Interleukin-2 inducible T-cell kinase: a potential prognostic biomarker and tumor microenvironment remodeling indicator for hepatocellular carcinoma

Binhua Pan[1,3,4,*], Modan Yang[1,3,4,*], Xuyong Wei[1,4], Wangyao Li[1,3,4], Kun Wang[1,3,4], Mengfan Yang[1,3,4], Di Lu[1,4], Rui Wang[1,4], Beini Cen[1,4], Xiao Xu[1,2,3,4,5]

[1]Department of Hepatobiliary and Pancreatic Surgery, The Center for Integrated Oncology and Precision Medicine, Affiliated Hangzhou First People's Hospital, Zhejiang University School of Medicine, Hangzhou 310006, China
[2]Zhejiang University Cancer Center, Hangzhou 310058, China
[3]Department of Hepatobiliary and Pancreatic Surgery, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310003, China
[4]NHC Key Laboratory of Combined Multi-Organ Transplantation, Hangzhou 310003, China
[5]Institute of Organ Transplantation, Zhejiang University, Hangzhou 310003, China
[*]Equal contribution

Correspondence to: Xiao Xu; **email:** zjxu@zju.edu.cn

ABSTRACT

Background: The heterogeneous tumor microenvironment (TME) contributes to poor prognosis of hepatocellular carcinoma (HCC). However, determining the modulation of TME during HCC progression remains a challenge.
Methods: Herein, the stromal score and immune score of HCC samples from The Cancer Genome Atlas database were calculated using the ESTIMATE algorithm and differentially expressed genes (DEGs) were obtained. Key DEGs were identified based on a protein-protein interaction network and survival analysis. Immunohistochemistry was carried out using primary samples to evaluate key DEGs expression. The *CIBERSORT* algorithm was applied to evaluate immune components. Gene Set Enrichment Analysis (GSEA) and correlation analysis were carried out to determine the relationship between key DEGs and tumor-infiltrating immune cells (TICs).
Results: The stromal score, immune score and estimate score correlated significantly with 1-year recurrence-free survival of patients with HCC. Interleukin-2 inducible T-cell kinase (ITK) was identified as the most prognostic DEG for patients with HCC. GSEA revealed that genes in the high ITK subgroup were enriched in inflammatory-immunological terms. CIBERSORT analysis identified nine TIC subsets that correlated with ITK expression.
Conclusion: We identified ITK as a novel indicator for early post-surgery tumor recurrence and microenvironment remodeling in HCC, providing a potential therapeutic target to treat HCC.

INTRODUCTION

Hepatocellular carcinoma (HCC) is the major form of primary liver cancer. According to authoritative statistics, liver cancer is the fourth most common cause of cancer-related death and ranks as sixth in accordance with incident cases worldwide [1]. The World Health Organization (WHO) claimed that over 1 million patients would die from liver cancer by the year 2030 [2]. HCC's occult symptoms and highly aggressive

properties mean that the majority of patients are diagnosed at an advanced stage, and the 5-year overall survival (OS) rate of patients with advanced HCC is only 25–39%, while the recurrence rate is nearly 80% [3]. Early post-operation recurrence and metastasis are the predominant cause of poor outcomes. Previous studies have reported that patients with HCC suffering early recurrence (within 1-year) after liver resection mainly had intrahepatic recurrence while patients with late recurrence commonly suffered from multi-extrahepatic metastasis, also, patients with early recurrence showed worse outcomes [4–7]. Thus, clinical management options and therapeutic efficacy for HCC are limited. Current clinical treatments for early-stage HCC include surgical resection, liver transplantation, and locoregional therapies, while systemic therapies are recommended for patients who have intermediate and advanced disease [8]. Promoted by the evolvement of multi-omics detection and analyzation techniques, reliable predictive biomarkers could guide scientific and clinical decisions that are crucial for patients with HCC [9].

