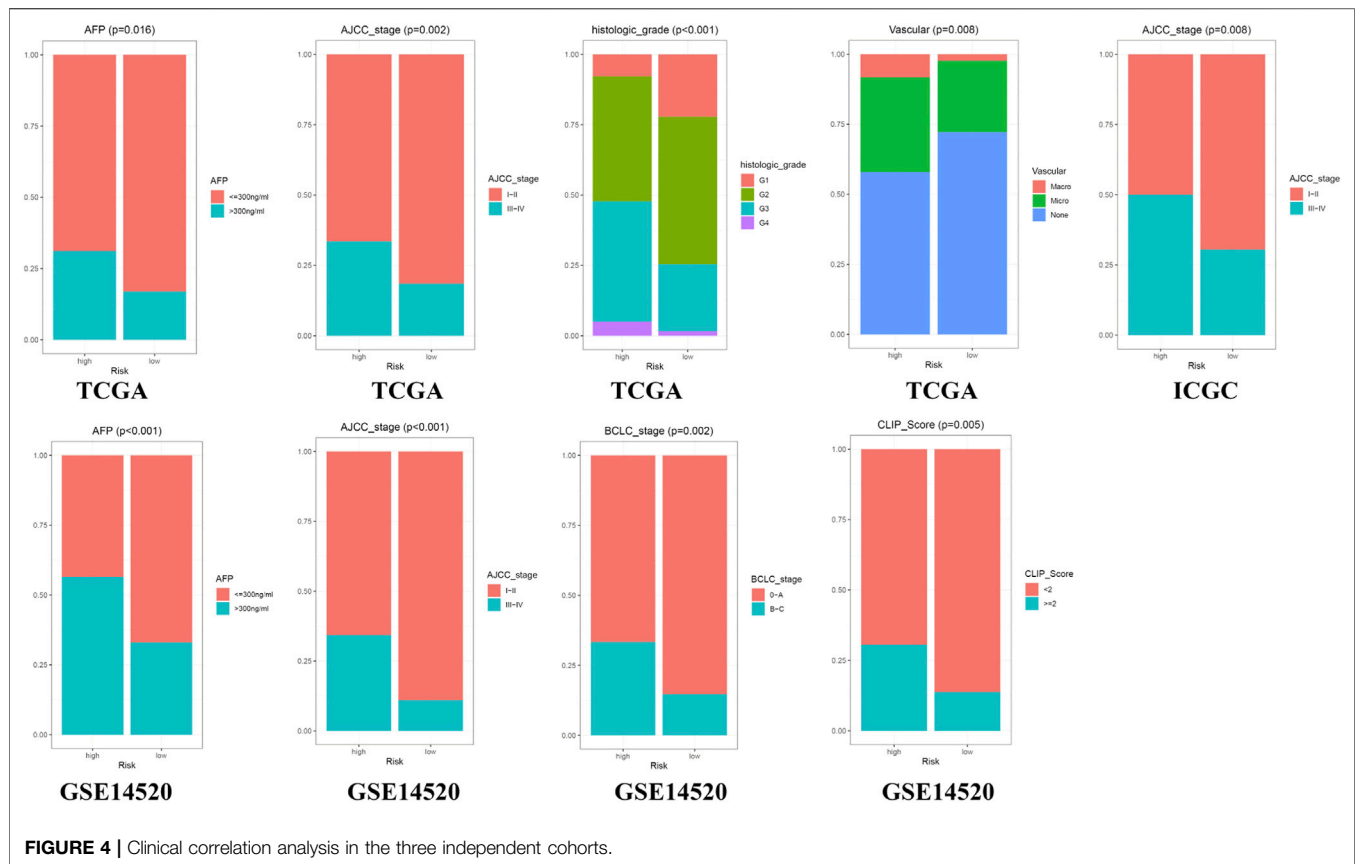
and 0.723 respectively (**Supplementary Figure S4B**). The principal component analysis (PCA) showed satisfactory separation between the high- and low-risk groups (**Figure 3C**). Among the five genes, only *PON1* was upregulated in the low-risk group, and the other four genes were upregulated in the high-risk group (**Figures 3A–C**; **Supplementary Figure S4C**).

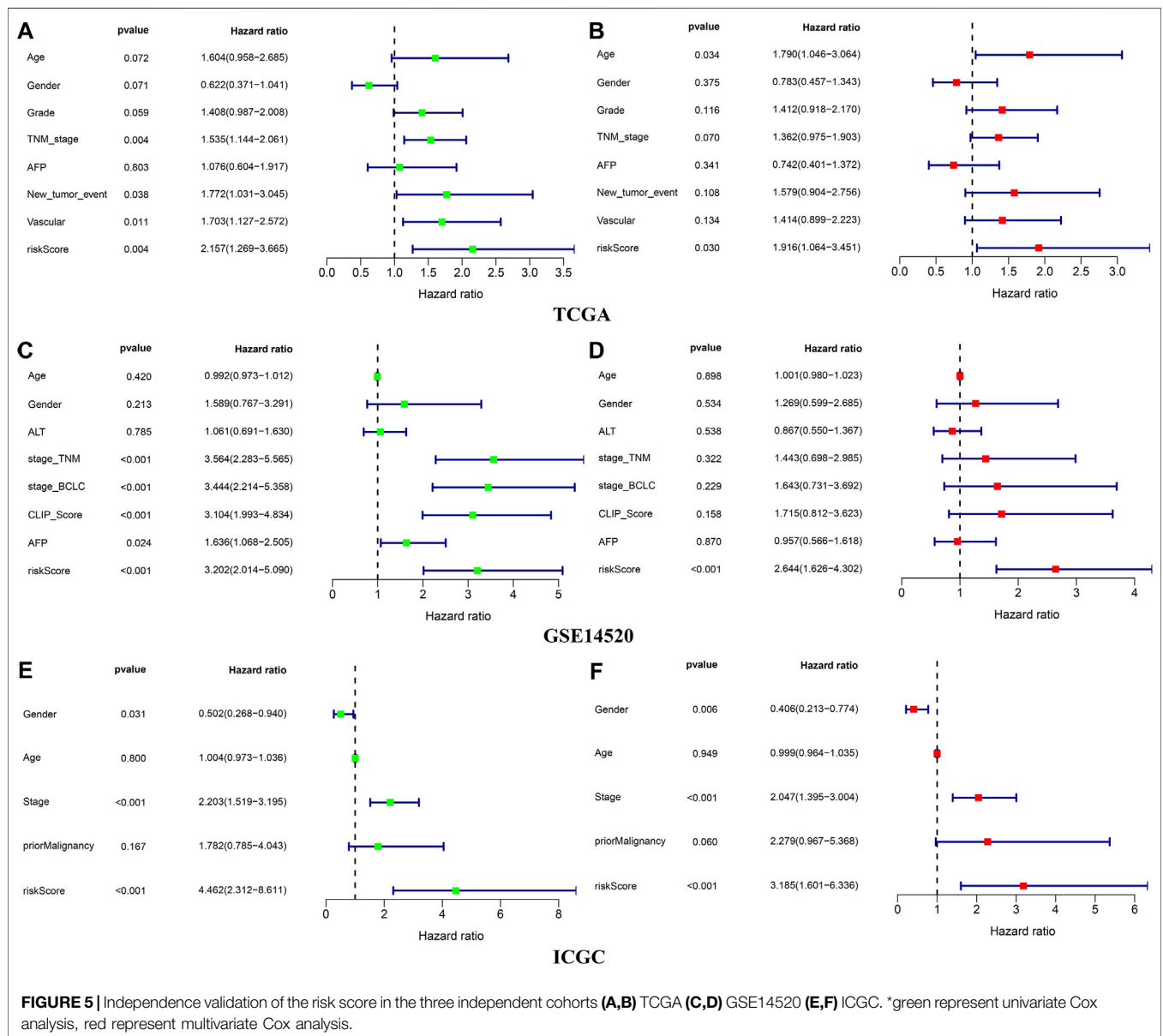## Clinical Correlation Analysis and Subgroup Survival Analysis

