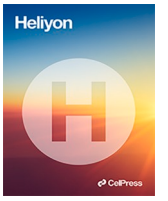
Heliyon 10 (2024) e28156

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Prognostic prediction using a gene signature developed based on exhausted T cells for liver cancer patients

Yu Zhou [a,1], Wanrui Wu [b,1], Wei Cai [a], Dong Zhang [a], Weiwei Zhang [a], Yunling Luo [a], Fujing Cai [a,**], Zhenjing Shi [b,*]

[a] Department of Infectious, The Third Affiliated Hospital of Wenzhou Medical University, Wenzhou, 325000, China
[b] Department of Vasointerventional, The Third Affiliated Hospital of Wenzhou Medical University, Wenzhou, 325000, China

ARTICLE INFO

Keywords:
Single cell profile
T cell exhaustion
Prognosis signature
Liver hepatocellular carcinoma (LIHC)

ABSTRACT

Background: Liver hepatocellular carcinoma (LIHC) is a solid primary malignancy with poor prognosis. This study discovered key prognostic genes based on T cell exhaustion and used them to develop a prognostic prediction model for LIHC.

Methods: SingleR's annotations combined with Seurat was used to automatically annotate the single-cell clustering results of the LIHC dataset GSE166635 downloaded from the Gene Expression Omnibus (GEO) database and to identify clusters related to exhausted T cells. Patients were classified using ConsensusClusterPlus package. Next, weighted gene co-expression network analysis (WGCNA) package was employed to distinguish key gene module, based on which least absolute shrinkage and selection operator (Lasso) and multi/univariate cox analysis were performed to construct a RiskScore system. Kaplan-Meier (KM) analysis and receiver operating characteristic curve (ROC) were employed to evaluate the efficacy of the model. To further optimize the risk model, a nomogram capable of predicting immune infiltration and immunotherapy sensitivity in different risk groups was developed. Expressions of genes were measured by quantitative real-time polymerase chain reaction (qRT-PCR), and immunofluorescence and Cell Counting Kit-8 (CCK-8) were performed for analyzing cell functions.

Results: We obtained 18,413 cells and clustered them into 7 immune and non-immune cell subpopulations. Based on highly variable genes among T cell exhaustion clusters, 3 molecular subtypes (C1, C2 and C3) of LIHC were defined, with C3 subtype showing the highest score of exhausted T cells and a poor prognosis. The Lasso and multivariate cox analysis selected 7 risk genes from the green module, which were closely associated with the C3 subtype. All the patients were divided into low- and high-risk groups based on the medium value of RiskScore, and we found that high-risk patients had higher immune infiltration and immune escape and poorer prognosis. The nomogram exhibited a strong performance for predicting long-term LIHC prognosis. *In vitro* experiments revealed that the 7 risk genes all had a higher expression in HCC cells, and that both liver HCC cell numbers and cell viability were reduced by knocking down MMP-9.

Conclusion: We developed a RiskScore model for predicting LIHC prognosis based on the scRNA-seq and RNA-seq data. The RiskScore as an independent prognostic factor could improve the clinical treatment for LIHC patients.

* Corresponding author.
** Corresponding author.
E-mail addresses: caifujing@qq.com (F. Cai), szj258@126.com (Z. Shi).
[1] Equal Contribution.

https://doi.org/10.1016/j.heliyon.2024.e28156
Received 4 January 2024; Received in revised form 4 February 2024; Accepted 13 March 2024
Available online 16 March 2024

1. Introduction

LIHC is a solid primary malignancy with high morbidity and aggressiveness [1] and accounts for 75–85% of all liver cancer cases [2]. According to the reports of World Health Organization (WHO), the morbidity and mortality of LIHC rank the sixth and third highest among all cancers worldwide, respectively [3,4]. There were 41,210 new cases and 29,380 deaths in 2023 [5] and 42,810 new cases and 30,160 deaths in 2020 [6], showing an increasing trend of morbidity and mortality. LIHC usually has an insidious onset and rapid progression [7,8], but the invasive and metastatic nature of the cancer is the primary cause of death to LIHC patients [9]. Even for patients who have received surgery resection at early stage, more than half of LIHC patients will recur within 2 years after diagnosis [10], with 5-year overall survival lower than 20% [5,11]. Viral hepatitis, liver cirrhosis, alcohol addiction, fatty liver and metabolic syndrome all contribute to the development of LIHC [12]. Current treatment options for LIHC patients include liver transplantation, radioembolization, hepatectomy and targeted therapy [13,14]. Drugs that contain immune checkpoint inhibitors and tyrosine kinase inhibitors [15] are proven to be effective in treating LIHC [15]. Thus, identifying novel biomarkers and developing a reliable prognostic model are crucial to further improve the current LIHC treatment.

The main components of tumor, stromal and immune cells in tumor microenvironment (TME) [16] interact with immune cells to inhibit tumor progression and serve as effective therapeutic targets [17]. It has been believed that the malignant potential of tumors is dominated by the aberrant expression of oncogenes and tumor suppressor genes [18]. However, the dynamic crosstalk between cancer cells and immune cells in TME [16] means that cancer progression is also associated with T cell exhaustion [19], which is a hypofunctional state of naïve CD8+T cells to gradually lose effector function and self-renewal capacity [19]. Major histocompatibility complex (MHC) class II (MHC II) persistently stimulates T cell receptor (TCR) in the chronic exposure of tumor antigen and then activates intracellular transcription factor NFAT [20] to upregulate exhaustion markers or co-inhibitory receptors including TIGIT, PD-1, LAG3, CD39, and CTLA4 [21], thereby inhibiting the production of inflammatory cytokines such as IFNγ, TNF, CTL, and IL-2 [22] as well as the effector function of T cells [19]. Additionally, the quantity of exhausted T cells is also dependent on other factors such as immune contexture, antigenic load and tissue oxygen concentration. Tumor-reactive CD8 + T cells undergo precursor-exhausted T cells to exhaust T cells (T_{EX}), which will ultimately die after overstimulation [23]. T cell exhaustion is related with poor prognosis in certain types of cancers, such as head and neck cell carcinoma [24], angioimmunoblastic cancer [25] and advanced LIHC [26]. Moreover, T cell exhaustion is also regarded as one of the major indicators of immune dysfunction. Studies have found that blocking the expression of surface co-inhibitory receptors of exhausted $CD8^+$ T cells such as PD-1 or CTLA4 can promote the activation of T cells [22]. T cell exhaustion is commonly regarded as a sign of immunotherapies resistance [27]. Therefore, T cell exhaustion can serve as an immunophenotype to help assess immunotherapy effect and patients' prognosis.

With the advance of next-generation sequencing (NGS) technologies, publicly available multi-omics datasets have emerged to facilitate the study of tumor heterogeneity, immunophenotype signature and tumor-TME interactions at molecular level. Bulk RNA-sequencing represents the average expression pattern of a great number of cells [28], while single-cell RNA sequencing (scRNA--seq) can provide subtle feature information of a single tumor or immune cell [29–31]. In this study, the scRNA-seq and bulk RNA-seq data of LIHC samples were integrated to develop a T cell exhaustion-related molecular signature and a reliable prognostic risk model for LIHC. ScRNA-seq data were filtrated, normalized and clustered to identify exhausted T cell clusters. Then highly variable genes from these cells were used for consistent clustering, genes with higher exhausted T cells score were further used to identify key gene modules. Univariate, Lasso and multivariate cox analysis were performed to distinguish key prognostic genes for constructing a RiskScore model. ROC analysis confirmed an accurate classification and long-term predictive performance of the RiskScore.

2. Material and methods

2.1. Data acquisition

The RNA sequencing data of LIHC samples and relevant clinical follow-up information were collected from the Cancer Genome Atlas Database (TCGA-LIHC) [32]. The datasets ICGC-LIRI-JP and GSE14520 were downloaded from the Hepatocellular Carcinoma Cell DataBase (HCCDB, http://lifeome.net/database/hccdb/home.html) [33] and GEO database [34], respectively.

2.2. Data preprocessing

Patients without survival status or time in the TCGA-LIHC and ICGC-LIRI-JP datasets were excluded. Ensembl_id was transformed into Genesymbol, and TPM data in the expression matrix of TCGA-LIHC were converted to log2 format. A total of 365 and 203 LIHC samples from the TCGA-LIHC and ICGC-LIRI-JP datasets were obtained, respectively. In addition, we downloaded the annotation information of GSE14520 and matched the probe label to gene symbol [35]. When one gene matched to multiple gene IDs, its expression value was shown by the mean value.

2.3. Single-cell RNA-seq profile of LIHC

The single-cell RNA-seq dataset GSE166635 containing 2 LIHC samples was downloaded from the GEO database [26]. The filtering settings were as follows: 1) Each gene was expressed in at least three cells, with each cell expressing a minimum of 200 genes; 2) The cells with mitochondrial content< 15% and expressed 200 to 8000 genes were kept [26]. Subsequently, the expression data of 18,413

cells were normalized to select hyper-variable genes via the FindVariableFeatures function in the Seurat R package [36], and the ScaleData function was used to scale gene expression matrix [36]. Principal Component Analysis (PCA) was conducted to detect the anchor point (setting dim = 20), and batch effect was removed by the harmony R package [37]. Then, the RunUMAP function was used for dimensionality reduction (dims = 1:20), followed by cell clustering (resolution = 0.01) using the FindNeighbors and FindClusters functions [36]. The SingleR package was applied to cluster immune cellclusters [38], from which exhausted T cells were identified by FindClusters function (resolution = 0.4). Finally, the FindAllMarkers function was used to screen specifically high-expressed genes from exhausted T cells (setting logfc.threshold = 0.5 and min.pct = 0.25).

2.4. Consistent clustering and immune-infiltration analysis

The key prognostic genes from high-expressed genes in exhausted T cells were selected by performing univariate cox regression analysis [39]. ConsensusClusterPlus R package was used to develop consistency matrix and classify samples in TCGA-LIHC cohort based on these key prognostic genes [40] under clusterAlg = "km", stance = "euclidean", and sampling was repeated for 500 times with each time running 80% of all the samples. The GSVA package was used to calculate the exhausted T cell score, and differences in immune infiltration among the molecular subtypes were analyzed using the ESTIAMTE and MCP-Counter [41]. Pathway activation was analyzed using the h.all.v7.5.1.entrez.gmt gene set by the GSEA in the GSEA R package [39].

2.5. WGCNA analysis

WGCNA was performed to determine gene modules related to the molecular subtypes [42]. Gene modules containing at least 50 genes (minModuleSize = 50) were selected by performing hierarchical clustering. Next, we analyzed the relationship between the modules and subtypes and determined the gene module related to T cell exhaustion. The Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) analysis were performed using the clusterProfiler R package to investigate gene function in each relevant module [43].

2.6. Development of a risk model related to T cell exhaustion and validation analysis

The survival R package was used to perform univariate cox proportional hazard regression on T cell exhaustion-related genes (filtered under the threshold of $p < 0.05$). Then Lasso and multivariate stepwise cox regression analysis was used to reduce the number of candidate signature genes and calculate the regression coefficient [42]. The risk formula was as follow: RiskScore = Σ(regression coefficient) $*$ (expression of gene) [32]. We calculated the RiskScore of patients and classified the patients based on the median value, and then prognostic classification was examined by ROC analysis [44]. Tumor Immune Dysfunction and Exclusion (TIDE) score is an indicator of blocking therapy using immune checkpoints (*anti*-PD-1/*anti*-CTLA-4) [42]. In addition, the Immune cell Proportion Score (IPS) was applied to evaluate the patients' potential benefit from taking immunotherapy [45]. According to patients' responses to the *anti*-PD-L1 receptor blockers, the IMvigor210 dataset was included to explore immunotherapy benefit in terms of complete response (CR), stable disease (SD), partial response (PR), and progressive disease (PD) [46].