The tumor microenvironment (TME) is the "soil" in which a tumor grows, as well as a severe obstacle in understanding and treating cancer. Not only cancer cells, but also immune cells, stromal cells, vascular networks, and many other components constitute the TME, representing a complex ecosystem in which cancer cells are formatted, proliferate, and progress [10]. To better determine the composition and function of the TME, several novel technologies, such as mass cytometry (CyTOF), single-cell RNA sequencing (scRNA-seq), and single-cell assay of transposase-accessible chromatin, have been adopted [11–15]. Most research has focus on tumor-infiltrating cells (TICs) especially tumor-infiltrating lymphocytes (TILs). A previous study demonstrated that the quantity of TILs in breast cancer is a robust prognostic factor for patient survival [16]. For HCC treatment, immunotherapy has recently become the new frontier of cancer treatment and the immunobiology of HCC is worthy of further exploration. Early in 2017, the landscape of infiltrating T cells was revealed by single-cell sequencing and has provided insights into immune modulation patterns [15]. Regulatory T cells (Tregs) promote tumor evasion via complicated mechanisms. The upregulation of TNF receptor superfamily member 4 (also known as OX40) expression on Tregs in the TME was reported to be associated with poor survival of patients with HCC [17]. Recent research reported that heterogeneity of exhausted T cells (Texs) in the TME is related to patient survival following resection in HCC. The authors stated that the high density of forkhead box P3 (FOXP3)[+] Tregs in the TME correlated strongly with early tumor recurrence [18]. Results derived from previous research

have increased our understanding of TME modulation of HCC and provided novel immunotherapeutic targets for further exploration.

To look deeper into the TME modulation of HCC and discover predictive biomarkers, as well as potential therapeutic targets, transcriptome-sequencing, and subsequent functional genomics analysis, are good available methods. In the current study, the ESTIMATE and CIBERSORT algorithms were used to determine the immune signatures of TICs from patients with HCC from The Cancer Genome Atlas (TCGA) database and identified interleukin-2 inducible T-cell kinase (ITK) as a predictive biomarker.

ITK belongs to the Tec family of kinases and is a crucial molecule in T cell development, differentiation, and effector function [19]. Besides its conventional participation in inflammatory activities, ITK has been reported to correlate with oncogenesis. ITK deficiency can promote severe Epstein-Barr-Virus (EBV) infection and lead to Hodgkin and non-Hodgkin lymphoma, lymphoproliferative disease, mononucleosis, and other diseases [20]. ITK expression in the TME of solid tumors is involved in the modulation of the micro-environment and is associated with prognosis. For instance, high ITK expression was found to predict better outcomes of patients with lung adenocarcinomas (LUAD) [21]. Controversially, it was reported that ITK upregulation is associated negatively with the prognosis of breast cancer [22]. As far as we know, there has been no research focusing on the correlation between ITK and the progression of HCC.

In the present study, the ESTIMATE and CIBERSORT algorithms were used to describe the immune landscape of patients with HCC. Further analysis revealed that ITK might be a potential indicator for post-operation prognosis and TME remodeling in HCC.

RESULTS

The stromal score, immune score, and estimate score were significantly associated with the prognosis of patients with HCC

Analysis workflow of this study was shown in Figure 1. To identify the correlations among the estimated stromal and immune scores with the RFS of patients with HCC, Kaplan–Meier survival analysis was used for the stromal score, immune score, and estimate score, respectively (the clinicopathological characteristics of the patients with HCC are shown in Supplementary Table 1). According to our analysis, all three scores showed a significant correlation with the 1-year RFS rate of patients with HCC who underwent liver

resection (stromal score, $p = 0.013$, Figure 2A, immune score, $p = 0.0016$, Figure 2B; estimate score $p = 0.001$, Figure 2C). We also evaluated the correlation between the scores and the long-term OS of patients with HCC; however, there was no significant difference between the low and high subgroups (Supplementary Figure 1A–1C).

The immune score and estimate score were consistent with the clinicopathological stages of HCC

To determine the correlation between the stromal and immune scores and the clinicopathological characteristics, we downloaded the clinical data of 373

patients with HCC from the TCGA database for further analysis (Figure 3A–3L). The immune score and estimate score correlated positively with the T classification of tumor-node-metastasis (TNM) stages (Figure 3F, $p = 0.033$; Figure 3J, $p = 0.048$). While the stromal score did not correlate with any classification of HCC (Figure 3A–3D).