The Chi-square test showed that the patients with AFP >300 ng/ml, tumor poor differentiation (grade 3–4), micro and macro vascular tumor invasion, advanced stage (AJCC III-IV, BCLC stage B-C, and CLIP score>2) exhibited a higher risk score (**Figure 4**). The independent-samples t-tests suggested that the risk score distributed significantly differently in patients with different AFP levels, histology grades, vascular tumor cell type, AJCC stage, BCLC stage, and CLIP score (**Supplementary Figure S5**). The univariate and multivariate Cox regression analysis demonstrated that the risk score was an independent risk factor for the OS of patients in the three independent cohorts (**Figures 5A–C**; **Supplementary Figures S4D,E**). In 22 subgroups assigned by clinical characteristics, the OS of high-risk patients were significantly reduced compared to the low-risk patients in each subgroup ($p < 0.05$, **Figures 6A–K**).

Uncovering the Potential Molecular Mechanism of the Prognostic Signature

We identified the DEGs between high- and low-risk groups (**Figure 7A**), GO annotation found that these DEGs were involved in nuclear division, mitotic nuclear division, and organelle fission, etc. (**Figure 7B**). KEGG enrichment analysis showed that cellular senescence, cell cycle, p53 signaling, TNF signaling-related pathways were upregulated in the high-risk group, and it is worth mentioning that the genes with a positive risk coefficient were also involved in the pathways which were significantly upregulated in the high-risk group. While chemical carcinogenesis and bile secretion related pathways were upregulated in the low-risk group (**Figure 7C**).

Comparison of the Immune Infiltration in Different Risk Groups

We performed quantification of 23 types of immune cells infiltration abundance via the ssGSEA algorithm (**Figure 8A**). Based on the median NES, the patients were divided high- and low- infiltration groups. The Kaplan-Meier survival curves showed that the patients with higher infiltration levels of activated CD4 T cell, CD56 bright natural killer cell, plasmacytoid dendritic cell, and type 2 T helper cells had poor prognosis ($p < 0.05$, **Figure 8B**), and these four kinds of

FIGURE 5 | Independence validation of the risk score in the three independent cohorts **(A,B)** TCGA **(C,D)** GSE14520 **(E,F)** ICGC. *green represent univariate Cox analysis, red represent multivariate Cox analysis.

immune cells exhibited a higher infiltration level in the high-risk group (**Figure 8C**). The patients with higher infiltration levels of eosinophil and type 1 T helper cells showed better prognosis (*p* < 0.05, **Figure 8B**), and these two kinds of immune cells exhibited a higher infiltration level in the low-risk group (**Figure 8C**).

Investigation Treatment Strategy for Different Risk Groups

First, we compared the expression level of immune checkpoints between the high-and low-risk groups, the expression level of *TIGIT, CTLA4, PDCD1, LAG3,* and *CD274(PDL1),* etc., were all