2.7. Cell culture and transfection

Human hepatocytes (THLE-2) purchased from Cellcook Biotech Company (Guangzhou, China) were cultured in BEGM (Gibco, USA) supplemented with fetal bovine serum (Gibco, USA) and penicillin/streptomycin. HCC cells (HepaRG and SK-HEP-1) purchased from Typical Culture Reserve Center of China (Shanghai, China) were cultured in cultured in DMEM (Gibco, USA) containing 10% fetal bovine serum. Cell culture was performed under the conditions of 5% CO_2 at 37 °C. Following the instructions of the manufacturer, MMP9 siRNA was transfected applying the Lipofectamine 2000 (Invitrogen, USA). The images of negative control siRNA-treated cells (si-NC) collected under a fluorescence microscope and qRT-PCR were used to verify the efficiency of transfection. The sequences of siRNA for MMP9 were as follows: si MMP9#1: 5′-ACCACAACATCACCTATTGGATC-3′, si MMP9#2: 5′-CACCTATTGGATCCAAAAC-TACT-3'.

Table 1
The primers information of genes.

Gene	Forward primer sequence (5′-3′)	Reverse primer sequence (5′-3′)
IL7R	GAAGCAGGGACATAGAGCAAC	TTCTAGCCAGGCATCTTAGGG
PIP4K2A	AAGAAGAAGCACTTCGTAGCG	ATGGCTCAGTTCATTGATCGAG
P2RY6	GTGAGGATTTCAAGCGACTGC	TCCCCTCTGGCGTAGTTATAGA
GZMH	CAGCAGAGCCACTGAGATTT	TACCTTGGGCTACGTCCTTA
WARS	AGCACCTACCAGTAATCATGGC	TCCAAACCGAACAATGAGCTT
CCR7	ATTTGTTTCGTGGGCCTACTG	TCATGGTCTTGAGCCTCTTGA
MMP9	GTGTGAGTACCCGGAACGAG	TGTATCCGGCAAACTGGTCC
GAPDH	AGGTCGGTGTGAACGGATTTG	TGTAGACCATGTAGTTGAGGTCA

2.8. QRT-PCR

Total RNA from HepaRG, SK-HEP-1 and THLE-2 cell lines was extracted utilizing TRIzol (Thermo Fisher, USA) reagent. Applying the HiScript II SuperMix (Vazyme, China), cDNA was synthesized from the RNA (5000 ng). The PCR was conducted under the amplification conditions of 45 cycles at 95 °C for 30 s (s), 95 °C for 10 s, and at 60 °C for 30 s. GAPDH was an internal reference. The sequences of primer pairs for the genes were displayed in Table 1.

2.9. Cell viability

Cell Counting Kit-8 assay (Beyotime, China) was carried out for detecting cell viability. Briefly, cells with different treatments were cultured in 96-well plates (1 × 103 cells per well) and added withCCK-8 solution at specific time points. The O.D 450 value was

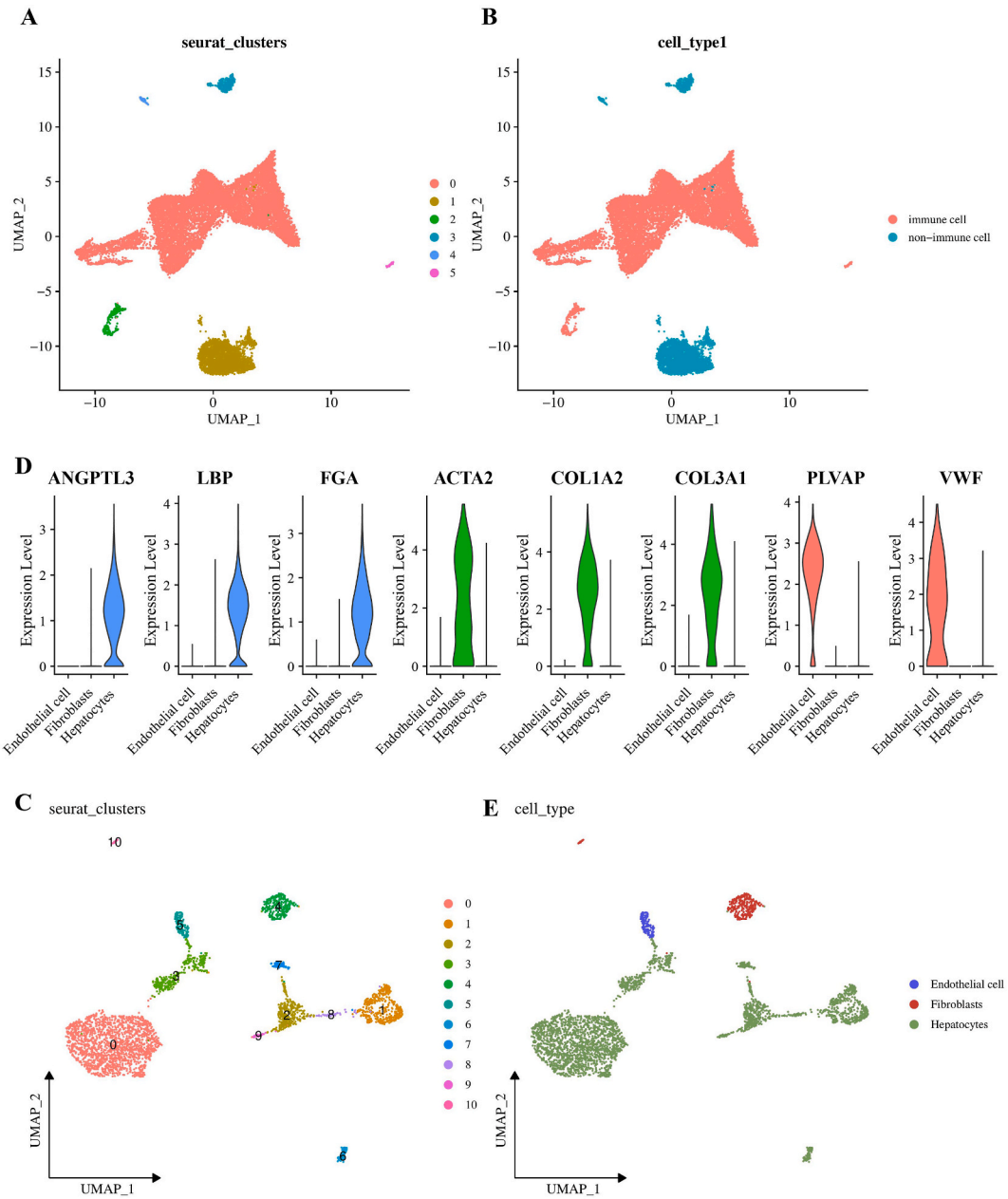


Fig. 1. Landscape of single cell RNA-sequencing of HILC. (A) UMAP plot of single cells clustering for 6 subsets. (B) Single cells clustering for immune cells and non-immune cells shown in UMAP plot. (C) Single cells clustering for non-immune cells shown in UMAP plot. (D) The violin plot of the marker genes expression in 3 non-immune cells. (E) Single cells clustering for 3 non-immune cells shown in UMAP plot.

determined after 2-h incubation at 37 °C utilizing a microplate reader (Thermo Fisher, USA).

2.10. Live/dead staining

The Live/Dead assay kit (L10119, Invitrogen, USA) was used to perform live/dead cell staining experiment. HepaRG and SK-HEP-1 (6×10^4 cells/well) treated with si-MMP9 were inoculated into 24-well plates. Staining solution prepared with Calcein-AM (2 μM) and ethidium homodimer-1 (4 μM) was added to the well for 30-min cell culture in an incubator with CO_2. Dead cells (red, Ex: 530 nm, Em: 645 nm) and live cells (green, Ex: 480 nm, Em: 530 nm) were photographed under an inverted fluorescence microscope (Olympus IMT-2/).

2.11. Transwell assay

The HCC cells, namely HepaRG and SK-HEP-1, were resuspended at a concentration of 2×10^4 cells/ml. Then, the cell suspension (200 μl) was added to the upper chamber of the Matrigel-coated serum-free transwell. As for the lower chamber, it contained 500 μl of complete medium that included 10% FBS. The Transwell system was incubated under the condition of 5% CO_2 for a duration of 24 h at 37 °C. Subsequently, the cells in the lower chamber were stained with a solution comprising 0.1% crystal violet for 15 min, and they were counted using an optical microscope.

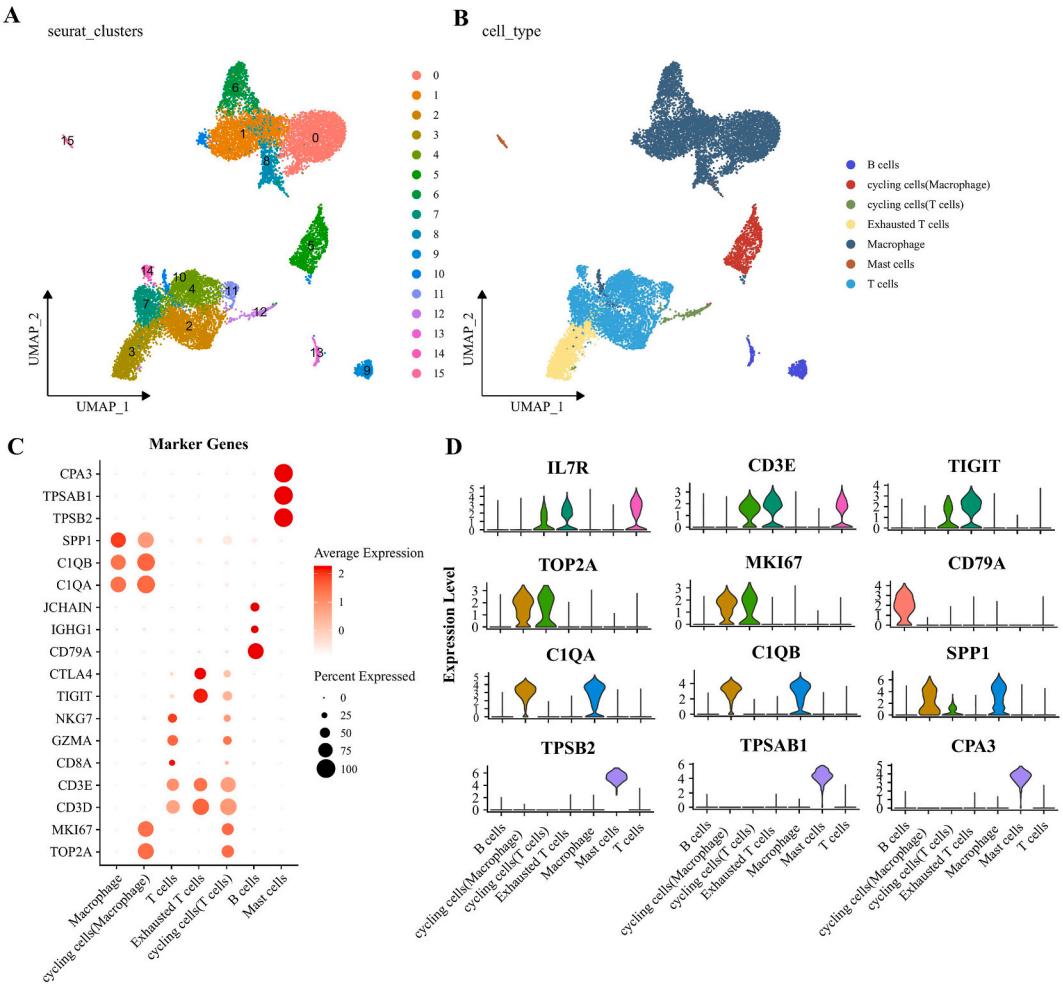


Fig. 2. *Identifying and definition of exhausted T cells.* (A) UMAP plot of single cells clustering for immune cells. (B) UMAP plot of single cells clustering for the annotated immune cells. (C) Bubble plot of the expression of marker genes in 7 immune cells. (D) The violin plot of the marker genes expression in 7 immune cells.

2.12. Statistical analysis

The R software (version 3.6.3) was used for statistical analysis and data visualization. SangerBox (http://sangerbox.com/home.html) provided online analyses. Significant difference between two sets of continuous variables was evaluated by the Wilcoxon rank-sum test [47]. The KM curves with log-rank test was used to display prognostic difference, and the Pearson method was used for the correlation analysis. *P*-value <0.05 was defined as a statistical significance.