Differentially expressed genes (DEGs) shared by the stromal score and immune score and their enrichment in immunity-related genes

To verify the key components in TME remodeling, we carried out comparisons between high and low stromal
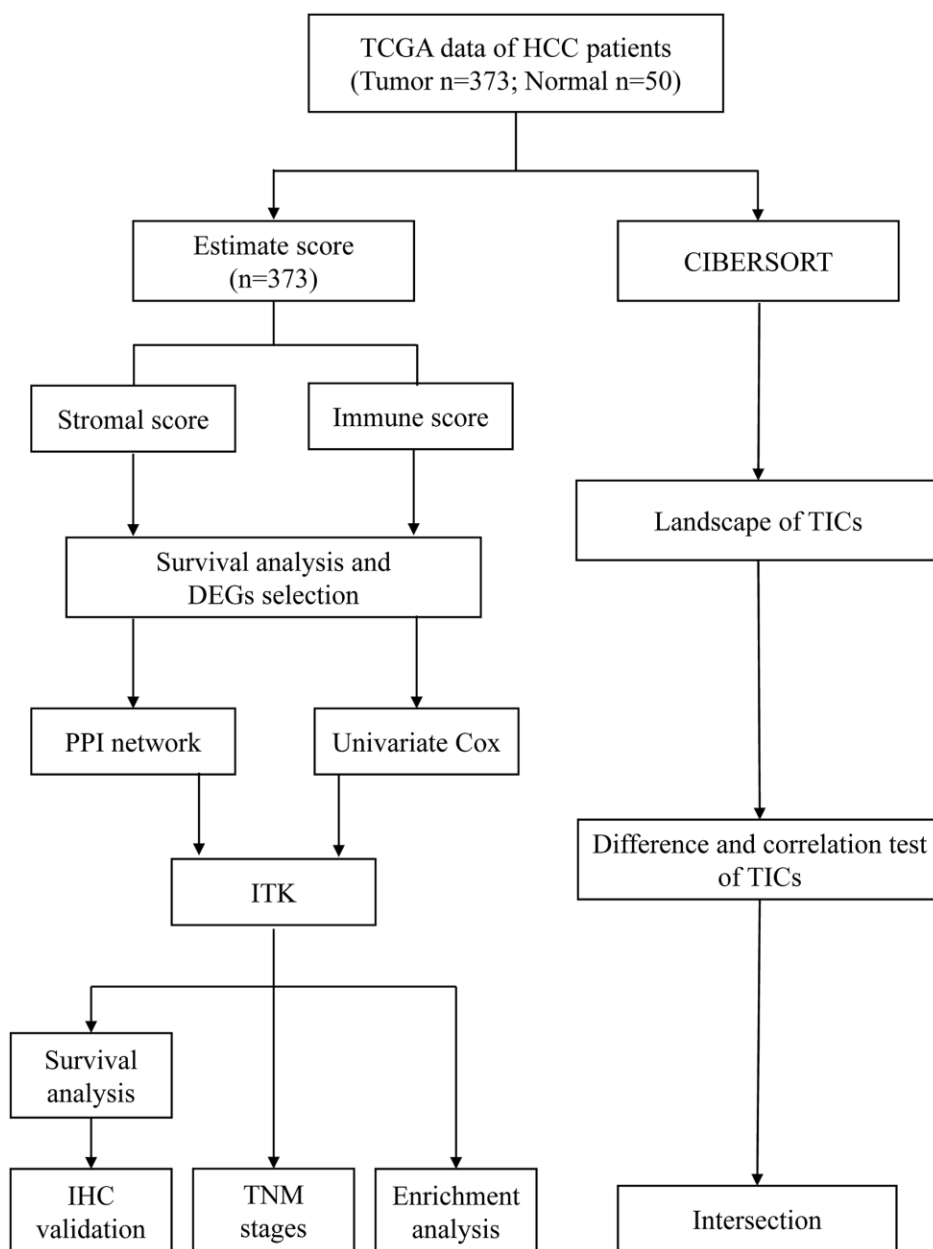

Figure 1. Analysis workflow of this study.

and immune score subgroups, respectively. 601 DEGs were obtained from the stromal score analysis, including 594 upregulated genes and 7 downregulated genes (Figure 4A). Similarly, 563 DEGs were obtained from the immune score when compared with the median, among which 557 genes were upregulated, and 6 were downregulated (Figure 4B). A Venn diagram showed that a total of 195 upregulated genes and 2 downregulated genes were shared by the high stromal score and high immune score groups (Figure 4C). These 197 DEGs were possibly the major components in TME remodeling. Gene ontology (GO) enrichment analysis showed that the DEGs were mainly enriched in terms that correlated with the immune response, such as T cell differentiation, lymphocyte differentiation, and T cell activation (Figure 4D). Likewise, the Kyoto Encyclopedia of Genes and Genomes (KEGG) and Reactome enrichment analysis showed that the DEGs were enriched in Th17 cell differentiation, the chemokine signaling pathway (Figure 4E), PD–1 signaling and immunoregulatory interactions between a Lymphoid and a non–Lymphoid cell (Supplementary Figure 2).