significantly upregulated in the high-risk group (**Figure 9A**). Next, we obtained the immunophenoscore (IPS) of patients in the TCGA cohort from The Cancer Immunome Atlas (https://tcia.at/home) (Charoentong et al., 2017). In general, the higher the IPS, the more sensitivity to immune checkpoint inhibitors (ICIs); the IPS of CTLA4 (+) and PD1 (−) in the low-risk group was higher than that of the high-risk group with statistical significance (**Figure 9B**), indicating that the low-risk group was more likely to benefit from immunotherapy with anti-CTLA4 ICIs. In addition, we also compared the difference of anticancer drug sensitivity between the high- and low-risk groups by calculating the IC50 of anticancer drugs with the R package pRRophetic. The IC50 of sorafenib, rapamycin, lapatinib, and gefitinib, etc. in the low-risk
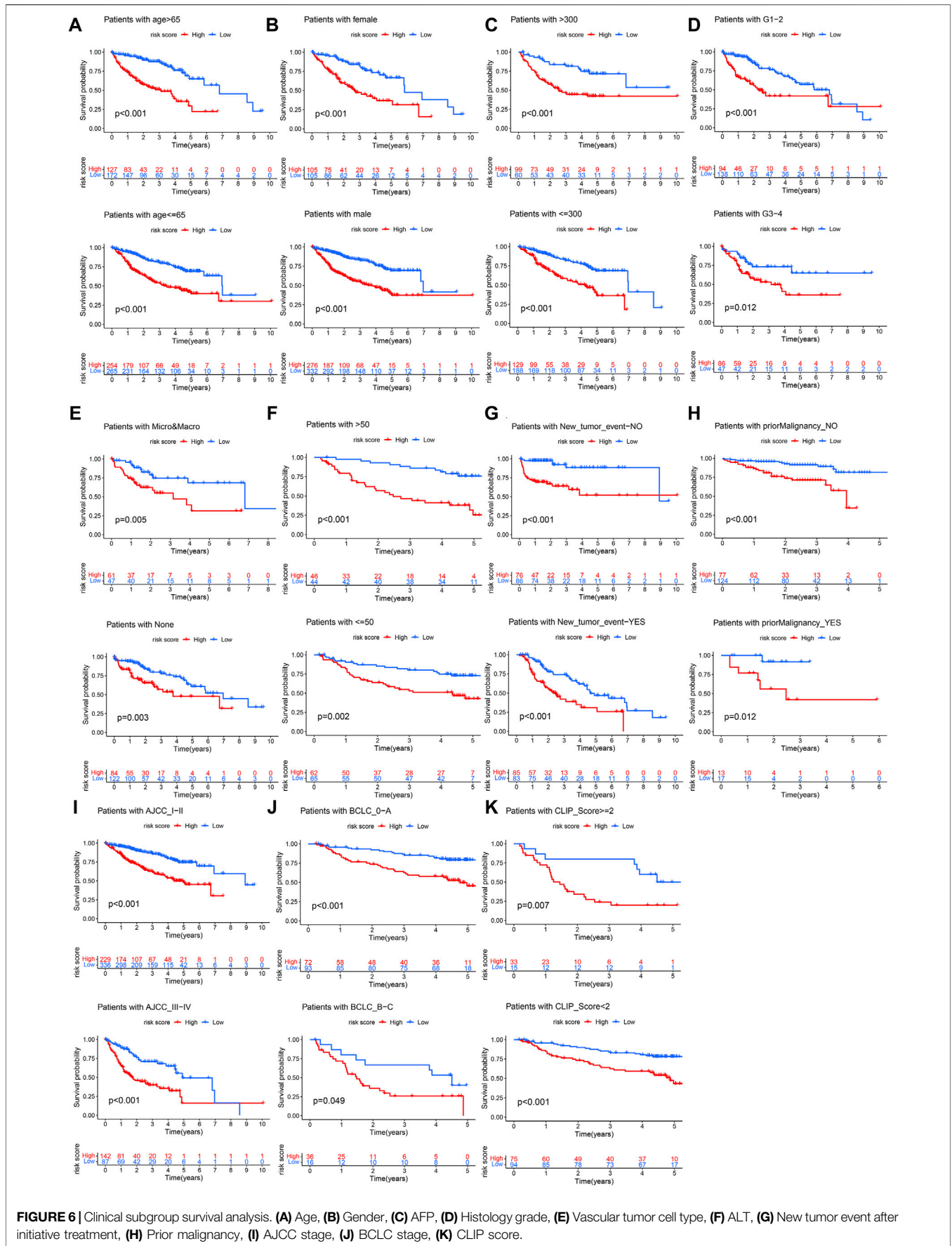
FIGURE 6 | Clinical subgroup survival analysis. **(A)** Age, **(B)** Gender, **(C)** AFP, **(D)** Histology grade, **(E)** Vascular tumor cell type, **(F)** ALT, **(G)** New tumor event after initiative treatment, **(H)** Prior malignancy, **(I)** AJCC stage, **(J)** BCLC stage, **(K)** CLIP score.
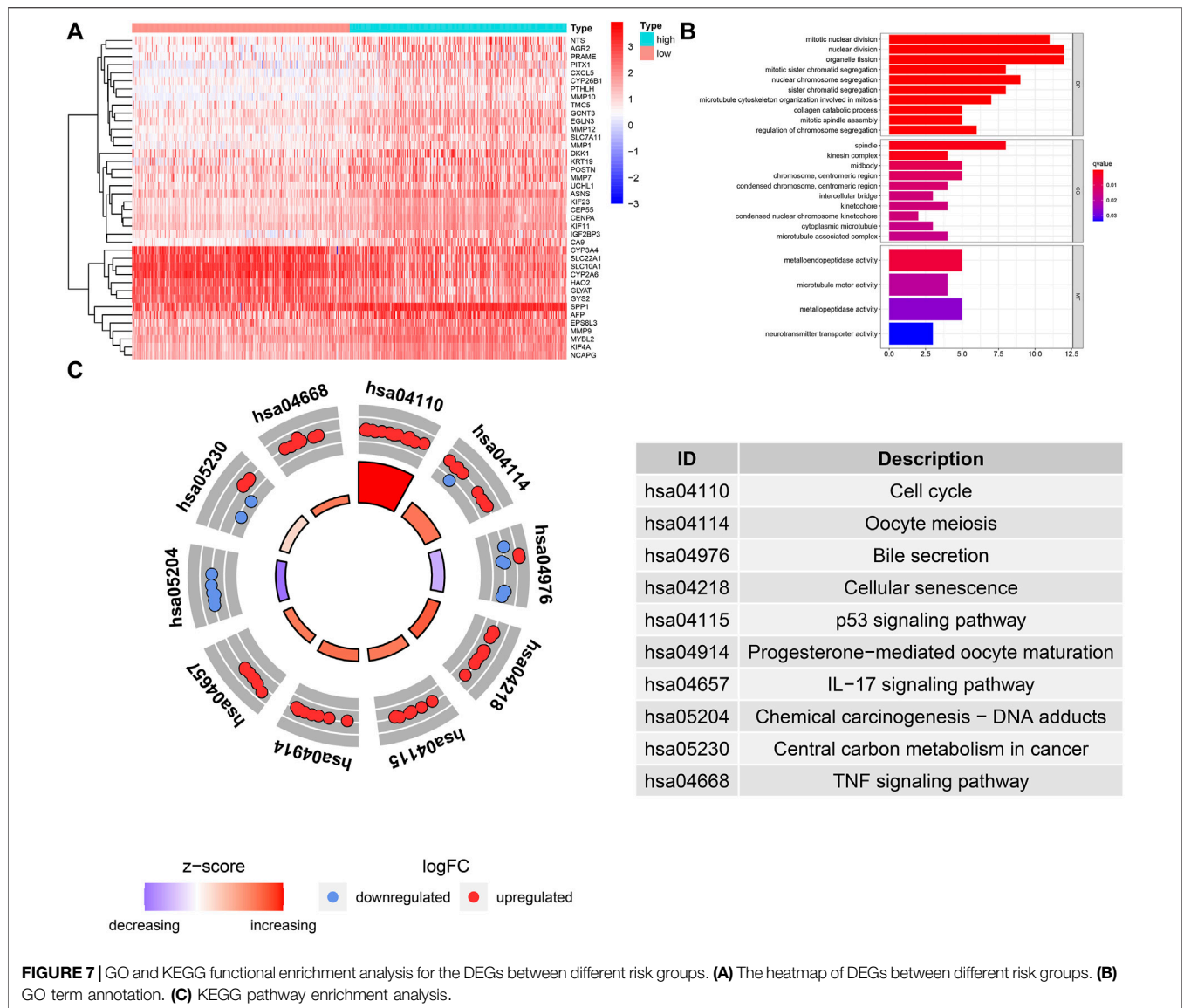
FIGURE 7 | GO and KEGG functional enrichment analysis for the DEGs between different risk groups. **(A)** The heatmap of DEGs between different risk groups. **(B)** GO term annotation. **(C)** KEGG pathway enrichment analysis.
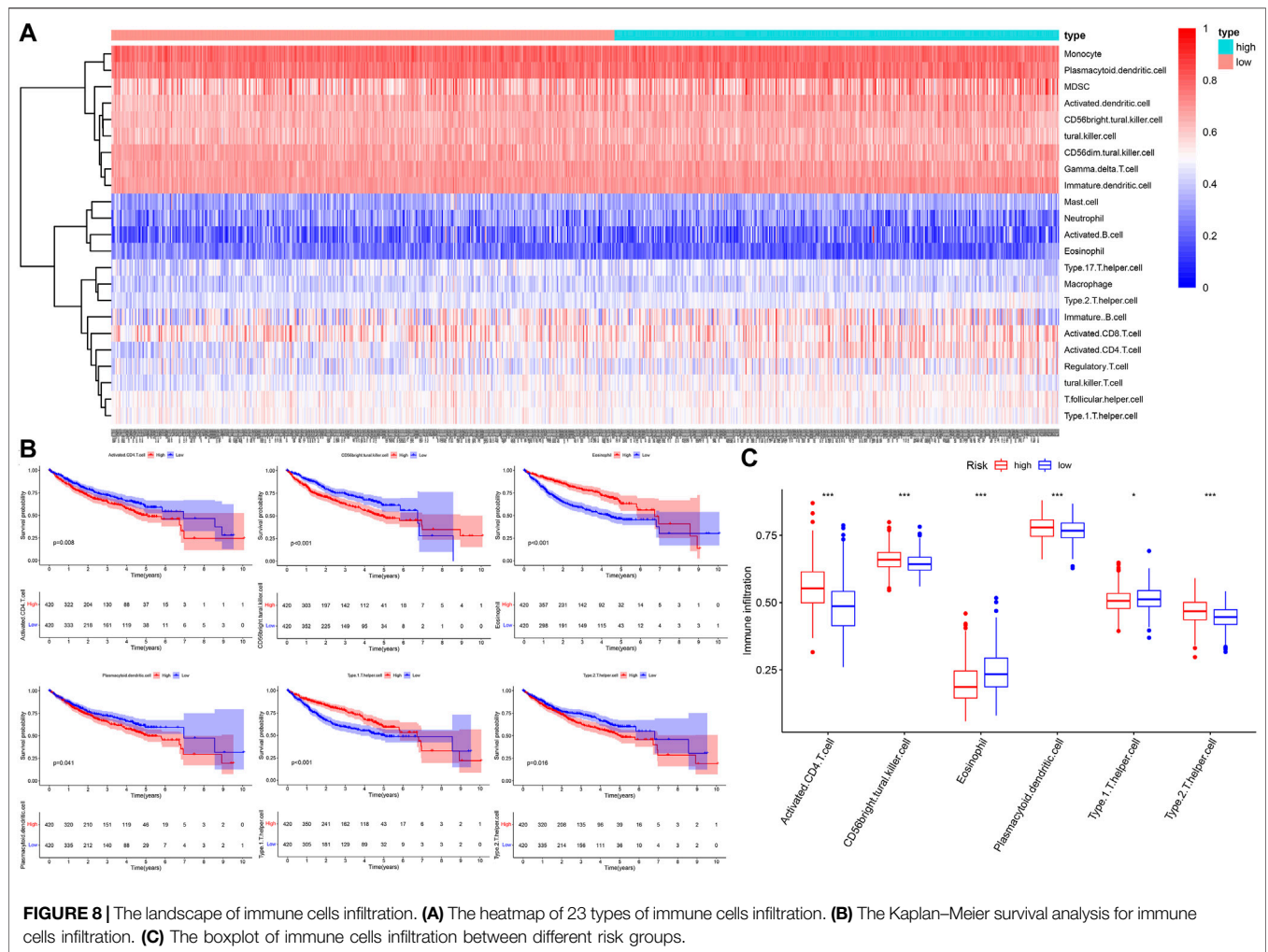
group was significantly lower than that of the high-risk group, but the IC50 of imatinib in the high-risk group was lower than that of the low-risk group with statistical significance (**Figure 9C**).