3. Results

3.1. Landscape of single-cell RNA-seq of LIHC

After data processing, a total of 18,413 cells from the GSE166635 dataset were clustered into 7 subsets (Fig. 1A) and then classified into immune cell (macrophage, B cells, T cells and mast cells) and non-immune cell (hepatocytes, fibroblasts and endothelial cells) clusters using the singleR package (Fig. 1B). FindClusters function (resolution = 0.4) was used to further cluster non-immune cell clusters into 11 cell clusters (Fig. 1C). Based on the expression of marker genes (Fig. 1D), these cell clusters were annotated and defined as 3 cell types (Fig. 1E), namely, endothelia cells, Fibroblast cells and hepatocytes.

3.2. Identification and definition of exhausted T cells

Immune cell clusters were separated into 16 subsets by unsupervised clustering at resolution = 0.4 (Fig. 2A), but based on the

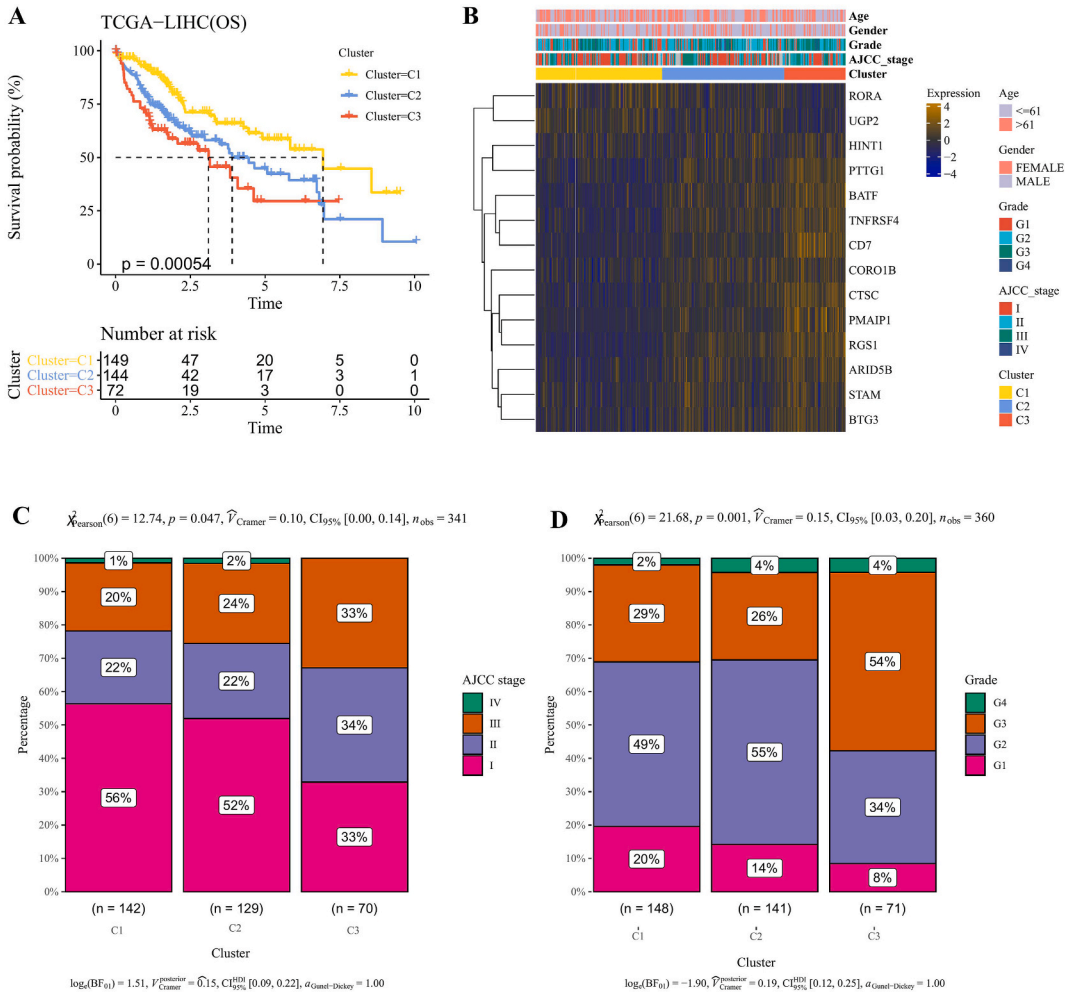


Fig. 3. Consistent clustering of TCGA-LIHC patients. (A) KM survival analysis of 3 clusters in TCGA-LIHC cohort that clustered by 14 exhausted T cells-related genes. (B) Heatmap of the 14 genes expression in 3 clusters. (C) The distribution of pathological feature of AJCC stage in 3 cluster. (D) The distribution of pathological feature of Grade in 3 cluster. (C1=Cluster1, C2=Cluster2 and C3=Cluster3).

expression of marker genes, these cell clusters were classified as 7 cell types (Fig. 2B), including B cells, two cycling cells, exhausted T cells, macrophage, mast cells and T cells. CTLA4 [48] and TIGIT [49] are classic exhaustive marker genes significantly expressed in exhausted T cells (Fig. 2C and D), indicating that the immune killing function of the two cells was impaired or suppressed in TME. Then, we used the FindAllMarkers function to select high-expressed genes in exhausted T cells and obtained 102 genes for further analysis. Univariate cox regression analysis revealed that 14 genes (defined as key genes) from the 102 high-expressed genes were closely related to LIHC prognosis.

3.3. Consistent clustering of TCGA-LIHC patients based on the key prognostic genes

Based on the 14 key prognostic genes, we performed consistent clustering of 365 LIHC patients in the TCGA-LIHC cohort. The Delta area curve of cumulative distribution function (CDF) indicated that the optimal number of clusters was 3 (K = 3), therefore patients

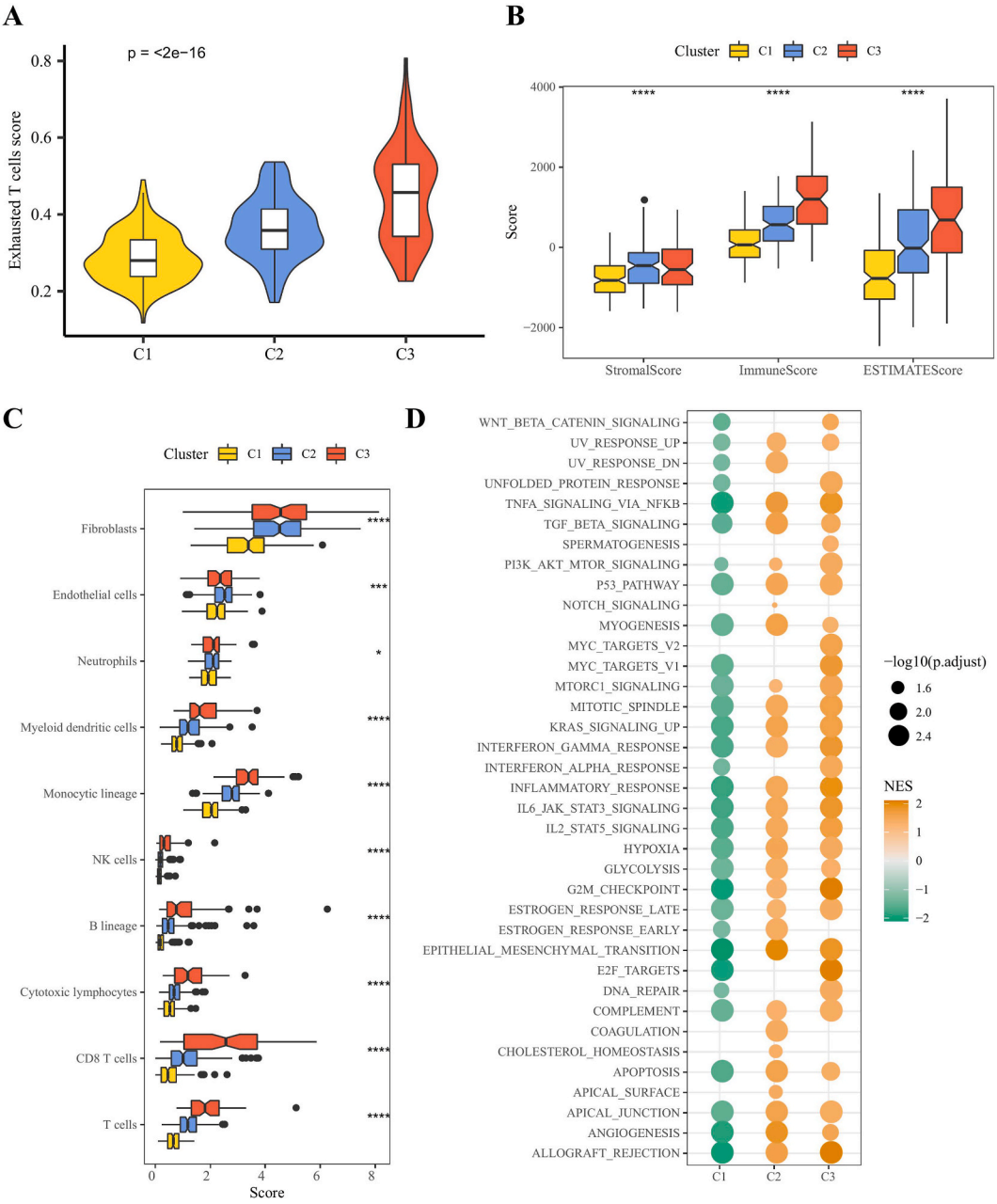


Fig. 4. *Immunity and pathway characteristics of 3 clusters.* (A) The distribution exhausted T cells score in 3 clusters. (B) ESTIMATE immune-infiltration analysis in 3 clusters. (C) MCP-Counter immune-infiltration analysis in 3 clusters. (D) The difference of GSEA score in 3 clusters.

were divided into 3 molecular subtypes (C1, C2 and C3). The KM survival analysis showed that the patients in C1 subtype had better prognosis, while the C3 patients exhibited the worst prognosis (Fig. 3A), moreover, most of these exhausted T cells-related genes were overexpressed in the C3 subtype (Fig. 3B). Cancer staging using American Joint Committee on Cancer (AJCC) and Grade also showed that the C3 subtype also had higher proportion of patients in advanced stages (Fig. 3C and D), suggesting that genes related to exhausted T cells could be an effective indicator to distinguish patients with LIHC.