A protein-protein interaction (PPI) network and univariate COX regression identified two significant factors

A PPI network was then constructed based on the STRING database to explore the potential mechanisms of the DEGs in TME modulation [23]. The interactions between the protein encoded by the 197 DEGs are shown in Figure 5A. Sixteen hub genes in the PPI network were identified using MCODE, a plug-in in Cytoscape (Supplementary Table 2). Univariate COX regression analysis for the OS and RFS of patients with HCC identified 27 and 152 DEGs respectively among the 197 DEGs (Figure 5B, Supplementary Figure 3). The intersection analysis for the PPI network hub gene set, the OS-related gene set, and the RFS-related gene set was carried out next, which identified two factors, ITK and HLA Class II histocompatibility antigen DRβ5 (HLA-DRB5), as being present in the three gene sets (Figure 5C).

ITK mRNA expression correlates with postoperative outcomes and TNM Stages for patients with HCC

The Mann–Whitney U test showed that ITK expression in HCC tumors was lower than that in normal tissues (P < 0.001; Figure 6A). Differential analysis of 50 paired HCC tumor and paracarcinoma tissues also revealed consistent results (P < 0.001; Figure 6B). All HCC samples (n = 365) were divided into two subgroups based on the median expression level of ITK. Kaplan–Meier survival analysis was carried out to evaluate the predictive capacity of ITK expression. According to the results, patients with HCC in the high ITK group had better 1-year, 3-year, and 5-year OS rates (88.4% vs.
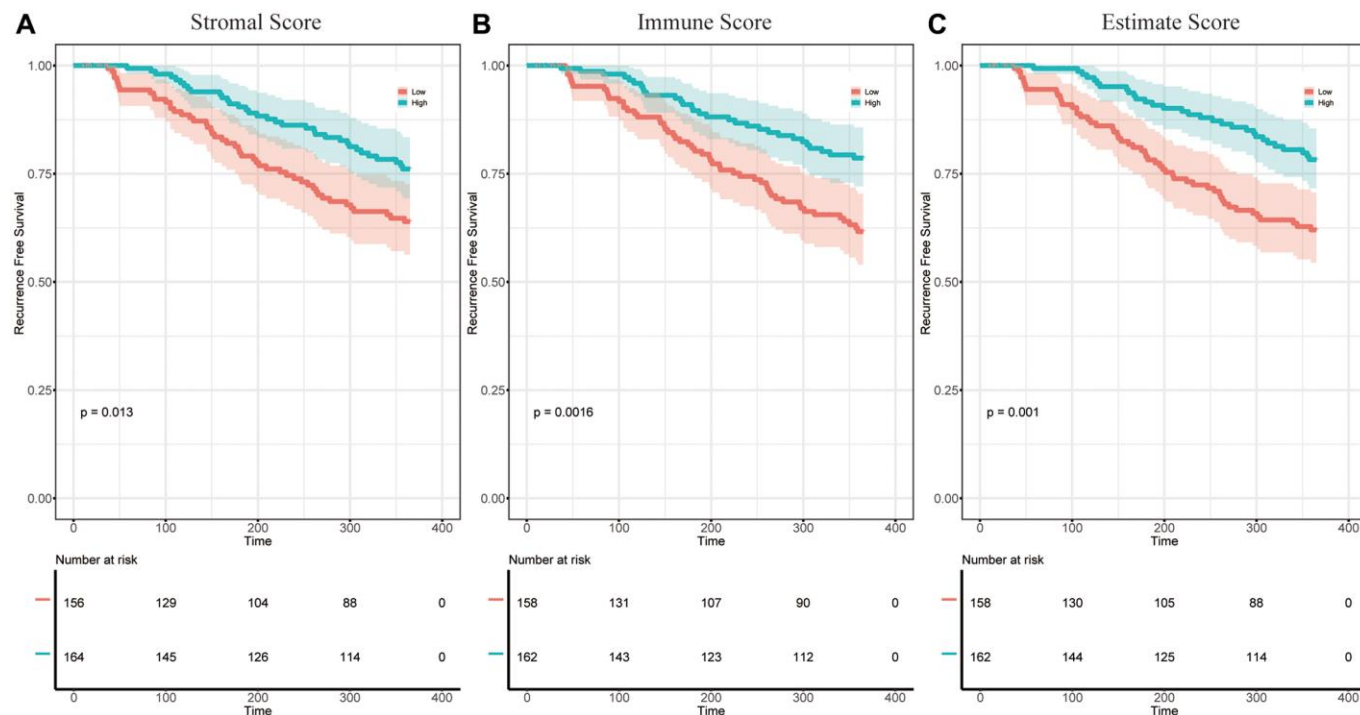