DISCUSSION

Hepatocellular carcinoma (HCC) is a common malignant tumor, and its etiology and pathogenesis have not been fully elucidated (Yang et al., 2019). It is closely related to hepatitis virus infection, aflatoxin, and bile acid to varying degrees (Fattovich et al., 2004), these factors could cause DNA damage in hepatocytes and then trigger a series of cellular reactions, mainly including damage signal transduction and apoptosis induced by DNA repair (Tanaka et al., 2008). If DNA damage could not be repaired correctly and accumulated continuously, it could lead to malignant transformation of hepatocytes and eventually lead

to HCC (Gillman et al., 2021). Therefore, DNA damage and repair (DDR) is an important molecular mechanism for the occurrence and development of HCC, and further study of it will lay a foundation for the comprehensive treatment of HCC.
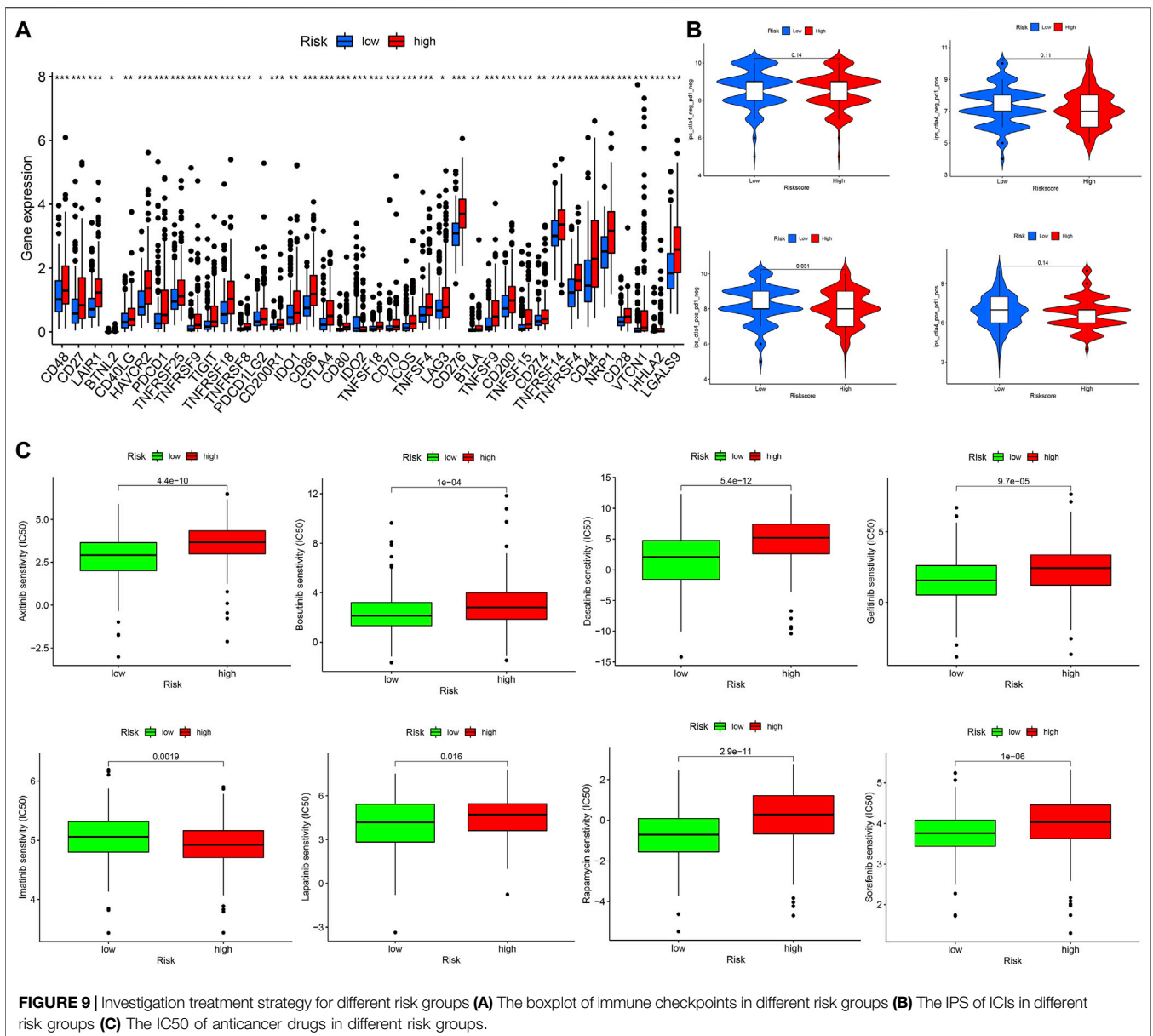
In our research, we found that the HCC patients with different DDR molecular subtypes had different clinical outcomes. C1 was defined as DDR active subtype, and its prognosis was significantly worse than C2, which confirmed that the activity of DDR was indeed related to the progression of HCC. Considering the complex correlation between DDR and genomic instability (Kerzendorfer and O'Driscoll, 2009), we identified the DEGs between C1 and C2 further. The results showed that 193 of 239 DEGs were significantly correlated with the OS of HCC ($p <$ 0.001), therefore, the difference in the expression of these DEGs may be the potential cause of the difference in prognosis between C1 and C2. After lasso and multivariate Cox regression analysis, a risk score consisting of five genes was established in the training