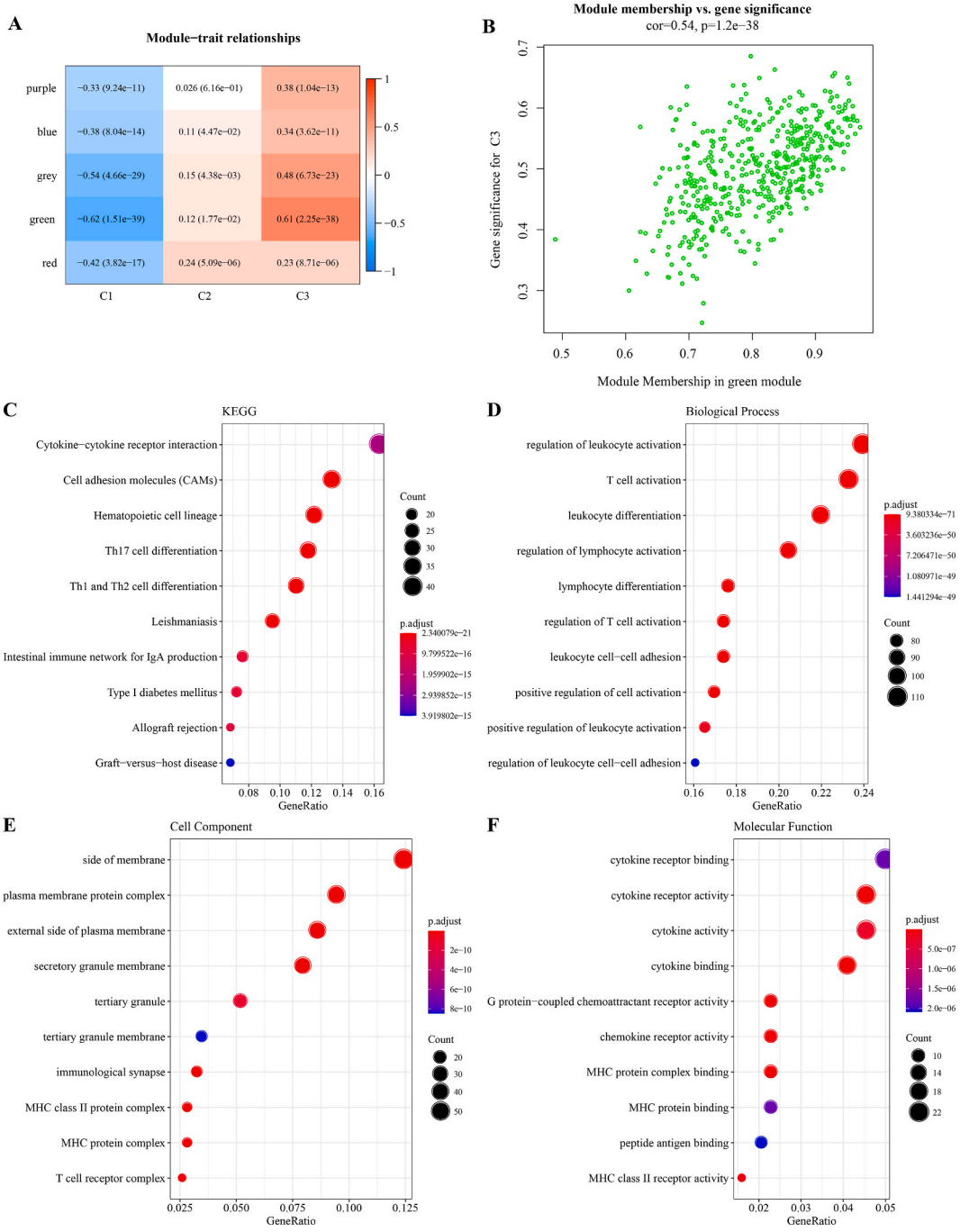
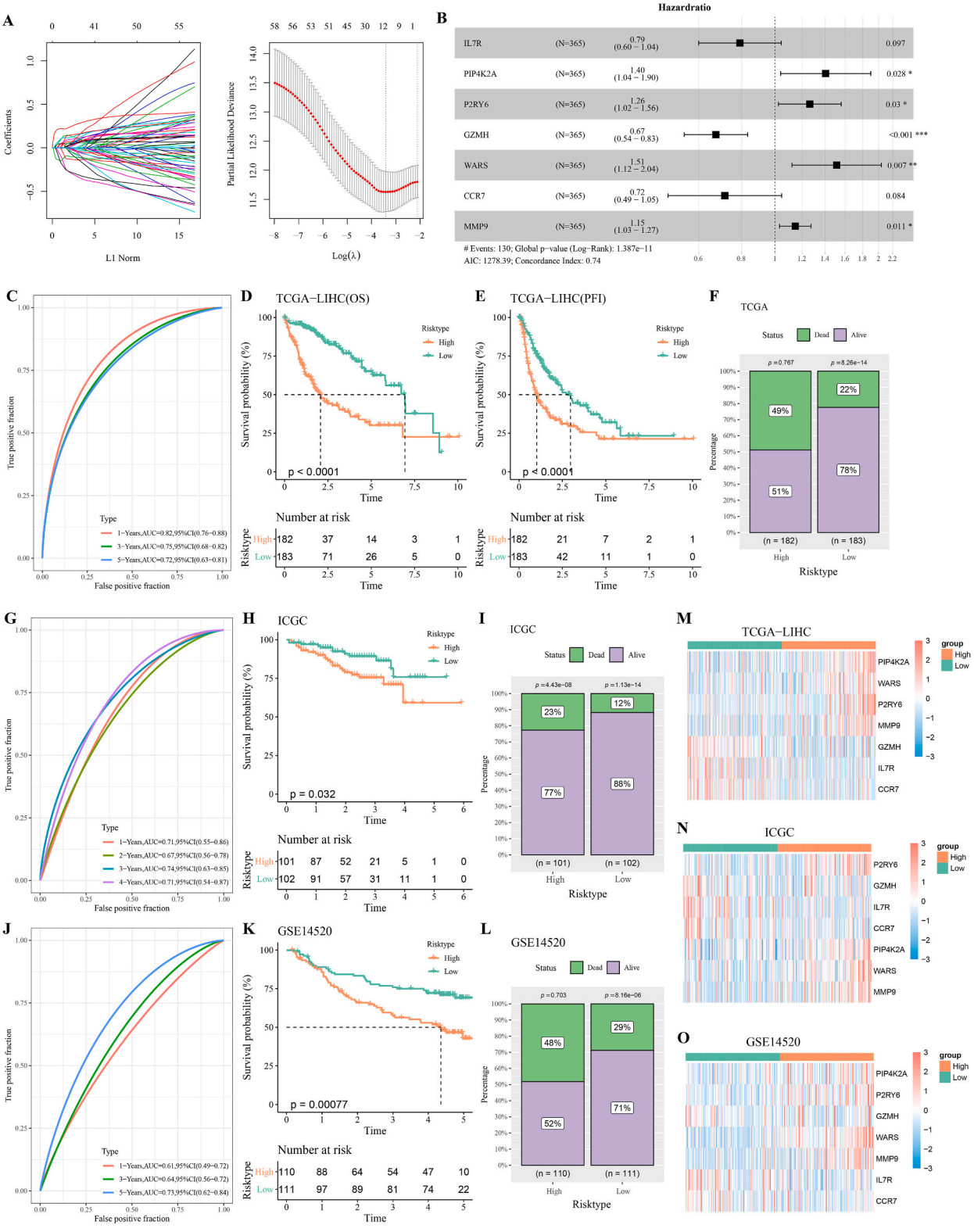


Fig. 5. WGCNA for the key exhausted T cells-related gene module. (A) The correlation between 3 clusters and the gene modules shown in heatmap. (B) Correlation of green module genes with the C3 cluster. (C) KEGG analysis of C3 cluster-related green module genes. (D) Biological process of GO analysis of C3 cluster-related green module genes. (E) Cell component of GO analysis of C3 cluster-related green module genes. (F) Molecular function of GO analysis of C3 cluster-related green module genes. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)



(caption on next page)

Fig. 6. Construction and validation the T cell exhaustion-related RiskScore model. (A) The trajectory of variables and lambda confidence interval analysis. (B) Multivariate cox regression determined 7 key prognostic genes. (C) ROC analysis of 1-, 3- and 5-year survival of RiskScore. (D) KM analysis of overall survival (OS) of low-risk and high-risk patients in the TCGA-LIHC cohort. (E) KM analysis of progression-free interval (PFI) of low-risk and high-risk patients in the TCGA-LIHC cohort. (F) Proportion of dead and alive in low-risk and high-risk patients. (G) ROC analysis of 1-, 3- and 5-year survival of RiskScore in ICGC cohort. (H) KM analysis of low-risk and high-risk patients in ICGC cohort. (I) Proportion of dead and alive in low-risk and high-risk patients in ICGC cohort. (J) ROC analysis of 1-, 3- and 5-year survival of RiskScore in GSE14520 cohort. (K) KM analysis of low-risk and high-risk patients in GSE14520 cohort. (L) Proportion of dead and alive in low-risk and high-risk patients in GSE14520 cohort. (M) The expression of 7 model genes in TCGA-LIHC cohort. (N) The expression of 7 model genes in ICGC cohort. (O) The expression of 7 model genes in GSE14520 cohort.

3.4. *Differences in immune-infiltration and pathway activation among the subtypes*

We calculated the exhausted T cell score for TCGA-LIHC patients using the GSVA package based on the expression of 102 high-expressed genes, and found that the C3 subtype had the highest score (Fig. 4A). The immune-infiltration analysis also revealed higher immune cell infiltration of the C3 subtype (Fig. 4B and C) and the worst outcome (Fig. 3A), indicating that long-term chronic inflammation suppressed the immune function of most immune cells and T cell exhaustion facilitated tumor progression. The results of

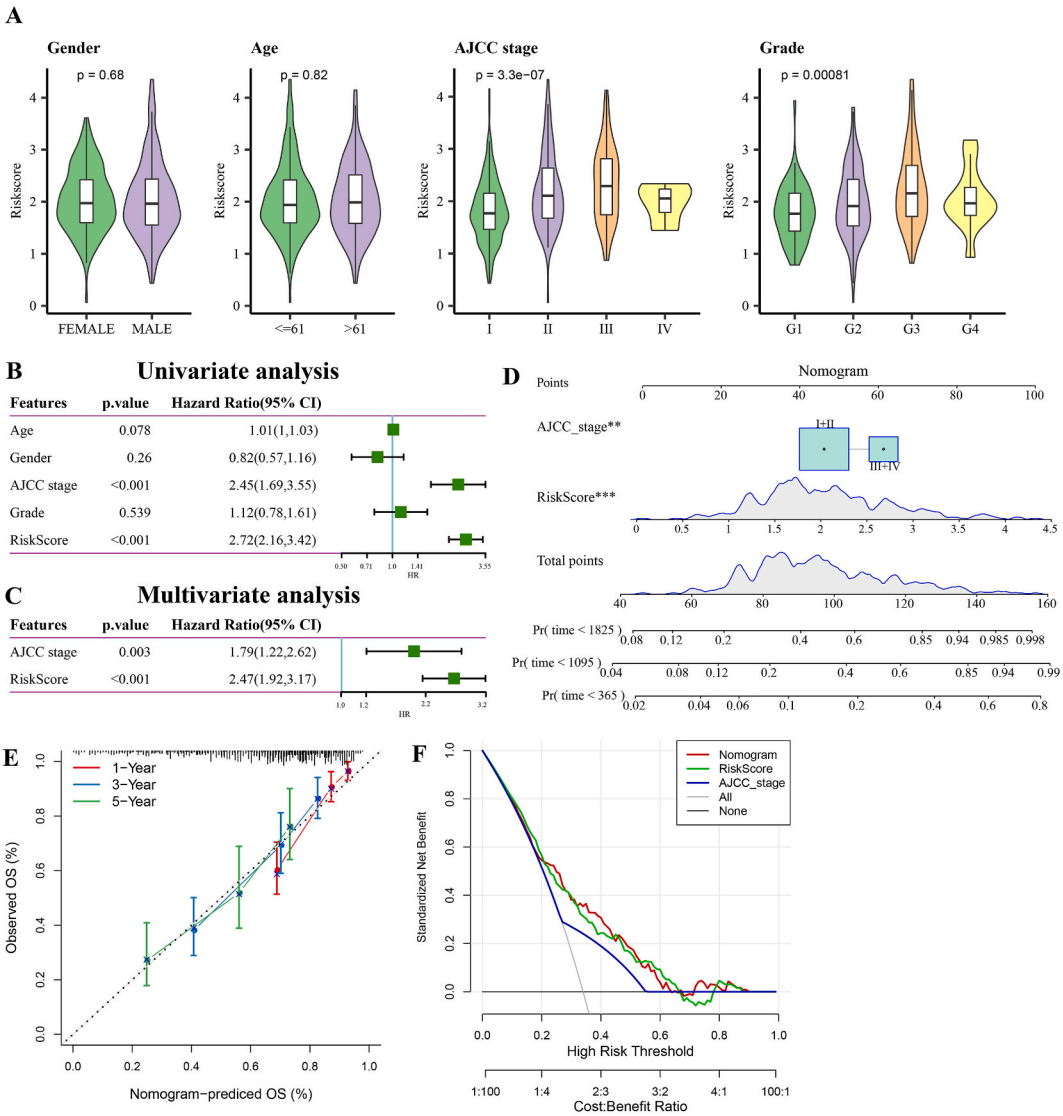


Fig. 7. Developing of independent factor and nomogram model. (A) The distribution of RiskScore in varying clinical classification. (B) Univariate analysis of RiskScore and other clinical features. (C) Multivariate analysis of RiskScore and AJCC stage. (D) A developing of nomogram model. (E) Calibration curve of nomogram. (F) Decision curve of nomogram and clinical features.