Figure 2. The correlation between estimate scores and the 1-year RFS of patients with HCC. (**A**) KM survival curves for the 1-year RFS of low/high stromal score subgroups (p = 0.013). (**B**) KM survival curves for the 1-year RFS of low/high immune score subgroups (p = 0.0016). (**C**) KM survival curves for the 1-year RFS of low/high estimate score subgroups (p = 0.001).

77.4%, 73.5% vs. 52.8%, and 54.0% vs. 40.8%; $p = 0.0085$, Figure 6C) and RFS (80.1% vs. 59.3%, 59.4% vs. 28.8%, and 46.2% vs. 21.7%; $p < 0.001$, Figure 6D) compared with those of the ITK low group. We subsequently analyzed the correlation between ITK expression and clinical characteristics. High expression of ITK correlated with earlier clinical stages and better T classification in the TNM stage system ($P < 0.05$, Figure 6E; $P < 0.05$, Figure 6F), while there was no significant difference in ITK expression for the N and M classifications (Figure 6G and 6H). HLA-DRB5 expression showed no significant correlation with OS ($p = 0.12$, Supplementary Figure 4C) or RFS ($p = 0.22$, Supplementary Figure 4D) according to Kaplan–Meier survival analysis. HLA-DRB5 expression was not correlated with TNM staging either (Supplementary Figure 4E–4H).

Immunohistochemistry (IHC) analysis proved that high ITK levels predict better postoperative outcomes of patients with HCC

