Check for updates

A Prognostic Gene Signature for Hepatocellular Carcinoma

Rong Chen[1], Meng Zhao[2], Yanli An[3,4], Dongfang Liu[3], Qiusha Tang[3] and Gaojun Teng[4*]

[1] Department of Oncology, Zhongda Hospital, Nanjing, China, [2] School of Basic Medicine, Zhengzhou University, Zhengzhou, China, [3] Medical School of Southeast University, Nanjing, China, [4] Department of Radiology, Medical School of Southeast University, Nanjing, China

Hepatocellular carcinoma is the third most common cause of cancer-related deaths in China and immune-based therapy can improve patient outcomes. In this study, we investigated the relationship between immunity-associated genes and hepatocellular carcinoma from the prognostic perspective. The data downloaded from The Cancer Genome Atlas Liver Hepatocellular Carcinoma (TCGA-LIHC) and the Gene Expression Omnibus (GEO) was screened for gene mutation frequency using the maftools package. Immunity-associated eight-gene signature with strong prognostic ability was constructed and proved as an independent predictor of the patient outcome in LIHC. Seven genes in the immune-related eight-gene signature were strongly associated with the infiltration of M0 macrophages, resting mast cells, and regulatory T cells. Our research may provide clinicians with a quantitative method to predict the prognosis of patients with liver cancer, which can assist in the selection of the optimal treatment plan.

Keywords: hepatocellular carcinoma, bioinformatics analysis, immune-related prognostic genes, immune cells, gene signature

INTRODUCTION

Hepatocellular carcinoma (HCC) is the fifth most common cancer in the world (1, 2). China accounts for 55% of new HCC cases and HCC-related deaths annually (3). Standard treatment approaches for HCC include surgery, liver transplantation, targeted therapy, radiotherapy, immunotherapy, and chemotherapy; however, the therapeutic effect is still not satisfactory (4). Thus, in China a 5-year survival rate for patients with HCC is only 14.1% (5). The overall poor outcome can be attributed to the fact that patients are already at the advanced stage when diagnosed and only less than 30% of them can be operated (6). Therefore, in order to improve patients' survival, it is important to explore new diagnostic and therapeutic targets, including disease-specific biomarkers and prognostic molecular models.

Mutations in the TP53 gene encoding an important tumour suppressor protein are commonly found in diverse human cancers (7–9). Wild-type TP53 can activate apoptosis-related pathways to induce cancer cell death and prevent tumour growth (10), whereas loss-of-function mutations in the TP53 gene can induce uncontrolled tumour cell proliferation (11, 12), as shown for oesophageal cancer (13, 14). In breast cancer, the frequency of TP53 mutations is as high as 80%, which exceeds even that of BRCA1 mutations (15) and which can account for shorter lifespan of patients with mutated TP53 (16). In high-grade ovarian cancer, the TP53 mutation rate is approximately 97% (17). Adavosertib can increase the chemotherapeutic drug sensitivity of cancer cells harbouring

TP53 mutations (18, 19), and it was reported that in patients with platinum-sensitive ovarian cancer, adavosertib combined with paclitaxel and carboplatin can improve progression-free survival (PFS).

TP53 mutations are also very common in HCC and have been detected in 13–48% of patients (20–24). Patients with HCC and mutated TP53 had shorter overall and relapse-free survival (25). Previous studies indicate that the TP53 mutation status is associated with distinct immune reactions (26, 27); thus, in HCC TP53 genetic alterations resulted in decreased immune response (28). It has been reported that mutations in both low-density lipoprotein receptor-related protein 1B (LRP1B) and TP53 may be a prognostic biomarker predicting a better effect of immunotherapy in patients with HCC (29). The predictive value of the TP53 mutation status was also shown in the treatment and prognosis of other cancers. Thus, in squamous cell carcinoma of the head and neck (HNSCC), mutations in methylguanine-DNA methyltransferase and TP53 were related to a poorer prognosis (30). Recent studies indicate that the cooperative effect of poly (ADP-ribose) polymerase (PARP) inhibitors and ionic radiation or chemotherapy depends on the TP53 function (31, 32). Moreover, reactivation of mutant TP53 combined with olaparib resulted in more efficient inhibition of tumour growth in the preclinical model of triple negative breast cancer with a TP53 mutation (33). PARP inhibitors showed activity in a subset of colorectal cancer cell lines and preservation of the TP53 function may increase the likelihood of a favourable response (34).

As TP53 mutations play a significant role in many types of cancers including HCC, previous studies have been limited to the investigation of this particular gene (22, 23, 28). However, it is established that cancer is a heterogeneous multi-stage disease caused by the interaction of numerous gene products and signalling factors. Therefore, an integrative functional genomics approach should help in deciphering the molecular features of liver cancer. In previous studies, either the sample size was insufficient, which undermined the reliability of conclusions (35), or the records on baseline clinical features and therapeutic regimens and even the information included in the dataset were incomplete (36). Therefore, comprehensive testing and analysis are required to identify more reliable diagnostic biomarkers and therapeutic targets in HCC.

In this study, we used The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) databases to obtain and screen highly mutated genes in primary liver cancer, construct an immune-related gene signature, and explore the relationship between immune cells and patient prognosis.

MATERIALS AND METHODS

Acquisition of Liver Hepatocellular Carcinoma (LIHC) Data and Screening of Highly Mutated Genes

The gene expression data on 364 cases of LIHC were downloaded from TCGA database (37) using RTCGAToolbox (38) and used

as a training set. The LIHC gene chip and clinical survival data of 115 LIHC cases contained in the GSE76427 dataset (39) were downloaded from the GEO database (40) and used as the validation set. Maftools (41) was used to identify the top 20 highly mutated genes and visualize mutations and their frequencies in all samples, which were then divided into two groups according to the presence of mutations in the gene with the highest mutation frequency.

Gene Set Enrichment Analysis (GSEA) and Gene Set Variation Analysis (GSVA)

To determine the pathways differentially expressed between patient groups, we performed GSEA, a computational method to detect functionally relevant genes (42, 43) and GSVA, a non-parametric approach to calculate sample-wise gene set enrichment scores for gene expression data (44). For enrichment analysis, we used GSEA 4.0.3 software and file 'c2.all.v.7.1.symbols.gmt' as the reference gene set, and performed 1,000 genome replacements to determine the standardized enrichment score for each analysis; P-values and false discovery rates less than 0.05 were considered to indicate statistical significance. File 'h.all.v7.1.symbols.gmt' was used as the reference gene set for GSVA performed with clusterProfiler (45). P-value less than 0.05 indicated statistical significance.

Determination of the Immunity-Associated Gene Signature

Univariate Cox regression was used to analyze the association between immunity-related genes and the prognosis of patients with LIHC; forest plots were constructed for visualization. Screening of immune genes correlated with disease prognosis was used as the basis for signature construction; P <0.05 was the selection criterion. We applied machine learning methods and Lasso regression, which is widely used in search of prognostic biomarkers (46), to generate a new gene combination for each iteration; 1,000 Lasso regressions were performed on candidate genes and the best gene signature was determined based on the area under the curve (AUC). Next, we calculated the risk score for each patient according to the gene expression level and divided patients into groups. The optimal prognostic immune-gene signature was verified by Cox regression analysis.

Gene Ontology (GO) and Kyoto Encyclopaedia of Genes and Genomes (KEGG) Analyses

GO is widely used to annotate gene functions (47), and KEGG is a common method to analyze pathway enrichment (48). For studying the functions of genes associated with LIHC prognosis and the related molecular mechanisms, an R package clusterProfiler (45) was applied to perform GO and KEGG analyses.

External Validation of the Immunity-Related Gene Signature

The risk score for each sample in the validation set was calculated according to the optimal gene signature and used to assign

patients to high- and low-risk groups. The receiver operating characteristic (ROC) curve at different time points was used to analyze the prognostic potential of the optimal gene signature in the validation set.

Clinical Subgroup Analysis and Nomogram Construction

We grouped patients with LIHC according to clinical characteristics (gender, age, pathological stage, TNM stage, and metastasis) and performed Kaplan-Meier analysis on the samples in each group. Next, we constructed a nomogram including the predictive information on clinical features and gene signature.