GSEA demonstrated that most oncogenic signaling pathways, such as the PI3K-AKT signaling, WNT/β-catenin signaling, Notch signaling, angiogenesis and IL6-JAK-STAT3 signaling pathway (Fig. 4D), were activated in the C2 and C3 subtypes, indicating that immune inflammation were inhibited and cancer-promoting inflammation was widely activated.

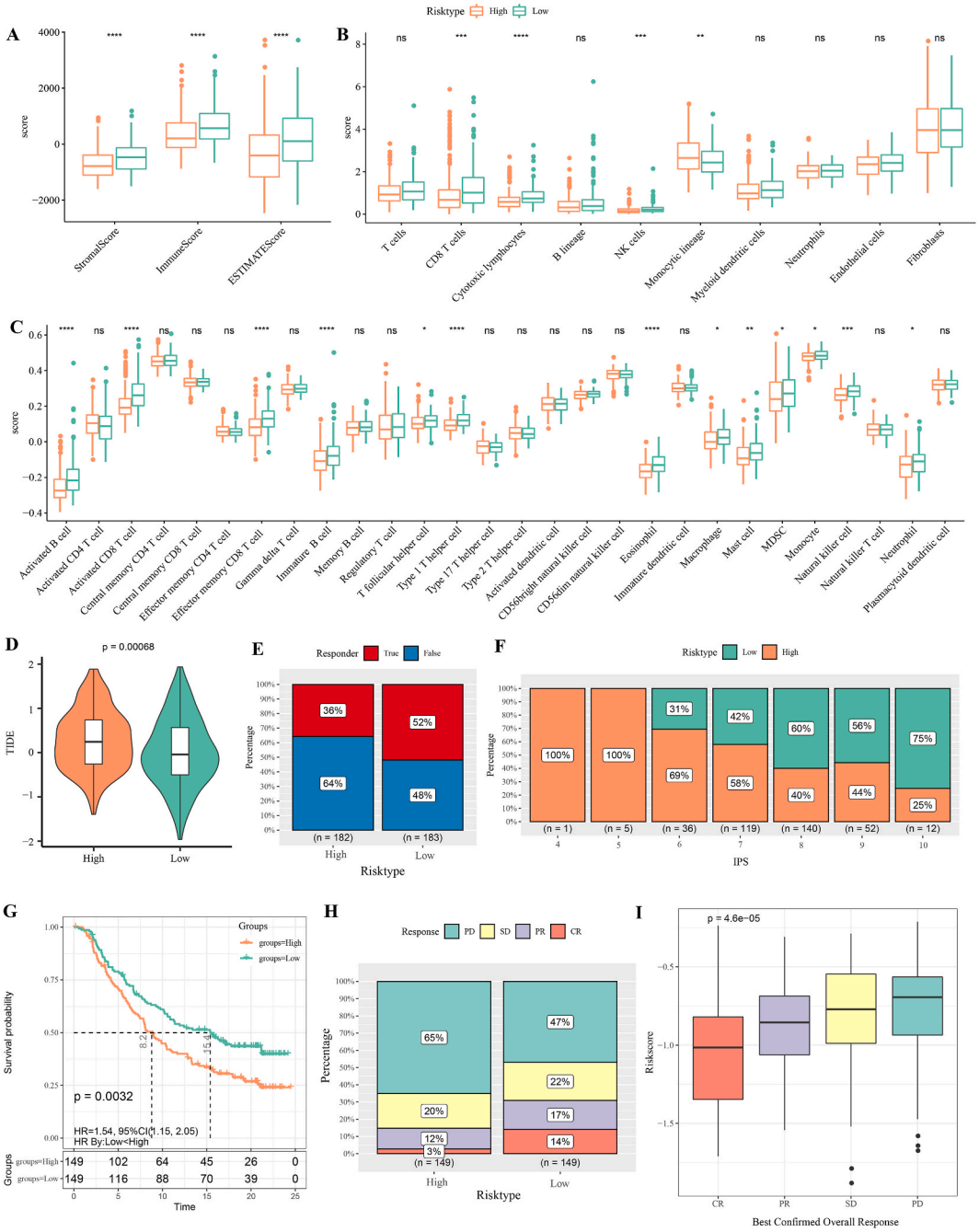


Fig. 8. *Immune-infiltration and immunotherapy response of varying risk patients*. (A) ESTIMATE immune-infiltration analysis of high- and low-risk patients. (B) MCPcounter immune-infiltration analysis of high- and low-risk patients. (C) ssGSEA score of high- and low-risk patients. (D) TIDE score of high- and low-risk patients. (E) The actually immune responder of high- and low-risk patients to immunotherapy. (F) IPS score analysis of varying risk patients in multiple sampling experiments. (G) KM analysis of high- and low-risk patients in IMvigor210 cohort. (H) The immunotherapy response ability of high- and low-risk patients in IMvigor210 cohort to the *anti*-PD-L1 receptor blockers. (I) The distribution of RiskScore in 4 immunotherapy response.

3.5. WGCNA for the key exhausted T cells-related gene module

The WGCNA was used to identify gene modules closely associated with different molecular subtype features. The top 50% genes with the greatest variation in the TCGA-LIHC cohort were selected and subjected to clustering analysis to determine co-expression module. The soft threshold β was 22 to ensure a scale-free nature of the network. Finally, 5 co-expression modules (the grey module was invalid) were obtained. The correlation analysis revealed that the green module exhibited significant correlation with C3 subtype (R = 0.61, Fig. 5A), and that the genes of green module also showed obvious correlation with C3 subtype (Fig. 5B). KEGG analysis indicated that these genes were mainly enriched in immune cell activation pathways, such as cytokine and receptor interaction, Th1 and Th2 cell differentiation (Fig. 5C). In biological process of GO, these genes were closely associated with the regulation of leukocyte activation, T cell activation and other lymphocytes activation (Fig. 5D). In cell component of GO, these genes were related to most of immune protein and complex (Fig. 5E). The molecular function of GO also revealed that these genes were involved in the activation of immune cell, such as MHC protein and peptide antigen binding, cytokine binding and receptor activity (Fig. 5F). The above results demonstrated that T cell exhaustion impaired immune response ability of T cells.

3.6. Construction and validation of a T cell exhaustion-related signature model for prognosis prediction

According to the trajectory of variables and lambda confidence interval analysis (Fig. 6A) and the multivariate cox regression (Fig. 6B), a total of 7 key prognostic genes were selected to establish a risk model: $\text{Riskscore} = (-0.234 * \text{IL7R}) + 0.339 * \text{PIP4K2A} + 0.233 * \text{P2RY6} + (-0.395 * \text{GZMH}) + 0.413 * \text{WARS} + (-0.335 * \text{CCR7}) + 0.136 * \text{MMP9}$. Based on the RiskScore of patients, patients in the TCGA-LIHC cohort were divided into high- and low-risk groups by the median value. The ROC analysis revealed that the RiskScore had excellent classification effectiveness, with an area under curve (AUC) [50] of 0.82, 0.75 and 0.72 for of 1-, 3- and 5-year survival, respectively (Fig. 6C). As shown in KM survival analysis, high-risk patients tended to have shorter overall survival (OS) and poor outcome (Fig. 6D). Patients with progression-free interval (PFI) in the high-risk group also exhibited poor outcome (Fig. 6E), and there were significantly more dead samples in the high-risk group (Fig. 6F). The same method was applied in the validation set (ICGC), the AUC for 1-, 3- and 5-year survival was 0.61, 0.64 and 0.73, respectively (Fig. 6G). High-risk patients had poor prognostic outcomes (Fig. 6H) and higher death rate (Fig. 6I). The GSE14520 dataset showed similar results to the ROC analysis (Fig. 6J), KM analysis (Fig. 6K) and proportion of dead samples (Fig. 6L). These results confirmed that our RiskScore was an accurate classifier and performed well in the survival evaluation for LIHC patients. It was also observed that the risk genes (PIP4K2A, P2RY6, WARS and MMP9) were high-expressed in the high-risk groups, while the rest protective genes were high-expressed in the low-risk groups (Fig. 6M–O). Here, the results were consistent with those from multivariate cox analysis (Fig. 6B).

3.7. Development of a nomogram model

Analysis on the distribution of RiskScore showed that patients with advanced stage had higher RiskScore (Fig. 7A). Univariate and multivariate cox regression analyses revealed that the AJCC stage and RiskScore were significant key prognostic factor ($p < 0.05$, Fig. 7B) and independent prognostic factor (Fig. 7C), respectively. In addition, to further improve the prediction accuracy, we established a nomogram model integrating the above two factors, and found that the RiskScore had the greatest influence on survival prediction (Fig. 7D). The calibration curve of the nomogram indicated that the 1-, 3- and 5-year curves were close to the standard ones (Fig. 7E), suggesting a high prediction accuracy of the nomogram. The decision curve analysis (DCA) showed that the net benefit of RiskScore and nomogram were higher than the extreme curves (Fig. 7F), indicating that the nomogram was reliable.

3.8. Immune-infiltration and immunotherapy sensitivity among the risk groups

The immune infiltration of different risk patients in the TCGA-LIHC cohort was measured to explore the TME feature. As shown by the results of the ESTIMATE analysis, the low-risk patients had higher stromal score, immune score and ESTIMATE score (Fig. 8A). The MCPcounter analysis demonstrated that CD8 T cells, cytotoxic lymphocytes and NK cells accounted for a higher proportion in the low-risk group (Fig. 8B). The ssGSEA results also showed that the low-risk patients had more effector memory CD8 T cells, activated B cells and CD4 T cells, immature B cells, eosinophils and natural killer cells (Fig. 8C), suggesting that the low-risk patients had more active immune response to tumor cells in their TME. Comparison on the TIDE score of different risk groups showed a higher TIDE score of high-risk patients (Fig. 8D), which indicated that those patients had greater possibility of immune escape and lower immunotherapy response. In clinical practice, the high-risk patients also exhibited less active immune response to immunotherapy (Fig. 8E). In multiple sampling analysis, higher IPS score was related to greater proportion of low-risk samples (Fig. 8F), which indicated a higher immunotherapy sensitivity of the low-risk group. We also incorporated an immunotherapy cohort (IMvigor210) of advanced urothelial carcinoma. The KM curve showed shorter time of patients with higher RiskScore (Fig. 8G) than those with higher PD proportion (Fig. 8H), while low-risk patients had better prognosis and higher CR proportion (Fig. 8G and H). Higher RiskScore also represented poorer response to immunotherapy (Fig. 8I).