FIGURE 8 | The landscape of immune cells infiltration. **(A)** The heatmap of 23 types of immune cells infiltration. **(B)** The Kaplan–Meier survival analysis for immune cells infiltration. **(C)** The boxplot of immune cells infiltration between different risk groups.

cohort. The training cohort was a cohort with a large sample size ($n = 609$) merged by TCGA and GSE14520 datasets. In order to test the reliability of the risk score, we carried out internal and external validation. The AUC values for the risk score to predict 1, 3, and 5 year OS were all greater than 0.6 in the four independent cohorts, indicating that the risk score had high accuracy in the OS prediction of HCC. The risk score could be regarded as an independent prognosis indicator as shown by univariate and multivariate Cox regression analysis. The patients were divided into 22 subgroups according to clinical characteristics, the high-risk patients' OS decreased obviously in each subgroup, which proved that the risk score was applicable to patients with different clinical characteristics.

To clarify the reasons for the difference in OS between high- and low-risk groups, we explored it from three aspects: clinicopathological features, molecular mechanism, and immune infiltration. First, from the perspective of clinical relevance, we found the presentation of risk score was higher in patients with AFP >300 ng/ml, tumor poor differentiation (grade 3–4), vascular micro and macro invasion, advanced stage (AJCC III-IV, BCLC stage B-C, and CLIP score >2), manifested

that the higher risk score represented rapid tumor progression and stronger invasiveness. Secondly, from the perspective of molecular mechanism, we found that the cell cycle, p53 signaling, TNF signaling related pathways were positively enriched in the high-risk group, while chemical carcinogenesis, bile secretion related pathways were upregulated in the low-risk group, these results suggested that the risk score may affect tumor progression by regulating tumor cell proliferation and metabolism. Interestingly, the expression of genes involved in cellular senescence was up-regulated in the high-risk groups, indicating that the aging tumor microenvironment may have an adverse impact on the prognosis of HCC and played a critical role in the malignant progression of HCC. Thirdly, in terms of immune infiltration, we found that the infiltration level of six kinds of immune cells had a significant impact on the OS of HCC. The patients with higher infiltration abundance of activated CD4 T cell, CD56 bright natural killer cell, plasmacytoid dendritic cell (pDC), and type 2 T helper (Th2) cells had adverse prognosis, and these four kinds of immune cells exhibited a higher infiltration level in the high-risk group. The patients with higher infiltration levels of eosinophil and type 1 T helper (Th1) cells showed a

FIGURE 9 | Investigation treatment strategy for different risk groups **(A)** The boxplot of immune checkpoints in different risk groups **(B)** The IPS of ICIs in different risk groups **(C)** The IC50 of anticancer drugs in different risk groups.

favorable prognosis, and these two kinds of immune cells exhibited a lower infiltration level in the high-risk group. Previous reports have pointed out that activated CD4 T cells have the ability to inhibit tumor, it could not only directly produce toxic effects on tumor cells, but also play an auxiliary role in the activation and proliferation of CD8 T cells (Toes et al., 1999; Gerloni and Zanetti, 2005). However, we found that the activated CD4 T cells showed higher infiltration in the high-risk group and were correlated with unfavorable survival outcomes. In the tumor microenvironment, pDC could not effectively activate T cells to kill tumor cells, but induced the production of various regulatory T cells (Tregs) and promoted the immune escape of tumor cells (Conrad and Gilliet, 2013). Th1 was considered to be the most important helper cell type in tumor immunity, it can

directly kill tumor cells by releasing cytokines that activate death receptors on the surface of tumor cells (LaCasse et al., 2011). Th2 mediated immunity was traditionally considered to be conducive to tumor growth, which can not only promote angiogenesis but also inhibit cell-mediated immunity and subsequent tumor cell killing (Ellyard et al., 2007). Therefore, the decreased antitumor immune response may be the potential reason that resulted in the poor prognosis of the high-risk group.

Up to now, the five genes have attracted extensive attention in the field of cancer. Zhang (Zhang et al., 2020) found that the upregulation of *STMN1* promoted the growth of HCC by triggering the MET pathway. Aronova (Aronova et al., 2018) found that *STMN1* was overexpressed in adrenocortical carcinoma and promoted a more invasive phenotype