Analysis of Correlation Between Immune Cell Infiltration and the Optimal Gene Signature

An online CIBERSORT tool (49) was used to analyze the distribution and infiltration of 22 types of immune cells in the high- and low-risk groups. Principal Component Analysis (PCA) was applied to the data to determine the difference in immune cell infiltration between the two groups. We also evaluated the inter-group differences in the composition, interaction, and infiltration of the 22 immune cell types. Further, the association between immune cell infiltration and LIHC prognosis was explored using Kaplan-Meier analysis.

q-RCR

Total RNA from cells was extracted with TRIzol reagent (Thermo Fisher Scientific, 15596026) following the manufacturer's instructions. Complementary DNA (cDNA) was synthesized and PCRs with cDNA as template were performed using a real-time detector (Analytik Jena AG, qTower 3.2G; Jena, Germany) using BeyoFast SYBR Green qPCR Mix (Bio-Rad, 1708882AP, Shanghai, China). The primer sequences were as follows: GAPDH Forward: 5'-AC AGCCTCAAGATCATCAGC-3'; GAPDH Reverse: 5'-GGT CATGAGTCCTTCCACGAT-3'; CCR3 Forward: 5'- CACA AGCCAGGGAGAAGTGA-3'; CCR3 Reverse: 5'- TTTTCA CAGAGCAGGCCCAC -3'; CHGA Forward: 5'- CAGCGGT TTTGAAGATGAACTC -3'; CHGA Reverse: 5'- ACTTTT CTCTGCCTCCTTGGAA -3'; EPO Forward: 5'- GCTGCAT GTGGATAAAGCCG -3'; EPO Reverse: 5'- TGATT GTTCGGAGTGGAGCA -3'; LECT2 Forward: 5'- CTGCTCA AAGAAGTCAGAGGC -3'; LECT2 Reverse: 5'- GCGTACA CAGTAGATCCAGCA -3'; NROB1 Forward: 5'- AGGGGGTA AAGAGGCGCTA -3'; NROB1 Reverse: 5'- CTTGA TTTGTGCTCGTGGGC -3'; S100A9 Forward: 5'- GGAACG CAACATAGAGACCA -3'; S100A9 Reverse: 5'- GATCTTTT CGCACCAGCTCTT -3'; SEMA4F Forward: 5'- CCTGC CTCCCACACACTTTA -3'; SEMA4F Reverse: 5'- ACC ATCCAGTCAATCCTGCG -3'; SPP1 Forward: 5'- CAAA TACCCAGATGCTGTGGC -3'; SPP1 Reverse: 5'- TGG TCATGGCTTTCGTTGGA -3'. Transcript levels were normalized against GAPDH levels as an internal reference and were evaluated using the 2-ΔΔCt method. All experiments were repeated three times.

Statistical Analysis

All statistical analyses were performed in R package. Cox regression analysis was applied to verify the association of patient survival with the gene signature and the expression of each signature gene. Kaplan-Meier analysis was used to evaluate the survival of patients in the high- and low-risk groups. Pearson correlation analysis was performed to determine the correlation between the prognostic gene signature and infiltration of prognosis-related immune cells. P-values and false discovery rates less than 0.05 were considered to indicate statistical significance.

RESULTS

Mutant Genes in LIHC

A flowchart of this study is shown in **Figure 1**. Using the maftools package (41), we identified the top 20 highly mutated genes in LIHC: SPTA1, CACNA1E, HMCN1, ARID1A, XIRP2, AXIN1, OBSCN, LRP1B, FLG, CSMD3, APOB, BCA13, RYR2, MUC4, PCLO, ALB, MUC16, CTNNB1, TTN, and TP53; among them TP53 had the highest mutation frequency (**Figure 2**).

GSEA and GSVA Results

According to the TP53 mutation status, a total of 364 samples were divided into the TP53 and NO_TP53 groups. GSEA showed that four immune-related pathways: Hoffmann-large-to-small-
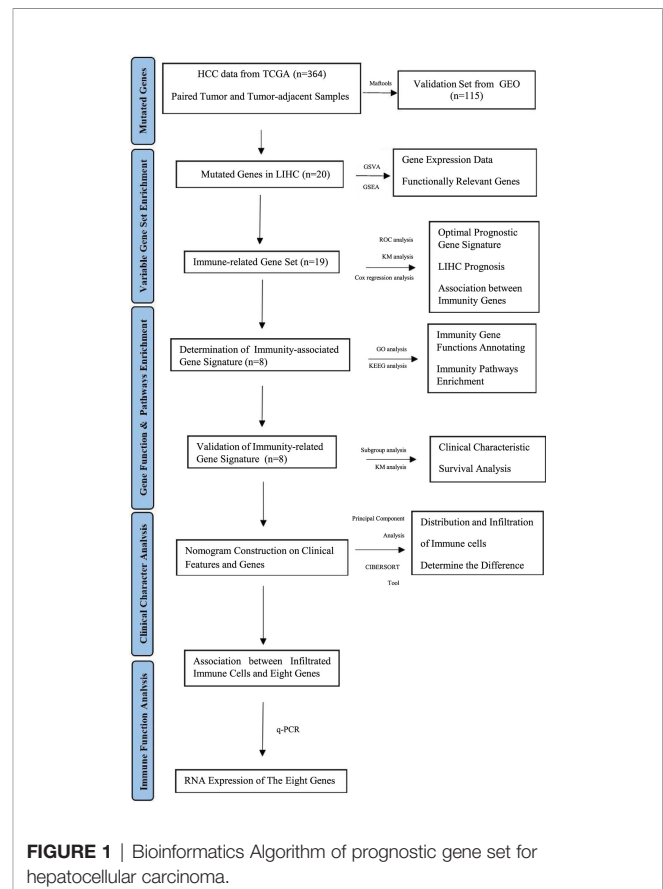
FIGURE 1 | Bioinformatics Algorithm of prognostic gene set for hepatocellular carcinoma.
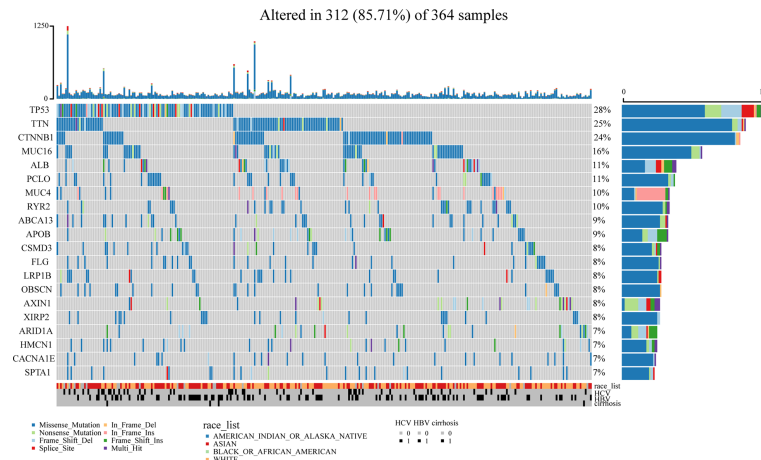
FIGURE 2 | Identification of mutated genes. The maftools package was used and we identified the top 20 highly mutated genes in LIHC. TP53 had the highest mutation frequency among them.

pre-bil-lymphocyte-up, croonquist-IL6-deprivation-dn, mori-large-pre-bil-lymphocyte-up and lee-early-t-lymphocyte-up were enriched in the TP53 group (**Supplementary Figure 1A**). GSVA confirmed that many immune-related KEGG pathways including myc-targets-v1, orc1-signalling, ical-junction, folded-protein-response, apoptotic-spindle, f-targets, 3-pathway, m-checkpoint, response-up, c-targets-v2, glycolysis, apoptosis, 2-stat5-signalling, 3k-akt-mtor-signalling, and complement were activated were enriched in the TP53 group, which further indicating that the activation of TP53 might participate in the process of immune process (**Supplementary Figure 1B**).