3.9. Experimental validation of predictive models

The seven indicators (IL7R, PIP4K2A, P2RY6, GZMH, WARS, CCR7, and MMP9) were high-expressed in liver cancer cell lines than those in human liver cells, among them MMP9 with the highest differential expression was selected for further study (Fig. 9A–G). The

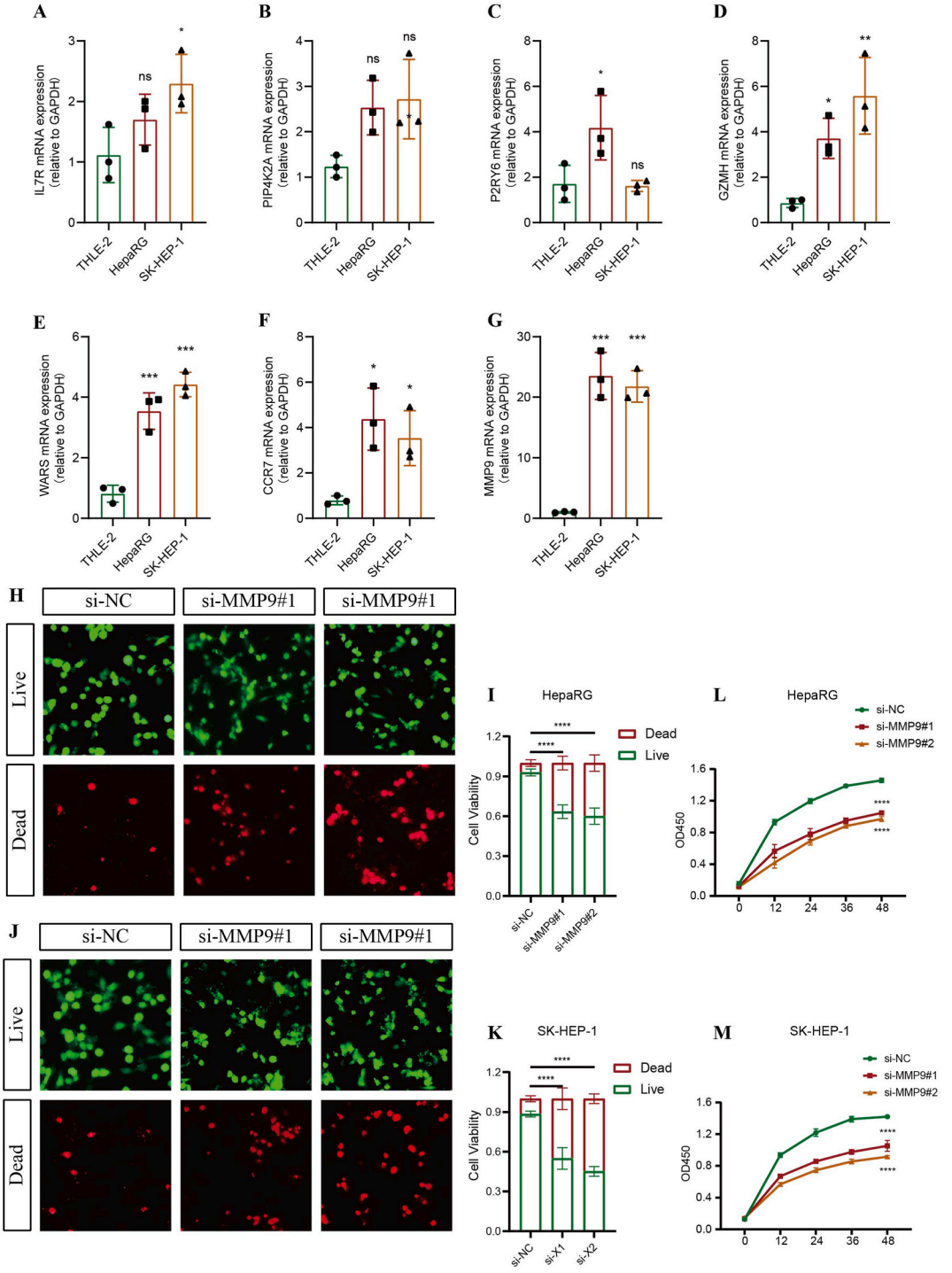


Fig. 9. *Experimental validation of predictive models.* (A–G) PCR was performed to detect the expression of IL7R, PIP4K2A, P2RY6, GZMH, WARS, CCR7, and MMP9 in THLE-2, HepaRG, and SK-HEP-1 cells. (H–K) The inhibitory efficiency of siRNA-MMP9 was verified in HepaRG, and SK-HEP-1 cell lines. (L–M) The CCK8 results of siRNA-MMP9 was verified in HepaRG, and SK-HEP-1 cell lines. n = 3, $p^{*} \leq 0.05$, $p^{**} \leq 0.01$, $p^{***} \leq 0.001$, $p^{****} \leq 0.0001$. The results are presented as mean ± SEM.

expression of MMP9 in HCC cell lines was inhibited using siRNA. The effect of inhibition of MMP9 on the viability of HCC cell lines was evaluated using live/dead cell staining. Inhibiting MMP9 promoted the death of HCC cells (Fig. 9H–K). The results of CCK8 further confirmed that inhibition of MMP9 could reduce cell viability (Fig. 9L and M). In addition, we evaluated the effect of MMP9 on HCC cells migration and invasion based on the transwell assay. The results showed that silencing MMP9 expression reduced the ability of HCC cells to migrate and invade (Fig. S1).

4. Discussion

LIHC is a frequent primary cancer that seriously threatens the life and health of patients [1,51]. Target therapy and immunotherapy have improved the survival outcome of LIHC patients. Tumor heterogeneity, postoperative recurrence, long-term chronic inflammation and immune tolerance feature of the liver as well as immune inhibition state and T cell exhaustion in TME of LIHC patients are all critical factors that decide the efficacy of immunotherapy [52]. Currently, we face a lack of effective biomarkers and prognostic models for guiding LIHC treatment [53]. To bridge such a gap, this study developed a RiskScore prognostic model based on genes related to T cell exhaustion and validated its predictive efficiency and reliability using varying cohorts. We found that our RiskScore model was accurate in predicting both short-term and long-term prognosis of LIHC and could be applied in clinical LIHC treatment.

T cell exhaustion is defined as a dysfunctional state of $CD8^+$ T cells under constant exposure to antigens or inflammatory signals [54]. In this state, T cells lose the ability to kill tumor cells and eliminate infection [55] and overexpresses suppressive receptor on cellular surface. However, blocking these receptors such as using *anti*-PD1 and *anti*-CTLA4 antibodies in immunotherapy can reactivate the immune-killing function of T cells and inhibit tumor progression [56,57], showing a great potential of exhausted T cells in immune checkpoint blockade therapies. Yang et al. examined surgical resection samples from patients with primary hepatocellular carcinoma for infiltration of exhausted T cells in the tumor and adjacent tissues and identified 10 clusters of cells with different distributions and characteristics by single-cell analysis. They found that the conversion from effector $CD8^+$ T cells to depleted $CD8^+$ T cells expressed both upregulated effector molecules and inhibitory receptors, suggesting altered expression of genes associated with the stress response and the cell cycle in the early stages of depletion [58]. Based on the expression of 14 key highly variable genes in exhausted T cells, patients with LIHC were divided into 3 clusters (C1, C2, and C3). Here, we found that the C3 cluster had poor prognosis, higher immune score, and many activated cancer-related pathways, indicating that a larger number of immune cells gathered in TME but most T cells were exhausted. This finding was consistent with previous discoveries that tumor-infiltrating T cells are often exhausted in different immune microenvironments, such as melanoma, lung and liver cancer, and that the exhaustion degree is closely correlated with prognosis [59,60]. A majority of T cell exhaustion-related module genes were involved in the activation of immune cells, indicating that the immune response sensitivity of T cell in immune microenvironments was impaired or inhibited, but this may be mediated by other complex mechanisms. Mounting clinical trials have shown that a large number of advanced LIHC patients could benefit from the *anti*-CTLA-4 and *anti*-PD-1 immune-based therapies [61]. The current findings demonstrated a higher immunotherapy sensitivity of low-risk patients, meaning that these therapies could be used to treat low-risk patients.

In addition, several key genes or transcription factors (such as NR4A and Tox) are involved in the epigenetics of T cell exhaustion [20], for example, overexpression of transcriptional repressor BLIMP1 in hypoxic microenvironment could promote T cell exhaustion phenotype [62]. Previous studies have provided potential molecular targets for gene editing or immunotherapy to suppress the expression of inhibitory receptors and reverse the exhausted state of T cells [63]. The current RiskScore model exhibited excellent prognostic performance, and the risk genes could be considered in the development of therapeutic LIHC targets. In our model, multivariate cox regression analysis defined IL7R, GZMH and CCR7 as protective factors and PIP4K2A, P2RY6, WARS and MMP9 as risk factors. Interleukin 7 (IL-7) is normally released by stromal cells under a stable state [64] and binds to interleukin 7 receptor (IL7R) to promote IL-7 response. This also means that a dynamic expression of IL7R is critical for the maturation and survival of T cells [65]. Granzyme H (GZMH) is regarded as an orphan granzyme that is constitutively expressed in NK cells at a high level to induce rapid apoptosis of tumor cells [66]. In addition, chemokines receptor CCR7 also acts as a tumor suppressor and is involved in initiating and orchestrating immune response via activating CCR7 axis [67]. Therapeutic strategies involving CCR7 axis can combat the metastasis and expansion of cancer [68]. PIP4K2A is a stress-regulated lipid kinase localized in cytoplasm and plasma membrane [69] that degrades p85 protein to negatively regulate the phosphorylation of 3-kinase (PI3K)-AKT signaling so as to inhibit tumor proliferation [70]. However, PIP4K2A could also act as a tumor promoter in tumor progression, and knockdown of PIP4K2A alone significantly increases the proliferation of cancer cells [71]. P2Y purinoceptor 6 (P2RY6) is high-expressed in many cancer tissues, which is associated with poor prognosis [72]. A high level of P2RY6 promotes the formation of a tumor-promoting microenvironment through enhancing resistance to apoptosis [73], while blocking Wnt/β-catenin and Hippo/YAP signaling pathways in P2RY6-deficient mice could inhibit skin tumorigenesis [74]. These findings indicate that P2RY6 is a positive regulator of cancers. Tryptophanyl-tRNA synthetase (WARS) is an enzyme that responds to tryptophan production [75] and is a cytokine involved in cancer proliferation and metastasis [76,77]. Several studies have revealed significantly higher level of WARS in tumor metastatic tissues than at primary sites [77,78], and that a high level of WARS is predictive of a poor survival [79]. Oncogenic matrix metalloproteinase-9 (MMP9) is an extracellular matrix (ECM) remodeling protein that acts as an indicator of a shorter survival in breast cancer [80]. Meanwhile, using targeted MMP9 hemopexin domain (PEX) could fully block neoplastic process (cancer cell migration and invasiveness) [81].

To conclude, we developed a reliable and effective RiskScore model based on T cell exhaustion to help evaluate the prognosis and immune infiltration and immunotherapy response of LIHC patients and provided theoretical foundation for precise and individual treatment. However, there were some limitations in this study. The sample size was relatively small and therefore we need more samples to further validate the present results. In addition, the robustness of current models and the specific functions of the model genes in LIHC should be verified by performing *in vivo* experiments.

5. Conclusion

This study integrated single-cell RNA-seq and bulk RNA-seq to analyze single-cell profiles and explore prognostic features together with T cell exhaustion. The RiskScore can achieve an accurate stratification of LIHC patients and effectively predict the prognosis, contributing to individualized treatment for LIHC patients.

Funding

This study was supported by Zhejiang Provincial Natural Science Foundation of China under Grant (No. LY21H160009), Zhejiang Province Traditional Chinese Medicine Science and Technology Project (No. 2024ZL1008), Basic medical and health science and technology project of Wenzhou Municipal Bureau of Science and Technology (No. Y20211119) and Wenzhou City Basic Scientific Research Project (No. Y2023896)

Data availability statement

The dataset analyzed in this study is available in [GSE149614] at [https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE149614].

Ethics approval and informed consent

None.

Consent for publication

None.

CRediT authorship contribution statement

Yu Zhou: Writing – original draft, Visualization, Supervision, Resources, Funding acquisition, Conceptualization. **Wanrui Wu:** Writing – review & editing, Visualization, Validation, Resources, Formal analysis. **Wei Cai:** Visualization, Software, Methodology, Investigation. **Dong Zhang:** Visualization, Project administration, Methodology. **Weiwei Zhang:** Visualization, Validation, Methodology, Investigation, Data curation. **Yunling Luo:** Visualization, Supervision, Methodology, Formal analysis. **Fujing Cai:** Visualization, Supervision, Software, Resources, Formal analysis. **Zhenjing Shi:** Writing – original draft, Supervision, Software, Methodology, Funding acquisition, Data curation, Conceptualization.

Declaration of competing interest

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

Acknowledgements

None.

Abbreviations

LIHC	Liver hepatocellular carcinoma
GEO	Gene Expression Omnibus
TCGA	the cancer genome atlas
HCCDB	Hepatocellular Carcinoma Cell DataBase
WGCNA	Weighted Gene Co-Expression Network Analysis operating characteristic curve (ROC)
LASSO	least absolute shrinkage and selection operator
KM	Kaplan-Meier
scRNA-seq	single-cell RNA sequencing
GSEA	Gene Set Enrichment Analysis
WHO	World Health Organization
KEGG	Kyoto Encyclopedia of Genes and Genomes
GO	Gene Ontology
FDA	Food and Drug Administration
TIDE	Tumor Immune Dysfunction and Exclusion
IPS	Immune cell Proportion Score

CR	complete response
PR	partial response
SD	stable disease
PD	progressive disease
TME	tumor microenvironment
MHC	major histocompatibility complex
TCR	T cell receptor
TNF	tumor necrosis factor
TEX	exhausted T cells
NGS	next-generation sequencing

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e28156.