in vitro. Jiang (Jiang et al., 2018) found that *STMN1* promotes the proliferation, migration, and invasion of esophageal squamous cell carcinoma by activating the PI3K pathway. He (He et al., 2016) found that *STMN1* promotes the growth and invasion of endometrial carcinoma by mediating the secretion and activation of *MMP2* and *MMP9* proteins. Li (Li et al., 2015) found that overexpression of *STMN1* was related to the proliferation, migration, invasion, and apoptosis of human skin squamous cell carcinoma. Bao (Bao et al., 2017) found that the increased expression of *STMN1* is associated with the progression and chemoresistance of lung squamous cell carcinoma. Bai (Bai et al., 2017) found that the high level of *STMN1* in patients with gastric cancer was related to chemoresistance and poor prognosis. Shu (Shu et al., 2017) and Ding (Ding et al., 2020) found that *PON1* had an important diagnostic reference value for AFP negative HCC and was helpful to predict the microvascular invasion of HCC. Yu (Yu et al., 2018) found that the decreased expression of *PON1* represented the high invasiveness of HCC and was closely related to the recurrence and metastasis of HCC. Cao (Cao et al., 2021) found that variations in *PON1* glycosylation may help to distinguish AFP negative HCC from cirrhosis. *PLOD2* has been regarded as an oncogene and its upregulation is closely associated with malignant behavior and poor prognosis in multiple cancers (Du et al., 2017b). For example, Du (Du et al., 2020) found that *PLOD2* promotes aerobic glycolysis and cell progression of colorectal cancer by up regulating *HK2*; Wan (Wan et al., 2020) found that *PLOD2* regulates the migration, invasion and EMT of endometrial cancer cells through *PI3K/Akt* signaling pathway; Kiyozumi (Kiyozumi et al., 2018) found that under hypoxia, *PLOD2* promoted the invasion and migration of gastric cancer cells and led to peritoneal dissemination of gastric cancer; Sheng (Sheng et al., 2019) found that *PLOD2* promotes drug resistance in laryngeal cancer by promoting tumor stem cell like characteristics; Okumura (Okumura et al., 2018) found that hypoxia induced *PLOD2* is a key regulator of epithelial mesenchymal transformation and chemotherapy tolerance in biliary cancer; Du (Du et al., 2017a) found that *PLOD2* is regulated by PI3K/akt-foxa1 axis and promotes the metastasis of non-small cell lung cancer; Noda (Noda et al., 2012) found that *PLOD2* expression was significantly correlated with tumor size and visible intrahepatic metastasis of HCC, which was an independent risk factor for poor prognosis. Liang (Liang et al., 2020) found that *MARCKSL1* promotes the progression of lung adenocarcinoma by regulating epithelial-mesenchymal transition. Zhang (Zhang et al., 2017) found that *SPP1* promotes immune escape of lung adenocarcinoma by mediating macrophage polarization. Zeng (Zeng et al., 2018) found that *SPP1* promotes ovarian cancer progression via Integrin β1/FAK/AKT signaling pathway. Wang (Wang

et al., 2019) found that *SPP1* can promote cell growth in miR-181c targeted HCC.

At present, tumor treatment has stepped into the era of precision medicine, that is, an emerging medical model based on individualized medicine and integrating gene detection, biological information, and big data science. As an important part of establishing a clinical decision-making system, tumor prognosis evaluation has become an important research content of precision medicine. Considering the prognosis of HCC is still a great challenge for medicine in its current stage, the five-gene risk score proposed in the research may provided new insights into the individualized evaluation of HCC prognosis. However, there were still some limitations that should be acknowledged in our work, the results were based on the public data that we did not validate in our samples, and the work was conducted without an experimental mechanism study.

CONCLUSION

The five-gene risk score proposed in the research may provide new insights into the individualized evaluation of HCC prognosis.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: The datasets analyzed for this study were obtained from The Cancer Genome Atlas (TCGA, https://portal.gdc.cancer.gov/), International Cancer Genome Consortium database (ICGC, https://dcc.icgc.org/releases/current/Projects/LIRI-JP), Gene Expression Omnibus(GEO, https://www.ncbi.nlm.nih.gov/geo/), and the molecular signatures database (MSigDB, http://www.gsea-msigdb.org/gsea).

AUTHOR CONTRIBUTIONS

JH and PS designed this study, XF and BQ collected data, JH analyzed the data in this study, interpreted the findings, and drafted the manuscript. PS carried out data management and revised the manuscript. All authors reviewed the final version of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.771819/full#supplementary-material

REFERENCES

Alhmoud, J. F., Woolley, J. F., Al Moustafa, A.-E., and Malki, M. I. (2020). DNA Damage/repair Management in Cancers. *Cancers* 12 (4), 1050. doi:10.3390/cancers12041050

Aravalli, R. N., Steer, C. J., and Cressman, E. N. K. (2008). Molecular Mechanisms of Hepatocellular Carcinoma. *Hepatology* 48 (6), 2047–2063. doi:10.1002/hep.22580

Aronova, A., Min, I. M., Crowley, M. J. P., Panjwani, S. J., Finnerty, B. M., Scognamiglio, T., et al. (2018). STMN1 Is Overexpressed in Adrenocortical Carcinoma and Promotes a More Aggressive Phenotype *In Vitro*. *Ann. Surg. Oncol.* 25 (3), 792–800. doi:10.1245/s10434-017-6296-2

Bai, T., Yokobori, T., Altan, B., Ide, M., Mochiki, E., Yanai, M., et al. (2017). High STMN1 Level Is Associated with Chemo-Resistance and Poor Prognosis in Gastric Cancer Patients. *Br. J. Cancer* 116 (9), 1177–1185. doi:10.1038/bjc.2017.76

Bao, P., Yokobori, T., Altan, B., Iijima, M., Azuma, Y., Onozato, R., et al. (2017). High STMN1 Expression Is Associated with Cancer Progression and Chemo-Resistance in Lung Squamous Cell Carcinoma. *Ann. Surg. Oncol.* 24 (13), 4017–4024. doi:10.1245/s10434-017-6083-0

Barbie, D. A., Tamayo, P., Boehm, J. S., Kim, S. Y., Moody, S. E., Dunn, I. F., et al. (2009). Systematic RNA Interference Reveals that Oncogenic KRAS-Driven Cancers Require TBK1. *Nature* 462 (7269), 108–112. doi:10.1038/nature08460

Cao, X., Cao, Z., Shao, Y., Liu, C., Yan, G., Meng, X., et al. (2021). Analysis of Serum Paraoxonase 1 Using Mass Spectrometry and Lectin Immunoassay in Patients with Alpha-Fetoprotein Negative Hepatocellular Carcinoma. *Front. Oncol.* 11 (1021). doi:10.3389/fonc.2021.651421

Charoentong, P., Finotello, F., Angelova, M., Mayer, C., Efremova, M., Rieder, D., et al. (2017). Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. *Cell Rep.* 18 (1), 248–262. doi:10.1016/j.celrep.2016.12.019

Conrad, C., and Gilliet, M. (2013). Plasmacytoid Dendritic Cells and Regulatory T Cells in the Tumor Microenvironment. *OncoImmunology* 2 (5), e23887. doi:10.4161/onci.23887

Ding, G. Y., Zhu, X. D., Ji, Y., Shi, G. M., Shen, Y. H., Zhou, J., et al. (2020). Serum PON1 as a Biomarker for the Estimation of Microvascular Invasion in Hepatocellular Carcinoma. *Ann. Transl Med.* 8 (5), 204. doi:10.21037/atm.2020.01.44