Identification of Immune-Related Prognostic Genes and Signature Construction

A single-factor Cox regression model showed that 19 immunity-related genes: BIRC5, CALCR, CCR3, CHGA, COLEC12, CXCL8, EPO, FABP6, FGF9, IKBKE, MAPT, NR0B1, S100A11, S100A2, S100A9, SEMA4F, SPP1, STC2, and TNFRSF11B were associated with LIHC prognosis (**Supplementary Figure 2**). For the accuracy of predicting the optimal gene signature for LIHC, we performed iterative Lasso Cox regression analysis, which identified a prognostic signature comprising eight genes: LECT2, SEMA4F, EPO, CHGA, NR0B1, S100A9, CCR3, and SPP1 (**Figure 3A**). ROC analysis showed that the eight-gene signature had a good predictive ability (**Figure 3B**), whereas Kaplan-Meier analysis revealed that the overall survival of patients in the low-risk group was significantly better than that in the high-risk group (P < 0.001; **Figure 3C**). **Figure 3D** shows the survival status, risk score distribution, and expression of the signature genes. The mortality rate was significantly higher in the high-risk than in the low-risk group and each signature gene was differentially expressed in the two groups. Cumulatively, these results indicated that the signature comprising eight immunity-associated genes could be a significant prognostic indicator in LIHC.

Verification of the Optimal Immune-Associated Gene Signature in the External Validation Set

Analysis of the survival status, risk score distribution, and gene expression of eight-gene signature in the validation set (**Supplementary Figure 3A**) confirmed that the prognosis of patients in the low-risk group was significantly better than that in the high-risk group, thus verifying the prognostic ability of the signature. Kaplan-Meier survival analysis showed that the eight-gene signature could predict the prognosis for patients with LIHC in the external verification set (P = 0.0017) (**Supplementary Figure 3B**). ROC analysis of survival prognosis indicated that the eight-gene signature had a strong ability to predict 3-year (AUC = 0.71), 5-year (AUC = 0.78), and 7-year (AUC = 0.68) survival of patients with LIHC (**Supplementary Figure 3C**). Comparison of the immune-related eight-gene signature with the established LIHC prognostic biomarkers showed that the prediction based on the gene signature was significantly more reliable (**Supplementary Figure 3D**).

Independent Predicting Ability of the Eight-Gene and Construction of a Prognostic Nomogram

Patients were regrouped and Kaplan-Meier survival analysis was performed based on clinicopathological characteristic. The results indicated that even if the clinical features were regrouped, the survival in the high-risk group was always poor (P < 0.05 for all; **Figure 4**). In addition, we combined clinicopathological characteristics and the immune-related eight-gene signature and constructed a prognostic nomogram (**Figure 5**), which could aid in the clinical decision regarding the treatment plan.

Immune Cell Infiltration in the Two Risk Groups

PCA revealed the difference in immune cell infiltration between the two risk groups (**Supplementary Figure 4A**). The results of
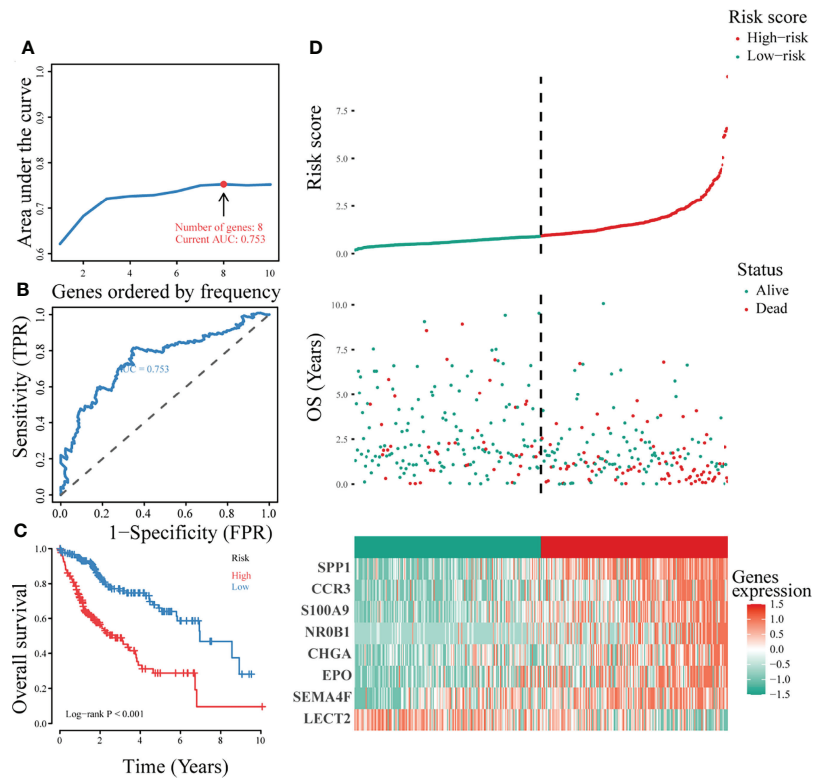
FIGURE 3 | Construction of the optimal immune gene signature associated with LIHC prognosis. **(A)** Iterative Lasso Cox regression analysis used to construct the immune gene signature based on the size of the AUC. **(B)** ROC analysis of the optimal immunity-associated gene signature. **(C)** Kaplan-Meier curves of different risk groups. **(D)** The risk factor association diagram showing risk score distribution, survival status, and expression of the signature genes in the two risk groups.

correlation analysis showed that the infiltration of CD8+ T cells was positively correlated with that of regulatory T cells (Tregs), M1 macrophages, and follicular helper T cells (**Supplementary Figure 4B**). However, the infiltration of naïve B cells was negatively correlated with that of CD8+ T cells, macrophages (M0, M1, and M2), monocytes, resting dendritic and NK cells. Immune cell interaction network revealed that M0 macrophages, activated NK cells, naïve B cells, and resting CD4+ memory T cells had the strongest, whereas activated dendritic and mast cells, naïve CD4+ T cells, and resting dendritic cells – the weakest association with other immune cells (**Supplementary Figure 4C**). Immune cell composition analysis revealed that activated NK cells had the highest infiltration rate and activated dendritic cells – the lowest infiltration rate (**Supplementary Figure 4D**).

Association of Immune Cell Infiltration With LIHC Prognosis and the Eight-Gene Signature

Analysis of the correlation between immune cell infiltration and prognosis showed that the infiltration of gamma delta T cells, eosinophils, and M0 and M2 macrophages indicated a poorer prognosis, whereas that of CD8+ T cells, M1 macrophages, and NK cells suggested a better prognosis for patients with LIHC

(**Figure 6**). The results of the constructed correlation heat map for the signature genes revealed that M0 macrophages, resting mast cells, and Tregs showed negative correlation with CCR3, EPO, NR0B1, S100A9, SEMA4F, and SPP1, and positive correlation with LECT2 (**Figure 7**).

The Relative RNA Expression Level of LECT2, SEMA4F, EPO, CHGA, NR0B1, S100A9, CCR3, and SPP1

The RNA expression of LECT2, SEMA4F, EPO, CHGA, NR0B1, S100A9, CCR3, and SPP1 in normal human hepatic epithelial cells HL-02 and human hepatoma cells BEL_7402 were compared by qPCR. It was found that LECT2, SEMA4F, EPO, CHGA, NR0B1, S100A9, CCR3, and SPP1 were low expressed in human hepatoma cells compared with normal human hepatic epithelial cells (Unpaired t-test, $p<0.01$) (**Figure 8**).