References

[1] L. Yu, N. Shen, Y. Shi, X. Shi, X. Fu, S. Li, B. Zhu, W. Yu, Y. Zhang, Characterization of cancer-related fibroblasts (CAF) in hepatocellular carcinoma and construction of CAF-based risk signature based on single-cell RNA-seq and bulk RNA-seq data, Front. Immunol. 13 (2022) 1009789.
[2] F. Bray, J. Ferlay, I. Soerjomataram, R.L. Siegel, L.A. Torre, A. Jemal, Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA A Cancer J. Clin. 68 (6) (2018) 394–424.
[3] R.L. Siegel, K.D. Miller, H.E. Fuchs, A. Jemal, Cancer statistics, 2022, CA A Cancer J. Clin. 72 (1) (2022) 7–33.
[4] H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, F. Bray, Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA A Cancer J. Clin. 71 (3) (2021) 209–249.
[5] R.L. Siegel, K.D. Miller, N.S. Wagle, A. Jemal, Cancer statistics, 2023, CA A Cancer J. Clin. 73 (1) (2023) 17–48.
[6] R.L. Siegel, K.D. Miller, A. Jemal, Cancer statistics, 2020, CA A Cancer J. Clin. 70 (1) (2020) 7–30.
[7] Y. Jiang, A. Sun, Y. Zhao, W. Ying, H. Sun, X. Yang, B. Xing, W. Sun, L. Ren, B. Hu, C. Li, L. Zhang, G. Qin, M. Zhang, N. Chen, M. Zhang, Y. Huang, J. Zhou, Y. Zhao, M. Liu, X. Zhu, Y. Qiu, Y. Sun, C. Huang, M. Yan, M. Wang, W. Liu, F. Tian, H. Xu, J. Zhou, Z. Wu, T. Shi, W. Zhu, J. Qin, L. Xie, J. Fan, X. Qian, F. He, Proteomics identifies new therapeutic targets of early-stage hepatocellular carcinoma, Nature 567 (7747) (2019) 257–261.
[8] Y.-C. Chen, H.-H. Chen, P.-M. Chen, Catalase expression is an independent prognostic marker in liver hepatocellular carcinoma, Oncologie 26 (1) (2024) 79–90.
[9] J.M. Llovet, J. Zucman-Rossi, E. Pikarsky, B. Sangro, M. Schwartz, M. Sherman, G. Gores, Hepatocellular carcinoma, Nat. Rev. Dis. Prim. 2 (2016) 16018.
[10] P. Tabrizian, G. Jibara, B. Shrager, M. Schwartz, S. Roayaie, Recurrence of hepatocellular cancer after resection: patterns, treatments, and prognosis, Ann. Surg. 261 (5) (2015) 947–955.
[11] J. Wang, Y. Zhang, L. Sun, Y. Liu, IMMT promotes hepatocellular carcinoma formation via PI3K/AKT/mTOR pathway, Oncologie 25 (6) (2023) 691–703.
[12] H.B. El-Serag, Hepatocellular carcinoma, N. Engl. J. Med. 365 (12) (2011) 1118–1127.
[13] K. Han, J.H. Kim, Transarterial chemoembolization in hepatocellular carcinoma treatment: Barcelona clinic liver cancer staging system, World J. Gastroenterol. 21 (36) (2015) 10327–10335.
[14] C. Yan, Y. Niu, L. Ma, L. Tian, J. Ma, System analysis based on the cuproptosis-related genes identifies LIPT1 as a novel therapy target for liver hepatocellular carcinoma, J. Transl. Med. 20 (1) (2022) 452.
[15] A. Huang, X.R. Yang, W.Y. Chung, A.R. Dennison, J. Zhou, Targeted therapy for hepatocellular carcinoma, Signal Transduct. Targeted Ther. 5 (1) (2020) 146.
[16] B. Arneth, Tumor microenvironment, Medicina 56 (1) (2019).
[17] F.R. Greten, S.I. Grivennikov, Inflammation and cancer: triggers, mechanisms, and consequences, Immunity 51 (1) (2019) 27–41.
[18] D.A. Spandidos, Oncogenes and tumor suppressor genes as paradigms in oncogenesis, J. B.U.ON. 12 (Suppl 1) (2007) S9–S12.
[19] J.A. Belk, B. Daniel, A.T. Satpathy, Epigenetic regulation of T cell exhaustion, Nat. Immunol. 23 (6) (2022) 848–860.
[20] H. Seo, J. Chen, E. González-Avalos, D. Samaniego-Castruita, A. Das, Y.H. Wang, I.F. López-Moyado, R.O. Georges, W. Zhang, A. Onodera, C.J. Wu, L.F. Lu, P. G. Hogan, A. Bhandoola, A. Rao, TOX and TOX2 transcription factors cooperate with NR4A transcription factors to impose CD8(+) T cell exhaustion, Proc. Natl. Acad. Sci. U.S.A. 116 (25) (2019) 12410–12415.
[21] M. Philip, A. Schietinger, CD8(+) T cell differentiation and dysfunction in cancer, Nat. Rev. Immunol. 22 (4) (2022) 209–223.
[22] E.J. Wherry, M. Kurachi, Molecular and cellular insights into T cell exhaustion, Nat. Rev. Immunol. 15 (8) (2015) 486–499.
[23] B.C. Miller, D.R. Sen, R. Al Abosy, K. Bi, Y.V. Virkud, M.W. LaFleur, K.B. Yates, A. Lako, K. Felt, G.S. Naik, M. Manos, E. Gjini, J.R. Kuchroo, J.J. Ishizuka, J. L. Collier, G.K. Griffin, S. Maleri, D.E. Comstock, S.A. Weiss, F.D. Brown, A. Panda, M.D. Zimmer, R.T. Manguso, F.S. Hodi, S.J. Rodig, A.H. Sharpe, W. N. Haining, Subsets of exhausted CD8(+) T cells differentially mediate tumor control and respond to checkpoint blockade, Nat. Immunol. 20 (3) (2019) 326–336.
[24] Y. Zhang, L. Li, W. Zheng, L. Zhang, N. Yao, CD8(+) T-cell exhaustion in the tumor microenvironment of head and neck squamous cell carcinoma determines poor prognosis, Ann. Transl. Med. 10 (6) (2022) 273.
[25] Q. Zhu, Y. Yang, X. Deng, N. Chao, Z. Chen, Y. Ye, W. Zhang, W. Liu, S. Zhao, High CD8(+)tumor-infiltrating lymphocytes indicate severe exhaustion and poor prognosis in angioimmunoblastic T-cell lymphoma, Front. Immunol. 14 (2023) 1228004.
[26] H. Chi, S. Zhao, J. Yang, X. Gao, G. Peng, J. Zhang, X. Xie, G. Song, K. Xu, Z. Xia, S. Chen, J. Zhao, T-cell exhaustion signatures characterize the immune landscape and predict HCC prognosis via integrating single-cell RNA-seq and bulk RNA-sequencing, Front. Immunol. 14 (2023) 1137025.
[27] A. Chow, K. Perica, C.A. Klebanoff, J.D. Wolchok, Clinical implications of T cell exhaustion for cancer immunotherapy, Nat. Rev. Clin. Oncol. 19 (12) (2022) 775–790.
[28] T.K. Olsen, N. Baryawno, Introduction to single-cell RNA sequencing, Curr. Protoc. Mol. Biol. 122 (1) (2018) e57.
[29] M. Su, T. Pan, Q.Z. Chen, W.W. Zhou, Y. Gong, G. Xu, H.Y. Yan, S. Li, Q.Z. Shi, Y. Zhang, X. He, C.J. Jiang, S.C. Fan, X. Li, M.J. Cairns, X. Wang, Y.S. Li, Data analysis guidelines for single-cell RNA-seq in biomedical studies and clinical applications, Milit. Med. Res. 9 (1) (2022) 68.
[30] X. Chang, J. Wang, C. Ni, Value of Mir-1271 and GPC3 in prognosis evaluation of liver cancer patients after liver transarterial chemoembolization, Oncologie 23 (1) (2021) 119–130.
[31] A. Zulibiya, J. Wen, H. Yu, X. Chen, L. Xu, X. Ma, B. Zhang, Single-cell RNA sequencing reveals potential for endothelial-to-mesenchymal transition in tetralogy of fallot, Congenit. Heart Dis. 18 (6) (2023) 611–625.
[32] J. Lu, Y. Chen, X. Zhang, J. Guo, K. Xu, L. Li, A novel prognostic model based on single-cell RNA sequencing data for hepatocellular carcinoma, Cancer Cell Int. 22 (1) (2022) 38.