Du, H., Chen, Y., Hou, X., Huang, Y., Wei, X., Yu, X., et al. (2017a). PLOD2 Regulated by Transcription Factor FOXA1 Promotes Metastasis in NSCLC. *Cell Death Dis* 8 (10), e3143. doi:10.1038/cddis.2017.553

Du, H., Pang, M., Hou, X., Yuan, S., and Sun, L. (2017b). PLOD2 in Cancer Research. *Biomed. Pharmacother.* 90, 670–676. doi:10.1016/j.biopha.2017.04.023

Du, W., Liu, N., Zhang, Y., Liu, X., Yang, Y., Chen, W., et al. (2020). PLOD2 Promotes Aerobic Glycolysis and Cell Progression in Colorectal Cancer by Upregulating HK2. *Biochem. Cell Biol.* 98 (3), 386–395. doi:10.1139/bcb-2019-0256

Ellyard, J. I., Simson, L., and Parish, C. R. (2007). Th2-mediated Anti-tumour Immunity: Friend or Foe? *Tissue Antigens* 70 (1), 1–11. doi:10.1111/j.1399-0039.2007.00869.x

Farazi, P. A., and DePinho, R. A. (2006). Hepatocellular Carcinoma Pathogenesis: from Genes to Environment. *Nat. Rev. Cancer* 6 (9), 674–687. doi:10.1038/nrc1934

Fattovich, G., Stroffolini, T., Zagni, I., and Donato, F. (2004). Hepatocellular Carcinoma in Cirrhosis: Incidence and Risk Factors. *Gastroenterology* 127 (5), S35–S50. doi:10.1053/j.gastro.2004.09.014

Gerloni, M., and Zanetti, M. (2005). "CD4 T Cells in Tumor Immunity," in *Springer Seminars in Immunopathology* (Springer). 27 (1), 37–48. doi:10.1007/s00281-004-0193-z

Gidron, Y., Russ, K., Tissarchondou, H., and Warner, J. (2006). The Relation between Psychological Factors and DNA-Damage: a Critical Review. *Biol. Psychol.* 72 (3), 291–304. doi:10.1016/j.biopsycho.2005.11.011

Gillman, R., Floro, K. L., Wankell, M., and Hebbard, L. (2021). The Role of DNA Damage and Repair in Liver Cancer. *Biochim. Biophys. Acta Rev. Cancer* 1875, 188493. doi:10.1016/j.bbcan.2020.188493

Gomaa, A. I., Khan, S. A., Toledano, M. B., Waked, I., and Taylor-Robinson, S. D. (2008). Hepatocellular Carcinoma: Epidemiology, Risk Factors and Pathogenesis. *Wjg* 14 (27), 4300. doi:10.3748/wjg.14.4300

He, X., Liao, Y., Lu, W., Xu, G., Tong, H., Ke, J., et al. (2016). Elevated STMN1 Promotes Tumor Growth and Invasion in Endometrial Carcinoma. *Tumor Biol.* 37 (7), 9951–9958. doi:10.1007/s13277-016-4869-5

Huo, J., Cai, J., Guan, G., Liu, H., and Wu, L. (2021a). A Ferroptosis and Pyroptosis Molecular Subtype-Related Signature Applicable for Prognosis and Immune Microenvironment Estimation in Hepatocellular Carcinoma. *Front. Cell Dev. Biol.* 9 (3118), 761839. doi:10.3389/fcell.2021.761839

Huo, J., Wu, L., and Zang, Y. (2021c). Construction and Validation of a Reliable Six-Gene Prognostic Signature Based on the TP53 Alteration for Hepatocellular Carcinoma. *Front. Oncol.* 11, 618976. doi:10.3389/fonc.2021.618976

Huo, J., Wu, L., and Zang, Y. (2021d). Construction and Validation of a Universal Applicable Prognostic Signature for Gastric Cancer Based on Seven Immune-Related Gene Correlated with Tumor Associated Macrophages. *Front. Oncol.* 11, 635324. doi:10.3389/fonc.2021.635324

Huo, J., Wu, L., and Zang, Y. (2021e). Development and Validation of a Metabolic-Related Prognostic Model for Hepatocellular Carcinoma. *J. Clin. Transl Hepatol.* 9 (2), 169–179. doi:10.14218/jcth.2020.00114

Huo, J., Wu, L., and Zang, Y. (2021g). Development and Validation of a Robust Immune-Related Prognostic Signature for Gastric Cancer. *J. Immunol. Res.* 2021, 5554342. doi:10.1155/2021/5554342

Huo, J., Wu, L., and Zang, Y. (2021b). Eight-gene Prognostic Signature Associated with Hypoxia and Ferroptosis for Gastric Cancer with General Applicability. *Epigenomics* 13 (11), 875–890. doi:10.2217/epi-2020-0411

Huo, J., Wu, L., and Zang, Y. (2021h). Eleven Immune-Gene Pairs Signature Associated with TP53 Predicting the Overall Survival of Gastric Cancer: a Retrospective Analysis of Large Sample and Multicenter from Public Database. *J. Transl Med.* 19 (1), 183. doi:10.1186/s12967-021-02846-x

Huo, J., Wu, L., Zang, Y., and Medicine, M. (2021f). Development and Validation of a CTNNB1-associated Metabolic Prognostic Model for Hepatocellular Carcinoma. *J. Cell Mol Med* 25 (2), 1151–1165. doi:10.1111/jcmm.16181

Jackson, S. P., and Bartek, J. (2009). The DNA-Damage Response in Human Biology and Disease. *Nature* 461 (7267), 1071–1078. doi:10.1038/nature08467

Jiang, W., Huang, S., Song, L., and Wang, Z. (2018). STMN1, a Prognostic Predictor of Esophageal Squamous Cell Carcinoma, Is a Marker of the Activation of the PI3K Pathway. *Oncol. Rep.* 39 (2), 834–842. doi:10.3892/or.2017.6145

Kerzendorfer, C., and O'Driscoll, M. (2009). Human DNA Damage Response and Repair Deficiency Syndromes: Linking Genomic Instability and Cell Cycle Checkpoint Proficiency. *DNA Repair* 8 (9), 1139–1152. doi:10.1016/j.dnarep.2009.04.018

Kiyozumi, Y., Iwatsuki, M., Kurashige, J., Ogata, Y., Yamashita, K., Koga, Y., et al. (2018). PLOD2 as a Potential Regulator of Peritoneal Dissemination in Gastric Cancer. *Int. J. Cancer* 143 (5), 1202–1211. doi:10.1002/ijc.31410

LaCasse, C. J., Janikashvili, N., Larmonier, C. B., Alizadeh, D., Hanke, N., Kartchner, J., et al. (2011). Th-1 Lymphocytes Induce Dendritic Cell Tumor Killing Activity by an IFN-γ-dependent Mechanism. *J.I.* 187 (12), 6310–6317. doi:10.4049/jimmunol.1101812