DISCUSSION

The current clinical problems of liver cancer are mainly associated with untimely diagnosis and treatment, which can be attributed to a special double blood supply structure of the liver allowing the formation of a microenvironment providing
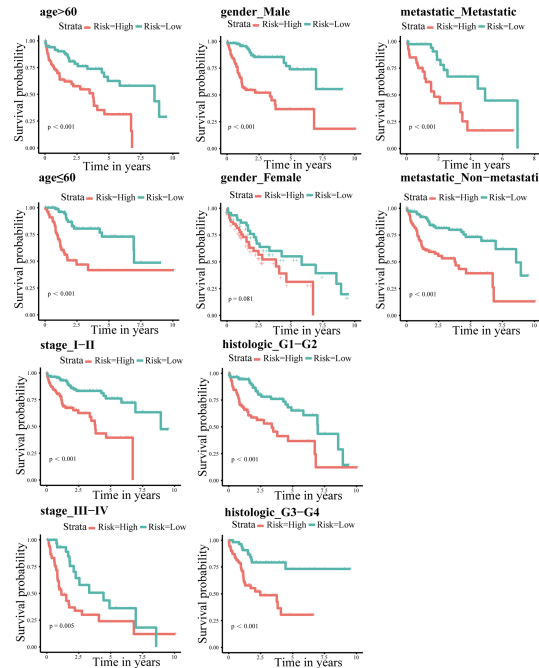
FIGURE 4 | Kaplan-Meier survival analysis according to individual clinicopathological characteristics (age, gender, metastasis, and TNM and pathological staging). Red and green indicate high- and low-risk groups, respectively.

autoimmune tolerance (50). This phenomenon, together with the immune escape of liver cancer cells, indicates that targeted immune-therapeutics should be an effective treatment for HCC (51). In recent years, the research on the mechanism of liver cancer immunotherapy has made great progress (52), but there are still many challenges. Our study identified an immunity-related eight-gene signature that can be used as an independent prognostic indicator of the LIHC outcome, offering a quantitative clinical method to predict patient's survival. Analysis of the association between LIHC prognosis and

immune cell infiltration showed that each gene in the immune-related eight-gene signature was strongly related to M0 macrophage infiltration.

Our analysis of TCGA database indicated that the mutation frequency of the TP53 gene was the highest among the genes mutated in LIHC, and GSEA and GSVA revealed important pathways enriched in patients harbouring TP53 mutations. We also determined an immunity-related eight-gene prognostic signature and performed analysis of its association with immune cell infiltration, which showed that the infiltration of M1
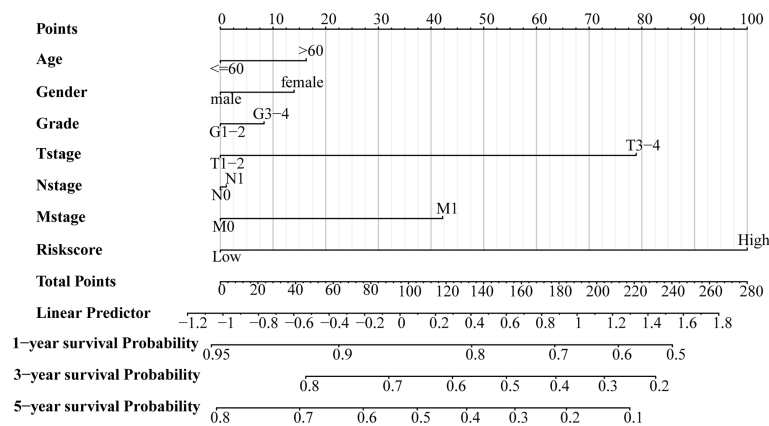
FIGURE 5 | A prognostic nomogram for the overall survival of patients with LIHC.
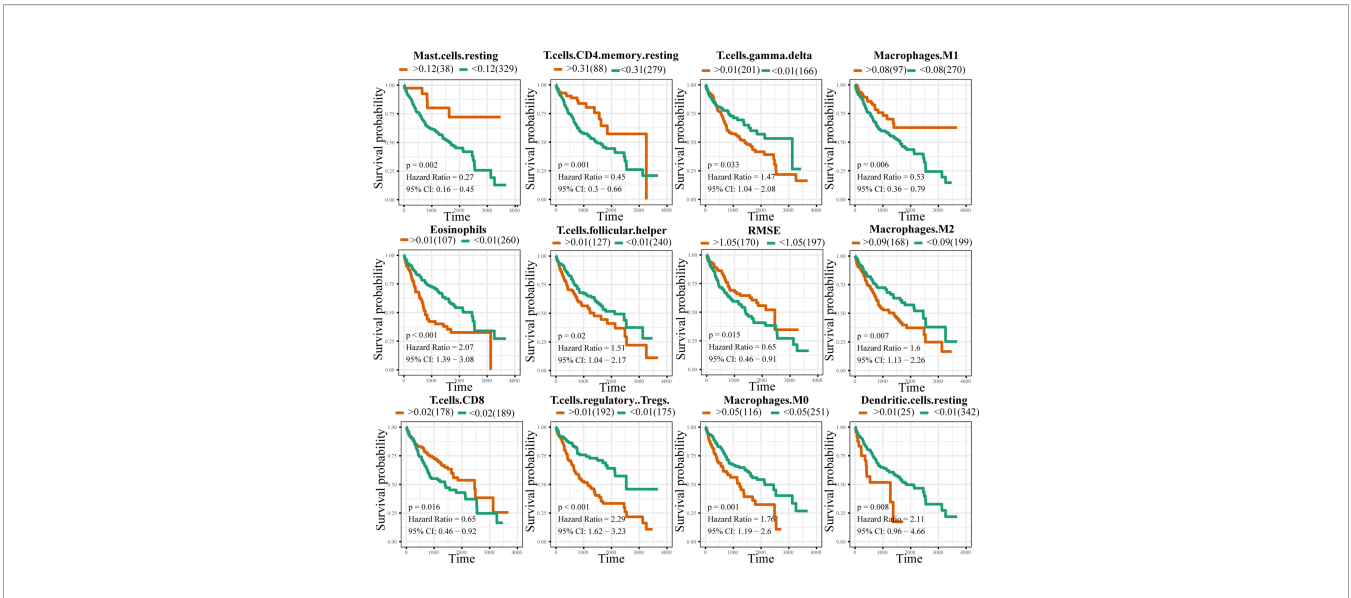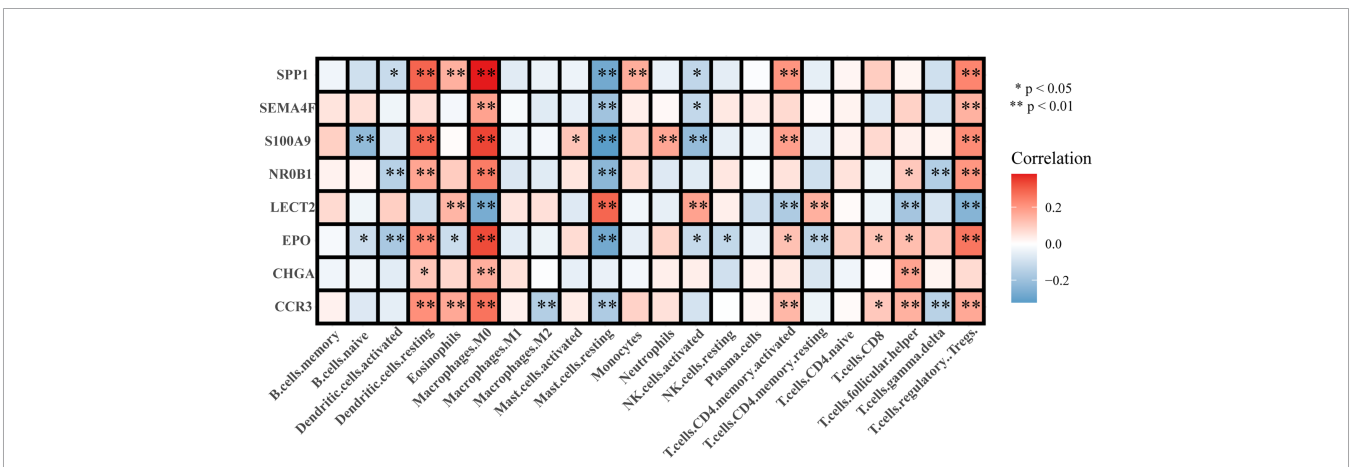
FIGURE 6 | Correlation between immune cell infiltration and LIHC prognosis. Orange and green colours indicate the high- and low-risk groups, respectively.

macrophages, resting CD4+ memory T cells, activated NK cells, and CD8+ T cell suggested a better prognosis. Previous reports indicate that liver cancer tissues are characterized with a high expression level of PD-L1, CTLA4, lymphocyte activation gene 3, and other immunosuppressive molecules, which is negatively associated with the tumour infiltration of IFNγ+ T lymphocytes. Antibody treatment could increase the rate of CD8+ tumour-infiltrating T lymphocytes and the production of cytokines in liver cancer tissues (53). The infiltration of T cells before and after immunotherapy could be used to evaluate the effect of drugs enhancing the response to immune checkpoint blockers and to determine whether T cell infiltration by itself could predict the outcome of immunotherapy (54). It was reported that ependymin related protein 1 (EPDR1) and BRCA1 are correlated with immune cell infiltration and prognosis in HCC (55, 56). Recently, a nine immune-related gene model with an

independent prognostic capability for HCC has been developed and shown to be associated with immune cell infiltration (57).