To further verify the prognostic capacity of ITK, we collected primary tumor samples from patients with HCC who underwent liver resection in our medical center between 2015.01.01 and 2017.12.31 ($n = 176$) and carried out IHC staining to determine ITK expression in the tissues. All the subjects were divided into a high ITK group ($n = 69$) and a low ITK group ($n = 107$) based on the IHC score, which was evaluated by two pathologists. Kaplan–Meier survival analysis showed that patients in the high ITK group had better 1-year and 3-year OS rates (97.0% vs. 89.6%, 82.9% vs. 69.0%, $p < 0.001$, Figure 7A) as well as RFS rates compared with patients in the low ITK group (63.5%
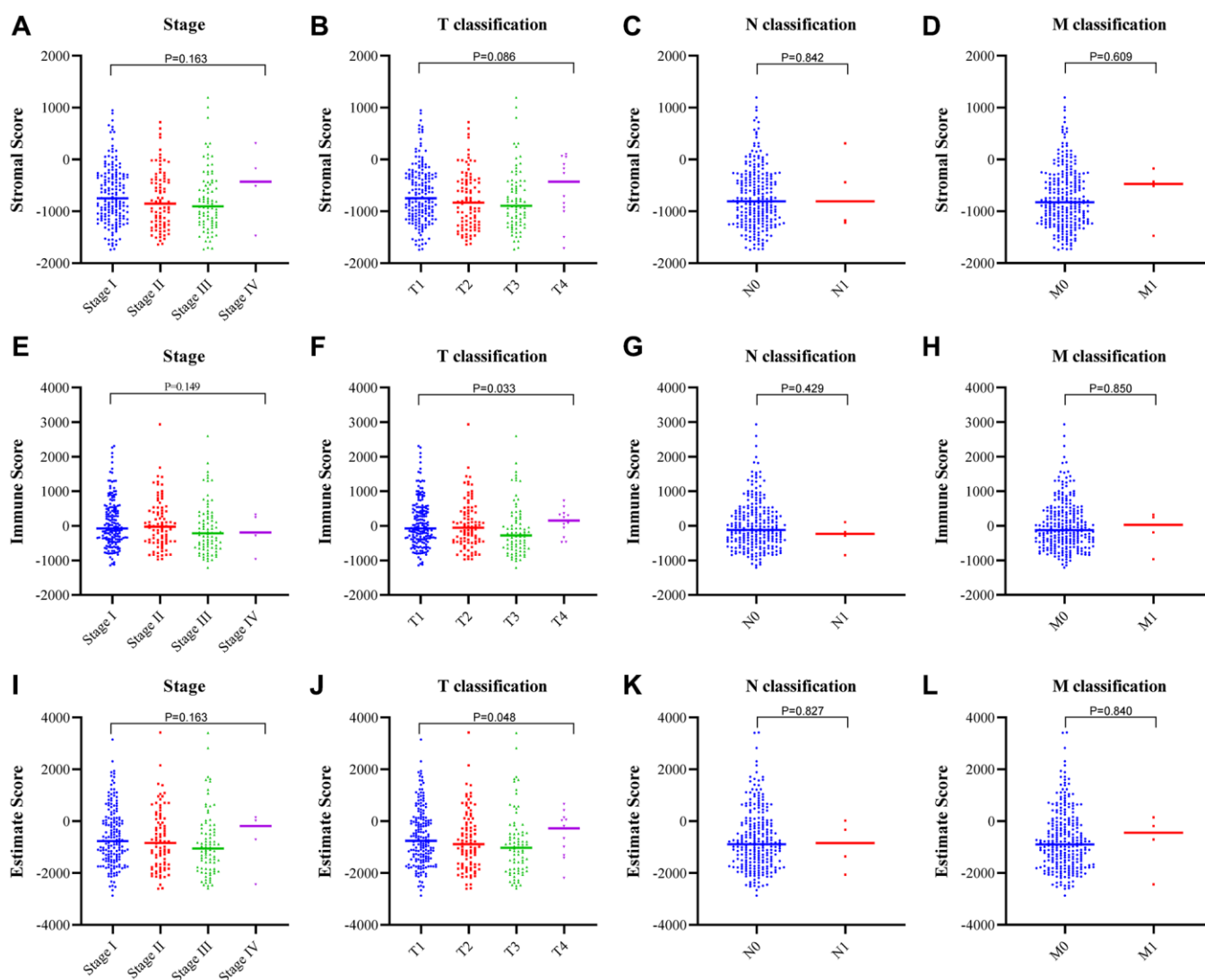

Figure 3. Correlation between the estimate scores and clinicopathological staging characteristics. (A–D) Correlation of the stromal score with the TNM stage. (E–H) Correlation of the immune score with the TNM stage. (I–L) Correlation of the estimate score with the TNM stage.

vs. 41.2%, 57.3% vs. 15.4%, *p* < 0.001, Figure 7B). Representative immunohistochemical pictures of ITK expression were shown in Figure 7C.

ITK potentially indicated TME modulation

The ITK expression level was related to the prognosis and clinicopathological stages of patients with HCC; therefore, we applied Gene Set Enrichment Analysis (GSEA) analysis to determine the underlying mechanisms. As shown in Figure 8A, in the ITK high group, the genes were mainly enriched in inflammatory activities, including interleukin (IL)-2/signal transducer and activator of transcription 5 (STAT5) signaling, interferon-alpha response, and IL-6/Janus kinase (JAK)/STAT3 signaling pathways.

The genes in the high ITK subgroup were enriched in multiple immunological gene sets according to the C7 collection defined by MSigDB (Figure 8B, the detailed information is shown in Supplementary Table 3).

ITK correlated with the proportion of TIC subsets

We subsequently analyzed the proportion of TICs using CIBERSORT to verify the correlation of ITK expression and the immune TME. As a result, the composition of 22 kinds of immune cells in HCC tissues was calculated (Figure 9A). The correlations among the 22 immune cell subpopulations are shown in Figure 9B. The results of principal component analysis (PCA) of 373 patients with HCC are presented in Figure 9C. The difference