Lewis, S. E. M., John Aitken, R., Conner, S. J., Iuliis, G. D., Evenson, D. P., Henkel, R., et al. (2013). The Impact of Sperm DNA Damage in Assisted conception and beyond: Recent Advances in Diagnosis and Treatment. *Reprod. BioMedicine Online* 27 (4), 325–337. doi:10.1016/j.rbmo.2013.06.014

Li, X., Wang, L., Li, T., You, B., Shan, Y., Shi, S., et al. (2015). STMN1 Overexpression Correlates with Biological Behavior in Human Cutaneous Squamous Cell Carcinoma. *Pathol. - Res. Pract.* 211 (11), 816–823. doi:10.1016/j.prp.2015.07.009

Liang, W., Gao, R., Yang, M., Wang, X., Cheng, K., Shi, X., et al. (2020). MARCKSL1 Promotes the Proliferation, Migration and Invasion of Lung Adenocarcinoma Cells. *Oncol. Lett.* 19 (3), 2272–2280. doi:10.3892/ol.2020.11313

Noda, T., Yamamoto, H., Takemasa, I., Yamada, D., Uemura, M., Wada, H., et al. (2012). PLOD2 Induced under Hypoxia Is a Novel Prognostic Factor for Hepatocellular Carcinoma after Curative Resection. *Liver Int.* 32 (1), 110–118. doi:10.1111/j.1478-3231.2011.02619.x

Okumura, Y., Noda, T., Eguchi, H., Sakamoto, T., Iwagami, Y., Yamada, D., et al. (2018). Hypoxia-induced PLOD2 Is a Key Regulator in Epithelial-Mesenchymal Transition and Chemoresistance in Biliary Tract Cancer. *Ann. Surg. Oncol.* 25 (12), 3728–3737. doi:10.1245/s10434-018-6670-8

Ozakyol, A. (2017). Global Epidemiology of Hepatocellular Carcinoma (HCC Epidemiology). *J. Gastrointest. Canc* 48 (3), 238–240. doi:10.1007/s12029-017-9959-0

Roos, W. P., and Kaina, B. (2006). DNA Damage-Induced Cell Death by Apoptosis. *Trends Mol. Med.* 12 (9), 440–450. doi:10.1016/j.molmed.2006.07.007

Sayiner, M., Golabi, P., and Younossi, Z. M. (2019). Disease burden of Hepatocellular Carcinoma: a Global Perspective. *Dig. Dis. Sci.* 64 (4), 910–917. doi:10.1007/s10620-019-05537-2

Schütte, K., Bornschein, J., and Malfertheiner, P. (2009). Hepatocellular Carcinoma-Eepidemiological Trends and Risk Factors. *Dig. Dis.* 27 (2), 80–92. doi:10.1159/000218339

Sheng, X., Li, Y., Li, Y., Liu, W., Lu, Z., Zhan, J., et al. (2019). PLOD2 Contributes to Drug Resistance in Laryngeal Cancer by Promoting Cancer Stem Cell-like Characteristics. *BMC Cancer* 19, 840. doi:10.1186/s12885-019-6029-y

Shu, H., Li, W., Shang, S., Qin, X., Zhang, S., and Liu, Y. (2017). Diagnosis of AFP-Negative Early-Stage Hepatocellular Carcinoma Using Fuc-PON1. *Discov. Med.* 23 (126), 163–168.

Soutoglou, E., and Misteli, T. (2008). Activation of the Cellular DNA Damage Response in the Absence of DNA Lesions. *Science* 320 (5882), 1507–1510. doi:10.1126/science.1159051

Tanaka, H., Fujita, N., Sugimoto, R., Urawa, N., Horiike, S., Kobayashi, Y., et al. (2008). Hepatic Oxidative DNA Damage Is Associated with Increased Risk for Hepatocellular Carcinoma in Chronic Hepatitis C. *Br. J. Cancer* 98 (3), 580–586. doi:10.1038/sj.bjc.6604204

Toes, R. E. M., Ossendorp, F., Offringa, R., and Melief, C. J. (1999). CD4 T Cells and Their Role in Antitumor Immune Responses. *J. Exp. Med.* 189 (5), 753–756. doi:10.1084/jem.189.5.753

Wan, J., Qin, J., Cao, Q., Hu, P., Zhong, C., Tu, C., et al. (2020). Hypoxia-induced PLOD2 Regulates Invasion and Epithelial-Mesenchymal Transition in Endometrial Carcinoma Cells. *Genes Genom* 42 (3), 317–324. doi:10.1007/s13258-019-00901-y

Wang, J., Hao, F., Fei, X., and Chen, Y. (2019). SPP1 Functions as an Enhancer of Cell Growth in Hepatocellular Carcinoma Targeted by miR-181c. *Am. J. Transl Res.* 11 (11), 6924–6937.

Yang, J. D., Hainaut, P., Gores, G. J., Amadou, A., Plymoth, A., and Roberts, L. R. (2019). A Global View of Hepatocellular Carcinoma: Trends, Risk, Prevention and Management. *Nat. Rev. Gastroenterol. Hepatol.* 16, 589–604. doi:10.1038/s41575-019-0186-y

Yang, S.-F., Chang, C.-W., Wei, R.-J., Shiue, Y.-L., Wang, S.-N., and Yeh, Y.-T. (2014). Involvement of DNA Damage Response Pathways in Hepatocellular Carcinoma. *Biomed. Res. Int.* 2014, 153867. doi:10.1155/2014/153867

Yu, Z., Ou, Q., Chen, F., Bi, J., Li, W., Ma, J., et al. (2018). Evaluation of the Prognostic Value of Paraoxonase 1 in the Recurrence and Metastasis of Hepatocellular Carcinoma and Establishment of a Liver-specific Predictive Model of Survival. *J. Transl Med.* 16 (1), 327. doi:10.1186/s12967-018-1707-0

Zeng, B., Zhou, M., Wu, H., and Xiong, Z. (2018). SPP1 Promotes Ovarian Cancer Progression via Integrin β1/FAK/AKT Signaling Pathway. *Onco Targets Ther.* 11, 1333–1343. doi:10.2147/ott.s154215

Zhang, R., Gao, X., Zuo, J., Hu, B., Yang, J., Zhao, J., et al. (2020). STMN1 Upregulation Mediates Hepatocellular Carcinoma and Hepatic Stellate Cell Crosstalk to Aggravate Cancer by Triggering the MET Pathway. *Cancer Sci.* 111 (2), 406–417. doi:10.1111/cas.14262

Zhang, Y., Du, W., Chen, Z., and Xiang, C. (2017). Upregulation of PD-L1 by SPP1 Mediates Macrophage Polarization and Facilitates Immune Escape in Lung Adenocarcinoma. *Exp. Cell Res.* 359 (2), 449–457. doi:10.1016/j.yexcr.2017.08.028