Using univariate Cox regression, we analysed the relationship among immune-associated genes and the prognosis of patients with LIHC, which were then divided into groups according to risk scores, and the optimal prognostic signature containing eight genes was established using iterative Lasso Cox regression analysis. ROC analysis performed in the external validation set revealed that the immune-associated eight-gene signature had a significant ability to predict 3-, 5-, and 7-year prognosis for patients with LIHC. Furthermore, compared with common prognostic biomarkers of LIHC, the eight-gene signature showed a superior predictive power and was proved to be an independent prognostic predictor of patient survival. We constructed a nomogram combining the clinicopathological characteristics and the immune-related eight-gene signature to offer clinicians a

FIGURE 7 | Correlation between the signature genes and immune cell infiltration.
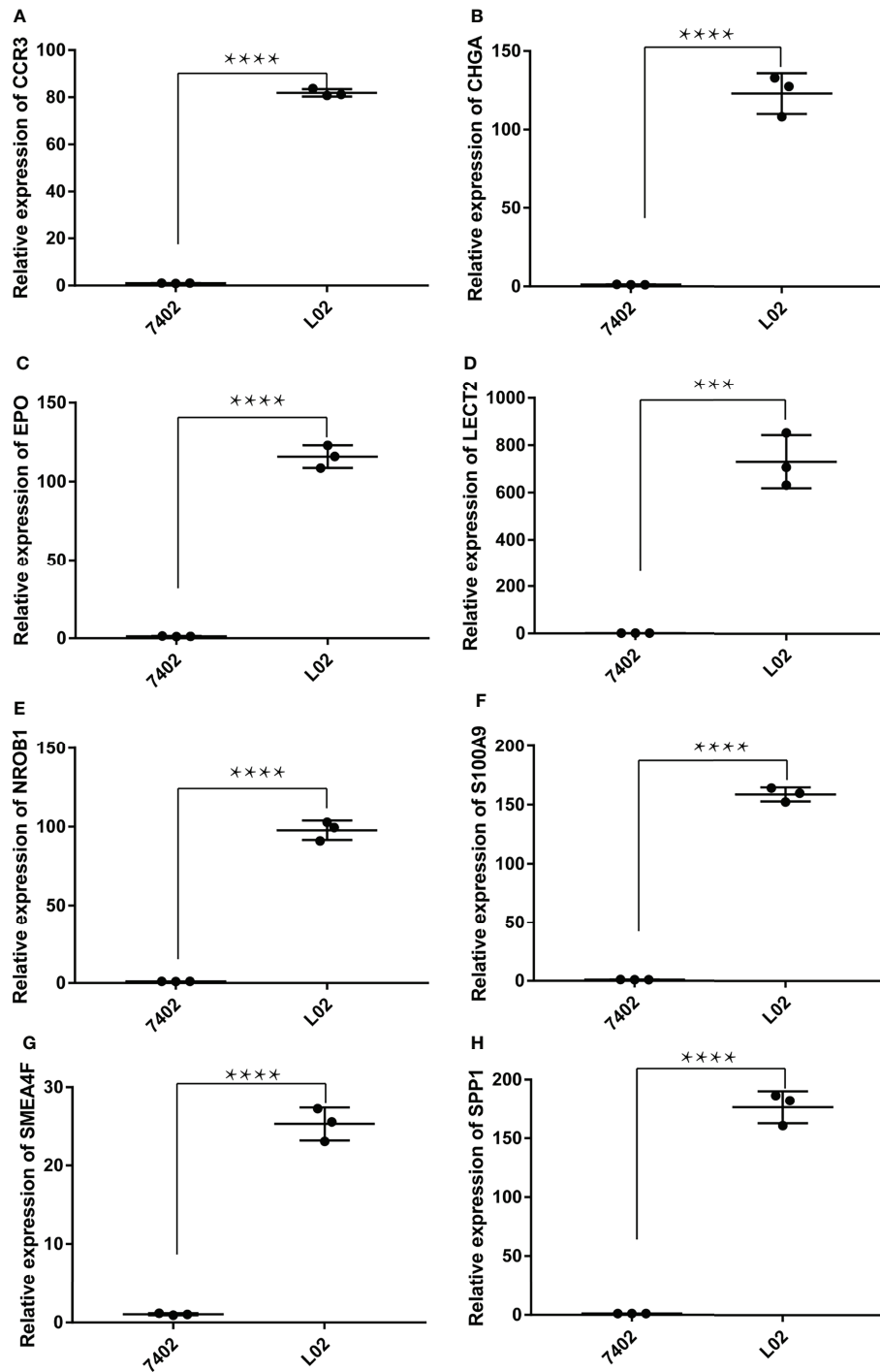
FIGURE 8 | The relative RNA expression level of LECT2, SEMA4F, EPO, CHGA, NR0B1, S100A9, CCR3, and SPP1. The RNA expression of CCR3 **(A)**, CHGA **(B)**, EPO **(C)**, LECT2 **(D)**, NR0B1 **(E)**, S100A9 **(F)**, SEMA4F **(G)** and SPP1 **(H)** were low expressed in human hepatoma cells compared with normal human hepatic epithelial cells (Unpaired t-test, ***$P < 0.001$, ****$P < 0.0001$).

quantitative method for predicting the LICH outcome, which should aid in the selection of optimal treatment approaches. Analysis of immune cell infiltration revealed that Tregs, activated NK cells, and M0 macrophages had the highest,

whereas activated dendritic cells – the lowest infiltration rate in the high-risk group.

The immunity-related genes composing our eight gene prognostic signature have been previously shown to be involved

in oncogenesis. Thus, SEMA4F encoding semaphorin 4F plays a role in axonal growth cone guidance (58) and induction of neurogenesis in prostate cancer (59). The expression of the erythropoietin-encoding EPO gene is related to apoptosis, survival, and proliferation in the early stages of clear cell renal cell carcinoma (60) and has been identified as a distinct prognostic factor for overall and metastases-free survival and locoregional control in locally advanced HNSCC (61). The overexpression of S100A9 encoding calgranulin B has been suggested to play a vital role in the progress of oral squamous cell carcinoma (OSCC) and may serve as a diagnostic and prognostic biomarker for OSCC (62) and nonsmall-cell lung carcinoma (63).

In order to determine the association between LIHC prognosis and immune cell infiltration, we constructed a correlation heat map, which showed that seven genes in the immune-related eight-gene signature were negatively associated with the infiltration of M0 macrophages, resting mast cells, and Tregs. Previous studies have shown that S100A9 plays a significant role in the regulation of immune response and inflammation in most tumours and that it promotes cancer metastasis by accelerating tumour cell proliferation and invasion (64–66), which is consistent with the role of tumour-associated inflammation in supporting metastasis and cancer progression (66–68). Leukocyte cell-derived chemotaxin-2, a 16-kDa secreted protein encoded by the LECT2 gene (69), is involved in the regulation of the tumour microenvironment (70) and plays a critical role in hepatic oncogenesis. Thus, LECT2 deletion modifies the tumour microenvironment and alters cancer phenotypes, suggesting that it is a promising immunotherapeutic target in liver cancer (71). Another study has found that LECT2 expression in HCC is strongly correlated with tumour angiogenesis (72).