[33] Q. Lian, S. Wang, G. Zhang, D. Wang, G. Luo, J. Tang, L. Chen, J. Gu, HCCDB: a database of hepatocellular carcinoma expression atlas, Dev. Reprod. Biol. 16 (4) (2018) 269–275.
[34] E. Clough, T. Barrett, The gene expression Omnibus database, Methods Mol. Biol. 1418 (2016) 93–110.
[35] G.M. Liu, H.D. Zeng, C.Y. Zhang, J.W. Xu, Identification of a six-gene signature predicting overall survival for hepatocellular carcinoma, Cancer Cell Int. 19 (2019) 138.
[36] Z. Tan, X. Chen, J. Zuo, S. Fu, H. Wang, J. Wang, Comprehensive analysis of scRNA-Seq and bulk RNA-Seq reveals dynamic changes in the tumor immune microenvironment of bladder cancer and establishes a prognostic model, J. Transl. Med. 21 (1) (2023) 223.
[37] N. Chen, B. Fan, Z. He, X. Yu, J. Wang, Identification of HBEGF+ fibroblasts in the remission of rheumatoid arthritis by integrating single-cell RNA sequencing datasets and bulk RNA sequencing datasets, Arthritis Res. Ther. 24 (1) (2022) 215.
[38] D. Aran, A.P. Looney, L. Liu, E. Wu, V. Fong, A. Hsu, S. Chak, R.P. Naikawadi, P.J. Wolters, A.R. Abate, A.J. Butte, M. Bhattacharya, Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage, Nat. Immunol. 20 (2) (2019) 163–172.
[39] M. Li, S. Xin, R. Gu, L. Zheng, J. Hu, R. Zhang, H. Dong, Novel diagnostic biomarkers related to oxidative stress and macrophage ferroptosis in atherosclerosis, Oxid. Med. Cell. Longev. 2022 (2022) 8917947.
[40] X. Hu, S. Ni, K. Zhao, J. Qian, Y. Duan, Bioinformatics-led discovery of osteoarthritis biomarkers and inflammatory infiltrates, Front. Immunol. 13 (2022) 871008.
[41] L. Liang, J. Yu, J. Li, N. Li, J. Liu, L. Xiu, J. Zeng, T. Wang, L. Wu, Integration of scRNA-seq and bulk RNA-seq to analyse the heterogeneity of ovarian cancer immune cells and establish a molecular risk model, Front. Oncol. 11 (2021) 711020.
[42] T. Wang, L. Dai, S. Shen, Y. Yang, M. Yang, X. Yang, Y. Qiu, W. Wang, Comprehensive molecular analyses of a macrophage-related gene signature with regard to prognosis, immune features, and biomarkers for immunotherapy in hepatocellular carcinoma based on WGCNA and the LASSO algorithm, Front. Immunol. 13 (2022) 843408.
[43] S. Chen, Y. Zhang, X. Ding, W. Li, Identification of lncRNA/circRNA-miRNA-mRNA ceRNA network as biomarkers for hepatocellular carcinoma, Front. Genet. 13 (2022) 838869.
[44] Y. He, J. Zhang, Z. Chen, K. Sun, X. Wu, J. Wu, L. Sheng, A seven-gene prognosis model to predict biochemical recurrence for prostate cancer based on the TCGA database, Front. Surg. 9 (2022) 923473.
[45] J. Wang, Z. Kang, Y. Liu, Z. Li, Y. Liu, J. Liu, Identification of immune cell infiltration and diagnostic biomarkers in unstable atherosclerotic plaques by integrated bioinformatics analysis and machine learning, Front. Immunol. 13 (2022) 956078.
[46] A.L. Cheng, S. Qin, M. Ikeda, P.R. Galle, M. Ducreux, T.Y. Kim, H.Y. Lim, M. Kudo, V. Breder, P. Merle, A.O. Kaseb, D. Li, W. Verret, N. Ma, A. Nicholas, Y. Wang, L. Li, A.X. Zhu, R.S. Finn, Updated efficacy and safety data from IMbrave150: atezolizumab plus bevacizumab vs. sorafenib for unresectable hepatocellular carcinoma, J. Hepatol. 76 (4) (2022) 862–873.
[47] P. Kourkoveli, S. Rammos, J. Parissis, A. Maillis, D. Kremastinos, I. Paraskevaidis, Depressive symptoms in patients with congenital heart disease: incidence and prognostic value of self-rating depression scales, Congenit. Heart Dis. 10 (3) (2015) 240–247.
[48] C.S. Pai, D.M. Simons, X. Lu, M. Evans, J. Wei, Y.H. Wang, M. Chen, J. Huang, C. Park, A. Chang, J. Wang, S. Westmoreland, C. Beam, D. Banach, D. Bowley, F. Dong, J. Seagal, W. Ritacco, P.L. Richardson, S. Mitra, G. Lynch, P. Bousquet, J. Mankovich, G. Kingsbury, L. Fong, Tumor-conditional anti-CTLA4 uncouples antitumor efficacy from immunotherapy-related toxicity, J. Clin. Investig. 129 (1) (2019) 349–363.
[49] Liu L, You X, Han S, Sun Y, Zhang J, Zhang Y. CD155/TIGIT, a novel immune checkpoint in human cancers (Review). Oncol Rep. 2021 Mar;45(3):835-845. doi: 10.3892/or.2021.7943. Epub 2021 Jan 19. PMID: 33469677.
[50] S.A. Aly, D. Zurakowski, P. Glass, K. Skurow-Todd, R.A. Jonas, M.T. Donofrio, Cerebral tissue oxygenation index and lactate at 24 hours postoperative predict survival and neurodevelopmental outcome after neonatal cardiac surgery, Congenit. Heart Dis. 12 (2) (2017) 188–195.
[51] S. Yu, H. Lv, J. Zhang, H. Zhang, W. Ju, Y. Jiang, L. Lin, Heparanase/Syndecan-1 Axis regulates the grade of liver cancer and proliferative ability of hepatocellular carcinoma cells, Oncologie 24 (3) (2022) 539–551.
[52] B. Sangro, P. Sarobe, S. Hervás-Stubbs, I. Melero, Advances in immunotherapy for hepatocellular carcinoma, Nat. Rev. Gastroenterol. Hepatol. 18 (8) (2021) 525–543.
[53] F. Piñero, M. Dirchwolf, M.G. Pessôa, Biomarkers in hepatocellular carcinoma: diagnosis, prognosis and treatment response assessment, Cells 9 (6) (2020).
[54] E.J. Wherry, T cell exhaustion, Nat. Immunol. 12 (6) (2011) 492–499.
[55] L.T. Nguyen, P.S. Ohashi, Clinical blockade of PD1 and LAG3–potential mechanisms of action, Nat. Rev. Immunol. 15 (1) (2015) 45–56.
[56] K.E. Pauken, E.J. Wherry, Overcoming T cell exhaustion in infection and cancer, Trends Immunol. 36 (4) (2015) 265–276.
[57] M. Hashimoto, A.O. Kamphorst, S.J. Im, H.T. Kissick, R.N. Pillai, S.S. Ramalingam, K. Araki, R. Ahmed, CD8 T cell exhaustion in chronic infection and cancer: opportunities for interventions, Annu. Rev. Med. 69 (2018) 301–318.
[58] Y. Yang, F. Liu, W. Liu, M. Ma, J. Gao, Y. Lu, L.H. Huang, X. Li, Y. Shi, X. Wang, D. Wu, Analysis of single-cell RNAseq identifies transitional states of T cells associated with hepatocellular carcinoma, Clin. Transl. Med. 10 (3) (2020) e133.
[59] C. Zheng, L. Zheng, J.K. Yoo, H. Guo, Y. Zhang, X. Guo, B. Kang, R. Hu, J.Y. Huang, Q. Zhang, Z. Liu, M. Dong, X. Hu, W. Ouyang, J. Peng, Z. Zhang, Landscape of infiltrating T cells in liver cancer revealed by single-cell sequencing, Cell 169 (7) (2017) 1342–1356.e1316.
[60] X. Guo, Y. Zhang, L. Zheng, C. Zheng, J. Song, Q. Zhang, B. Kang, Z. Liu, L. Jin, R. Xing, R. Gao, L. Zhang, M. Dong, X. Hu, X. Ren, D. Kirchhoff, H.G. Roider, T. Yan, Z. Zhang, Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing, Nat. Med. 24 (7) (2018) 978–985.
[61] M. Hilmi, C. Neuzillet, J. Calderaro, F. Lafdil, J.M. Pawlotsky, B. Rousseau, Angiogenesis and immune checkpoint inhibitors as therapies for hepatocellular carcinoma: current knowledge and future research directions, J. Immunother. Cancer 7 (1) (2019) 333.
[62] I.Y. Jung, V. Narayan, S. McDonald, A.J. Rech, R. Bartoszek, G. Hong, M.M. Davis, J. Xu, A.C. Boesteanu, J.S. Barber-Rotenberg, G. Plesa, S.F. Lacey, J. K. Jadlowsky, D.L. Siegel, D.M. Hammill, P.F. Cho-Park, S.L. Berger, N.B. Haas, J.A. Fraietta, BLIMP1 and NR4A3 transcription factors reciprocally regulate antitumor CAR T cell stemness and exhaustion, Sci. Transl. Med. 14 (670) (2022) eabn7336.
[63] N. Budimir, G.D. Thomas, J.S. Dolina, S. Salek-Ardakani, Reversing T-cell exhaustion in cancer: lessons learned from PD-1/PD-L1 immune checkpoint blockade, Cancer Immunol. Res. 10 (2) (2022) 146–153.
[64] D.T. Patton, A.W. Plumb, N. Abraham, The survival and differentiation of pro-B and pre-B cells in the bone marrow is dependent on IL-7Rα Tyr449, J. Immunol. 193 (7) (2014) 3446–3455.
[65] C. Hong, M.A. Luckey, J.H. Park, Intrathymic IL-7: the where, when, and why of IL-7 signaling during T cell development, Semin. Immunol. 24 (3) (2012) 151–158.
[66] Q. Hou, T. Zhao, H. Zhang, H. Lu, Q. Zhang, L. Sun, Z. Fan, Granzyme H induces apoptosis of target tumor cells characterized by DNA fragmentation and Bid-dependent mitochondrial damage, Mol. Immunol. 45 (4) (2008) 1044–1055.
[67] R. Förster, A.C. Davalos-Misslitz, A. Rot, CCR7 and its ligands: balancing immunity and tolerance, Nat. Rev. Immunol. 8 (5) (2008) 362–371.
[68] A. Salem, M. Alotaibi, R. Mroueh, H.A. Basheer, K. Afarinkia, CCR7 as a therapeutic target in Cancer, Biochim. Biophys. Acta, Rev. Cancer 1875 (1) (2021) 188499.
[69] R. Fiume, Y. Stijf-Bultsma, Z.H. Shah, W.J. Keune, D.R. Jones, J.G. Jude, N. Divecha, PIP4K and the role of nuclear phosphoinositides in tumour suppression, Biochim. Biophys. Acta 1851 (6) (2015) 898–910.
[70] Y.J. Shin, J.K. Sa, Y. Lee, D. Kim, N. Chang, H.J. Cho, M. Son, M.Y.T. Oh, K. Shin, J.K. Lee, J. Park, Y.K. Jo, M. Kim, P.J. Paddison, V. Tergaonkar, J. Lee, D. H. Nam, PIP4K2A as a negative regulator of PI3K in PTEN-deficient glioblastoma, J. Exp. Med. 216 (5) (2019) 1120–1134.
[71] J.G. Jude, G.J. Spencer, X. Huang, T.D.D. Somerville, D.R. Jones, N. Divecha, T.C.P. Somervaille, A targeted knockdown screen of genes coding for phosphoinositide modulators identifies PIP4K2A as required for acute myeloid leukemia cell proliferation and survival, Oncogene 34 (10) (2015) 1253–1262.
[72] H. Wu, X. Dong, Immunological role and clinical prognostic significance of P2RY6 in lung adenocarcinoma: a multi-omics studies and single-cell sequencing analysis, World J. Surg. Oncol. 21 (1) (2023) 341.

[73] M. Placet, G. Arguin, C.M. Molle, J.P. Babeu, C. Jones, J.C. Carrier, B. Robaye, S. Geha, F. Boudreau, F.P. Gendron, The G protein-coupled P2Y$_6$ receptor promotes colorectal cancer tumorigenesis by inhibiting apoptosis, Biochim. Biophys. Acta, Mol. Basis Dis. 1864 (5 Pt A) (2018) 1539–1551.
[74] P. Xu, C. Wang, W. Xiang, Y. Liang, Y. Li, X. Zhang, C. Guo, M. Liu, Y. Shi, X. Ye, Y. Dang, P2RY6 has a critical role in mouse skin carcinogenesis by regulating the YAP and β-catenin signaling pathways, J. Invest. Dermatol. 142 (9) (2022) 2334–2342.e2338.
[75] E.L. Paley, L. Smelyanski, V. Malinovskii, P.R. Subbarayan, Y. Berdichevsky, N. Posternak, J.M. Gershoni, O. Sokolova, G. Denisova, Mapping and molecular characterization of novel monoclonal antibodies to conformational epitopes on NH2 and COOH termini of mammalian tryptophanyl-tRNA synthetase reveal link of the epitopes to aggregation and Alzheimer's disease, Mol. Immunol. 44 (4) (2007) 541–557.
[76] A. Morita, E. Miyagi, H. Yasumitsu, H. Kawasaki, H. Hirano, F. Hirahara, Proteomic search for potential diagnostic markers and therapeutic targets for ovarian clear cell adenocarcinoma, Proteomics 6 (21) (2006) 5880–5890.
[77] C.W. Lee, K.P. Chang, Y.Y. Chen, Y. Liang, C. Hsueh, J.S. Yu, Y.S. Chang, C.J. Yu, Overexpressed tryptophanyl-tRNA synthetase, an angiostatic protein, enhances oral cancer cell invasiveness, Oncotarget 6 (26) (2015) 21979–21992.
[78] L.M. Chi, C.W. Lee, K.P. Chang, S.P. Hao, H.M. Lee, Y. Liang, C. Hsueh, C.J. Yu, I.N. Lee, Y.J. Chang, S.Y. Lee, Y.M. Yeh, Y.S. Chang, K.Y. Chien, J.S. Yu, Enhanced interferon signaling pathway in oral cancer revealed by quantitative proteome analysis of microdissected specimens using 16O/18O labeling and integrated two-dimensional LC-ESI-MALDI tandem MS, Mol. Cell. Proteomics : MCP 8 (7) (2009) 1453–1474.
[79] P.P. Yang, X.H. Yu, J. Zhou, Tryptophanyl-tRNA synthetase (WARS) expression in uveal melanoma - possible contributor during uveal melanoma progression, Biosci. Biotechnol. Biochem. 84 (3) (2020) 471–480.
[80] C. Joseph, M. Alsaleem, N. Orah, P.L. Narasimha, I.M. Miligy, S. Kurozumi, I.O. Ellis, N.P. Mongan, A.R. Green, E.A. Rakha, Elevated MMP9 expression in breast cancer is a predictor of shorter patient survival, Breast Cancer Res. Treat. 182 (2) (2020) 267–282.
[81] G. Yosef, H. Hayun, N. Papo, Simultaneous targeting of CD44 and MMP9 catalytic and hemopexin domains as a therapeutic strategy, Biochem. J. 478 (5) (2021) 1139–1157.