It was shown that the expression of C motif chemokine receptor 3 (CCR3) was correlated with malignancy of tumour cells (73). CCR3 ligand may be up-regulated by tumour-related inflammation and involved in the progress of renal cell carcinoma (74), whereas the CCR3/eotaxin-1 loop could induce malignant cell growth in T-cell lymphomas (75, 76). Secreted phosphoprotein 1 (SPP1), also known as osteopontin, is a multifunctional protein first characterized as a biomarker in epithelial cell transformation (77) and suggested to function as an enhancer of HCC growth targeted by miR-181c, thus representing a potential candidate biomarker for HCC diagnosis and therapy (78). We also found that LECT2, SEMA4F, EPO, CHGA, NR0B1, S100A9, CCR3, and SPP1 were low expressed in hepatocellular carcinoma, which may be new cancer therapeutic targets.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

RC analyzed the data and wrote the manuscript. MZ designed the framework of the paper. YA, DL, and QT interpreted data. GT revised the manuscript. The final version of the manuscript has been approved by all authors.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.841530/full#supplementary-material

Supplementary Figure 1 | Comparison of the TP53 and NO_TP53 groups using GSEA and GSVA. **(A)** File 'c2.all.v7.1.symbols.gmt' was used as the reference gene set for GSEA. **(B)** File 'h.all.v7.1.symbols.gmt' was used as the reference gene set for GSVA.

Supplementary Figure 2 | Single-factor Cox regression model.

Supplementary Figure 3 | Confirmation of the optimal gene signature using an external validation set. **(A)** A centralized risk factor correlation diagram for the validation set. **(B)** Kaplan-Meier analysis of the correlation between the eight-gene signature and patient survival. **(C)** ROC analysis of 3-, 5-, and 7-year prognosis for patients with LIHC. **(D)** Time-conditioned ROC curve of the eight-gene signature and common prognostic biomarkers of LIHC.

Supplementary Figure 4 | Immune cell infiltration analysis. **(A)** Correlations in immune cell infiltration. **(B)** Immune cell interaction network; circle size (large to small) is proportional to the intensity of interaction (strong to weak). **(C)** Differences in immune cell infiltration between high- and low-risk groups. **(D)** Immune cell composition analysis.

REFERENCES

1. Jemal A, Siegel R, Ward E, Murray T, Xu J, Thun MJ. Cancer Statistics, 2007. *CA Cancer J Clin* (2007) 57:43–66. doi: 10.3322/canjclin.57.1.43
2. Bruix J, Reig M, Sherman M. Evidence-Based Diagnosis, Staging, and Treatment of Patients With Hepatocellular Carcinoma. *Gastroenterology* (2016) 150:835–53. doi: 10.1053/j.gastro.2015.12.041
3. Wang Q, Zhang P, Li Z, Feng X, Lv C, Zhang H, et al. Evaluation of Polymer Nanoformulations in Hepatoma Therapy by Established Rodent Models. *Theranostics* (2019) 9:1426–52. doi: 10.7150/thno.31683
4. Bruix J, Qin S, Merle P, Granito A, Huang YH, Bodoky G, et al. Regorafenib for Patients With Hepatocellular Carcinoma Who Progressed on Sorafenib Treatment (RESORCE): A Randomised, Double-Blind, Placebo-Controlled, Phase 3 Trial. *Lancet* (2017) 389:56–66. doi: 10.1016/S0140-6736(16)32453-9
5. Allemani C, Matsuda T, Di Carlo V, Harewood R, Matz M, Niksic M, et al. Global Surveillance of Trends in Cancer Survival 2000-14 (Concord-3): Analysis of Individual Records for 37 513 025 Patients Diagnosed With One of 18 Cancers From 322 Population-Based Registries in 71 Countries. *Lancet* (2018) 391:1023–75. doi: 10.1016/S0140-6736(17)33326-3

6. Forner A, Llovet JM, Bruix J. Hepatocellular Carcinoma. *Lancet* (2012) 379:1245–55. doi: 10.1016/S0140-6736(11)61347-0

7. Stratton MR. Exploring the Genomes of Cancer Cells: Progress and Promise. *Science* (2011) 331:1553–8. doi: 10.1126/science.1204040

8. Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, et al. Mutational Landscape and Significance Across 12 Major Cancer Types. *Nature* (2013) 502:333–9. doi: 10.1038/nature12634

9. Hollstein M, Sidransky D, Vogelstein B, Harris CC. p53 Mutations in Human Cancers. *Science* (1991) 253:49–53. doi: 10.1126/science.1905840

10. da Silva GN, Evangelista AF, Magalhaes DA, Macedo C, Bufalo MC, Sakamoto-Hojo ET, et al. Expression of Genes Related to Apoptosis, Cell Cycle and Signaling Pathways Are Independent of TP53 Status in Urinary Bladder Cancer Cells. *Mol Biol Rep* (2011) 38:4159–70. doi: 10.1007/s11033-010-0536-x

11. Cardin R, Piciocchi M, Tieppo C, Maddalo G, Zaninotto G, Mescoli C, et al. Oxidative DNA Damage in Barrett Mucosa: Correlation With Telomeric Dysfunction and p53 Mutation. *Ann Surg Oncol* (2013) 20 Suppl 3:S583–9. doi: 10.1245/s10434-013-3043-1

12. Di Agostino S, Strano S, Blandino G. Gender, Mutant p53 and PML: A Growing "Affaire" in Tumor Suppression and Oncogenesis. *Cell Cycle* (2013) 12:1824–5. doi: 10.4161/cc.25174

13. Greenblatt MS, Bennett WP, Hollstein M, Harris CC. Mutations in the p53 Tumor Suppressor Gene: Clues to Cancer Etiology and Molecular Pathogenesis. *Cancer Res* (1994) 54:4855–78.

14. Song Y, Li L, Ou Y, Gao Z, Li E, Li X, et al. Identification of Genomic Alterations in Oesophageal Squamous Cell Cancer. *Nature* (2014) 509:91–5. doi: 10.1038/nature13176

15. Lalloo F, Evans DG. Familial Breast Cancer. *Clin Genet* (2012) 82:105–14. doi: 10.1111/j.1399-0004.2012.01859.x

16. Olivier M, Langerod A, Carrieri P, Bergh J, Klaar S, Eyfjord J, et al. The Clinical Value of Somatic TP53 Gene Mutations in 1,794 Patients With Breast Cancer. *Clin Cancer Res* (2006) 12:1157–67. doi: 10.1158/1078-0432.CCR-05-1029

17. Ahmed AA, Etemadmoghadam D, Temple J, Lynch AG, Riad M, Sharma R, et al. Driver Mutations in TP53 Are Ubiquitous in High Grade Serous Carcinoma of the Ovary. *J Pathol* (2010) 221:49–56. doi: 10.1002/path.2696

18. Rajeshkumar NV, De Oliveira E, Ottenhof N, Watters J, Brooks D, Demuth T, et al. Mk-1775, a Potent Wee1 Inhibitor, Synergizes With Gemcitabine to Achieve Tumor Regressions, Selectively in p53-deficient Pancreatic Cancer Xenografts. *Clin Cancer Res* (2011) 17:2799–806. doi: 10.1158/1078-0432.CCR-10-2580

19. 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

20. Cancer Genome Atlas Research Network. Electronic address and N. Cancer Genome Atlas Research. Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma. *Cell* (2017) 169:1327–41.e23. doi: 10.1016/j.cell.2017.05.046

21. Zhou X, Hao Q, Lu H. Mutant p53 in Cancer Therapy-the Barrier or the Path. *J Mol Cell Biol* (2019) 11:293–305. doi: 10.1093/jmcb/mjy072

22. Parrales A, Iwakuma T. Targeting Oncogenic Mutant p53 for Cancer Therapy. *Front Oncol* (2015) 5:288. doi: 10.3389/fonc.2015.00288

23. Takai A, Dang HT, Wang XW. Identification of Drivers From Cancer Genome Diversity in Hepatocellular Carcinoma. *Int J Mol Sci* (2014) 15:11142–60. doi: 10.3390/ijms150611142

24. 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

25. Liu J, Ma Q, Zhang M, Wang X, Zhang D, Li W, et al. Alterations of TP53 are Associated With a Poor Outcome for Patients With Hepatocellular Carcinoma: Evidence From a Systematic Review and Meta-Analysis. *Eur J Cancer* (2012) 48:2328–38. doi: 10.1016/j.ejca.2012.03.001

26. Dong ZY, Zhong WZ, Zhang XC, Su J, Xie Z, Liu SY, et al. Potential Predictive Value of TP53 and KRAS Mutation Status for Response to PD-1 Blockade Immunotherapy in Lung Adenocarcinoma. *Clin Cancer Res* (2017) 23:3012–24. doi: 10.1158/1078-0432.CCR-16-2554

27. Biton J, Mansuet-Lupo A, Pecuchet N, Alifano M, Ouakrim H, Arrondeau J, et al. Tp53, STK11, and EGFR Mutations Predict Tumor Immune Profile and the Response to Anti-PD-1 in Lung Adenocarcinoma. *Clin Cancer Res* (2018) 24:5710–23. doi: 10.1158/1078-0432.CCR-18-0163

28. Long J, Wang A, Bai Y, Lin J, Yang X, Wang D, et al. Development and Validation of a TP53-Associated Immune Prognostic Model for Hepatocellular Carcinoma. *EBioMedicine* (2019) 42:363–74. doi: 10.1016/j.ebiom.2019.03.022

29. Tan LL, Jiang XL, Xu LX, Li G, Feng CX, Ding X, et al. TP53-Induced Glycolysis and Apoptosis Regulator Alleviates Hypoxia/Ischemia-Induced Microglial Pyroptosis and Ischemic Brain Damage. *Neural Regen Res* (2021) 16:1037–43. doi: 10.4103/1673-5374.300453

30. Matsuda S, Mafune A, Kohda N, Hama T, Urashima M. Associations Among Smoking, MGMT Hypermethylation, TP53-Mutations, and Relapse in Head and Neck Squamous Cell Carcinoma. *PloS One* (2020) 15:e0231932. doi: 10.1371/journal.pone.0231932

31. Sizemore ST, Mohammad R, Sizemore GM, Nowsheen S, Yu H, Ostrowski MC, et al. Synthetic Lethality of PARP Inhibition and Ionizing Radiation Is P53-Dependent. *Mol Cancer Res* (2018) 16:1092–102. doi: 10.1158/1541-7786.MCR-18-0106

32. Augustine T, Maitra R, Zhang J, Nayak J, Goel S. Sensitization of Colorectal Cancer to Irinotecan Therapy by PARP Inhibitor Rucaparib. *Invest New Drugs* (2019) 37:948–60. doi: 10.1007/s10637-018-00717-9

33. Na B, Yu X, Withers T, Gilleran J, Yao M, Foo TK, et al. Therapeutic Targeting of BRCA1 and TP53 Mutant Breast Cancer Through Mutant p53 Reactivation. *NPJ Breast Cancer* (2019) 5:14. doi: 10.1038/s41523-019-0110-1

34. Smeby J, Kryeziu K, Berg KCG, Eilertsen IA, Eide PW, Johannessen B, et al. Molecular Correlates of Sensitivity to PARP Inhibition Beyond Homologous Recombination Deficiency in Pre-Clinical Models of Colorectal Cancer Point to Wild-Type TP53 Activity. *EBioMedicine* (2020) 59:102923. doi: 10.1016/j.ebiom.2020.102923

35. Niyaz M, Ainiwaer J, Abudureheman A, Zhang L, Sheyhidin I, Turhong A, et al. Association Between TP53 Gene Deletion and Protein Expression in Esophageal Squamous Cell Carcinoma and Its Prognostic Significance. *Oncol Lett* (2020) 20:1855–65. doi: 10.3892/ol.2020.11709

36. Wang L, Yan K, He X, Zhu H, Song J, Chen S, et al. LRP1B or TP53 Mutations Are Associated With Higher Tumor Mutational Burden and Worse Survival in Hepatocellular Carcinoma. *J Cancer* (2021) 12:217–23. doi: 10.7150/jca.48983

37. Cancer Genome Atlas Research Network, Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, et al. The Cancer Genome Atlas Pan-Cancer Analysis Project. *Nat Genet* (2013) 45:1113–20. doi: 10.1038/ng.2764

38. Samur MK. RTCGAToolbox: A New Tool for Exporting TCGA Firehose Data. *PloS One* (2014) 9:e106397. doi: 10.1371/journal.pone.0106397

39. Grinchuk OV, Yenamandra SP, Iyer R, Singh M, Lee HK, Lim KH, et al. Tumor-Adjacent Tissue Co-Expression Profile Analysis Reveals Pro-Oncogenic Ribosomal Gene Signature for Prognosis of Resectable Hepatocellular Carcinoma. *Mol Oncol* (2018) 12:89–113. doi: 10.1002/1878-0261.12153

40. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: Archive for Functional Genomics Data Sets–Update. *Nucleic Acids Res* (2013) 41:D991–5. doi: 10.1093/nar/gks1193

41. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: Efficient and Comprehensive Analysis of Somatic Variants in Cancer. *Genome Res* (2018) 28:1747–56. doi: 10.1101/gr.239244.118

42. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene Set Enrichment Analysis: A Knowledge-Based Approach for Interpreting Genome-Wide Expression Profiles. *Proc Natl Acad Sci USA* (2005) 102:15545–50. doi: 10.1073/pnas.0506580102

43. Subramanian A, Kuehn H, Gould J, Tamayo P, Mesirov JP. Gsea-P: A Desktop Application for Gene Set Enrichment Analysis. *Bioinformatics* (2007) 23:3251–3. doi: 10.1093/bioinformatics/btm369

44. Hanzelmann S, Castelo R, Guinney J. GSVA: Gene Set Variation Analysis for Microarray and RNA-seq Data. *BMC Bioinform* (2013) 14:7. doi: 10.1186/1471-2105-14-7

45. Yu G, Wang LG, Han Y, He QY. clusterProfiler: An R Package for Comparing Biological Themes Among Gene Clusters. *OMICS* (2012) 16:284–7. doi: 10.1089/omi.2011.0118

46. Tibshirani R. The Lasso Method for Variable Selection in the Cox Model. *Stat Med* (1997) 16:385–95. doi: 10.1002/(SICI)1097-0258(19970228)16:4<385::AID-SIM380>3.0.CO;2-3

47. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene Ontology: Tool for the Unification of Biology. The Gene Ontology Consortium. *Nat Genet* (2000) 25:25–9. doi: 10.1038/75556

48. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* (2000) 28:27–30. doi: 10.1093/nar/28.1.27

49. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust Enumeration of Cell Subsets From Tissue Expression Profiles. *Nat Methods* (2015) 12:453–7. doi: 10.1038/nmeth.3337

50. Jenne CN, Kubes P. Immune Surveillance by the Liver. *Nat Immunol* (2013) 14:996–1006. doi: 10.1038/ni.2691

51. Wu ZJL, Li K, Zhang K, Gong WD. [Research Progress of Immunotherapy Alone and in Combination for Liver Cancer]. *Zhonghua Gan Zang Bing Za Zhi* (2020) 28:471–4. doi: 10.3760/cma.j.cn501113-20200520-00262

52. Kroemer G, Galluzzi L, Kepp O, Zitvogel L. Immunogenic Cell Death in Cancer Therapy. *Annu Rev Immunol* (2013) 31:51–72. doi: 10.1146/annurev-immunol-032712-100008

53. Chew V, Lai L, Pan L, Lim CJ, Li J, Ong R, et al. Delineation of an Immunosuppressive Gradient in Hepatocellular Carcinoma Using High-Dimensional Proteomic and Transcriptomic Analyses. *Proc Natl Acad Sci USA* (2017) 114:E5900–9. doi: 10.1073/pnas.1706559114

54. Lu LC, Lee YH, Chang CJ, Shun CT, Fang CY, Shao YY, et al. Increased Expression of Programmed Death-Ligand 1 in Infiltrating Immune Cells in Hepatocellular Carcinoma Tissues After Sorafenib Treatment. *Liver Cancer* (2019) 8:110–20. doi: 10.1159/000489021

55. Chen R, Zhang Y. EPDR1 Correlates With Immune Cell Infiltration in Hepatocellular Carcinoma and Can Be Used as a Prognostic Biomarker. *J Cell Mol Med* (2020) 24:12107–18. doi: 10.1111/jcmm.15852

56. Mei J, Wang R, Xia D, Yang X, Zhou W, Wang H, et al. BRCA1 Is a Novel Prognostic Indicator and Associates With Immune Cell Infiltration in Hepatocellular Carcinoma. *DNA Cell Biol* (2020) 39:1838–49. doi: 10.1089/dna.2020.5644

57. Wang D, Liu J, Liu S, Li W. Identification of Crucial Genes Associated With Immune Cell Infiltration in Hepatocellular Carcinoma by Weighted Gene Co-expression Network Analysis. *Front Genet* (2020) 11:342. doi: 10.3389/fgene.2020.00342

58. Francks C, Fisher SE, Olson RK, Pennington BF, Smith SD, DeFries JC, et al. Fine Mapping of the Chromosome 2p12-16 Dyslexia Susceptibility Locus: Quantitative Association Analysis and Positional Candidate Genes SEMA4F and OTX1. *Psychiatr Genet* (2002) 12:35–41. doi: 10.1097/00041444-200203000-00005

59. Ayala GE, Dai H, Powell M, Li R, Ding Y, Wheeler TM, et al. Cancer-Related Axonogenesis and Neurogenesis in Prostate Cancer. *Clin Cancer Res* (2008) 14:7593–603. doi: 10.1158/1078-0432.CCR-08-1164

60. Stoyanoff TR, Rodriguez JP, Todaro JS, Espada JD, Colavita JP, Brandan NC, et al. Tumor Biology of Non-Metastatic Stages of Clear Cell Renal Cell Carcinoma; Overexpression of Stearoyl Desaturase-1, EPO/EPO-R System and Hypoxia-Related Proteins. *Tumour Biol* (2016) 37:13581–93. doi: 10.1007/s13277-016-5279-4

61. Seibold ND, Schild SE, Gebhard MP, Noack F, Schroder U, Rades D. Prognosis of Patients With Locally Advanced Squamous Cell Carcinoma of the Head and Neck. Impact of Tumor Cell Expression of EPO and EPO-R. *Strahlenther Onkol* (2013) 189:559–65. doi: 10.1007/s00066-013-0320-7

62. Qu ZF, Ma H, Duan XF, Wu R, Zou Y. [The Expression and Significance of S100A9 in Oral Squamous Cell Carcinoma]. *Lin Chung Er Bi Yan Hou Tou Jing Wai Ke Za Zhi* (2017) 31:219–22. doi: 10.13201/j.issn.1001-1781.2017.03.013

63. Koh HM, An HJ, Ko GH, Lee JH, Lee JS, Kim DC, et al. Prognostic Role of S100A8 and S100A9 Protein Expressions in Non-Small Cell Carcinoma of the Lung. *J Pathol Transl Med* (2019) 53:13–22. doi: 10.4132/jptm.2018.11.12

64. Lang B, Shang C, Meng L. Targeted Silencing of S100A8 Gene by miR-24 to Increase Chemotherapy Sensitivity of Endometrial Carcinoma Cells to Paclitaxel. *Med Sci Monit* (2016) 22:1953–8. doi: 10.12659/MSM.899179

65. Srikrishna G. S100A8 and S100A9: New Insights Into Their Roles in Malignancy. *J Innate Immun* (2012) 4:31–40. doi: 10.1159/000330095

66. Lim SY, Yuzhalin AE, Gordon-Weeks AN, Muschel RJ. Tumor-Infiltrating Monocytes/Macrophages Promote Tumor Invasion and Migration by Upregulating S100A8 and S100A9 Expression in Cancer Cells. *Oncogene* (2016) 35:5735–45. doi: 10.1038/onc.2016.107

67. Zhang X, Ai F, Li X, She X, Li N, Tang A, et al. Inflammation-Induced S100A8 Activates Id3 and Promotes Colorectal Tumorigenesis. *Int J Cancer* (2015) 137:2803–14. doi: 10.1002/ijc.29671

68. Funk S, Mark R, Bayo P, Flechtenmacher C, Grabe N, Angel P, et al. High S100A8 and S100A12 Protein Expression Is a Favorable Prognostic Factor for Survival of Oropharyngeal Squamous Cell Carcinoma. *Int J Cancer* (2015) 136:2037–46. doi: 10.1002/ijc.29262

69. Ito M, Nagata K, Kato Y, Oda Y, Yamagoe S, Suzuki K, et al. Expression, Oxidative Refolding, and Characterization of Six-Histidine-Tagged Recombinant Human LECT2, a 16-kDa Chemotactic Protein With Three Disulfide Bonds. *Protein Expr Purif* (2003) 27:272–8. doi: 10.1016/S1046-5928(02)00634-4

70. Hwang HJ, Jung TW, Hong HC, Seo JA, Kim SG, Kim NH, et al. LECT2 Induces Atherosclerotic Inflammatory Reaction *Via* CD209 Receptor-Mediated JNK Phosphorylation in Human Endothelial Cells. *Metabolism* (2015) 64:1175–82. doi: 10.1016/j.metabol.2015.06.001

71. L'Hermitte A, Pham S, Cadoux M, Couchy G, Caruso S, Anson M, et al. Lect2 Controls Inflammatory Monocytes to Constrain the Growth and Progression of Hepatocellular Carcinoma. *Hepatology* (2019) 69:160–78. doi: 10.1002/hep.30140

72. Chen CK, Yu WH, Cheng TY, Chen MW, Su CY, Yang YC, et al. Inhibition of VEGF165/VEGFR2-Dependent Signaling by LECT2 Suppresses Hepatocellular Carcinoma Angiogenesis. *Sci Rep* (2016) 6:31398. doi: 10.1038/srep31398

73. Jin L, Liu WR, Tian MX, Jiang XF, Wang H, Zhou PY, et al. CCL24 Contributes to HCC Malignancy *Via* RhoB- Vegfa-VEGFR2 Angiogenesis Pathway and Indicates Poor Prognosis. *Oncotarget* (2017) 8:5135–48. doi: 10.18632/oncotarget.14095

74. Johrer K, Zelle-Rieser C, Perathoner A, Moser P, Hager M, Ramoner R, et al. Up-Regulation of Functional Chemokine Receptor CCR3 in Human Renal Cell Carcinoma. *Clin Cancer Res* (2005) 11:2459–65. doi: 10.1158/1078-0432.CCR-04-0405

75. Kleinhans M, Tun-Kyi A, Gilliet M, Kadin ME, Dummer R, Burg G, et al. Functional Expression of the Eotaxin Receptor CCR3 in CD30+ Cutaneous T-Cell Lymphoma. *Blood* (2003) 101:1487–93. doi: 10.1182/blood-2002-02-0475

76. Wang C, Wang Y, Hong T, Cheng B, Gan S, Chen L, et al. Blocking the Autocrine Regulatory Loop of Gankyrin/STAT3/CCL24/CCR3 Impairs the Progression and Pazopanib Resistance of Clear Cell Renal Cell Carcinoma. *Cell Death Dis* (2020) 11:117. doi: 10.1038/s41419-020-2306-6

77. Han X, Wang W, He J, Jiang L, Li X. Osteopontin as a Biomarker for Osteosarcoma Therapy and Prognosis. *Oncol Lett* (2019) 17:2592–8. doi: 10.3892/ol.2019.9905

78. Wang J, Hao F, Fei X, Chen Y. SPP1 Functions as an Enhancer of Cell Growth in Hepatocellular Carcinoma Targeted by miR-181c. *Am J Transl Res* (2019) 11:6924–37.