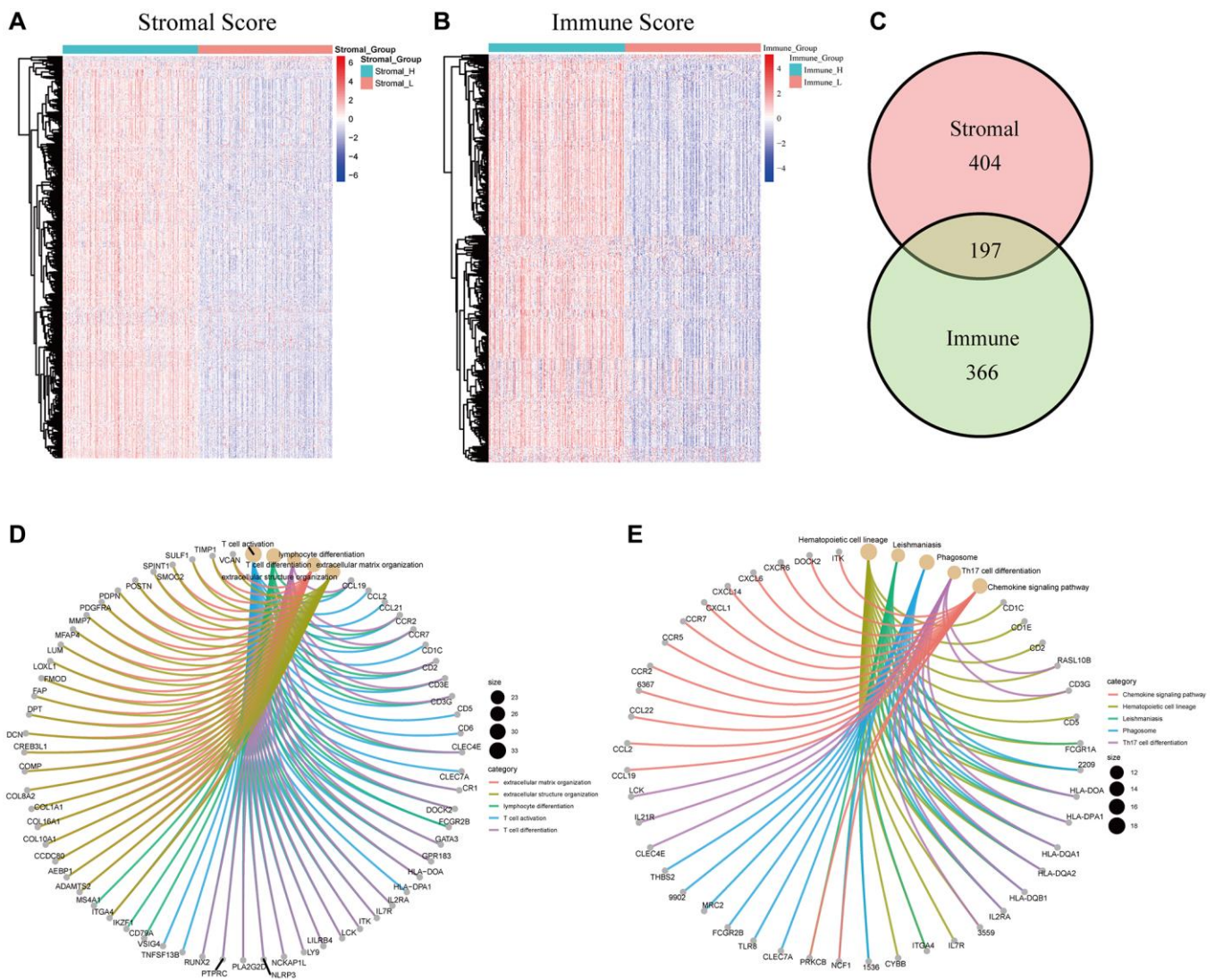**Figure 4. Cluster analysis, intersection analysis, GO, and KEGG enrichment analysis of the DEGs.** (**A**) Heatmap of 601 DEGs between the high/low stromal score subgroups. (**B**) Heatmap of 563 DEGs between the high/low immune score subgroups. (**C**) Venn diagram of the DEGs commonly shared by the two groups. (**D**) GO enrichment analysis of the common DEGs. (**E**) KEGG enrichment analysis of the common DEGs.

analysis screened out 10 types of TICs (Figure 10A) and the correlation analysis identified 13 types of TICs (Figure 10B, the TICs with no correlation with ITK expression are shown in Supplementary Figure 5). Nine types of TICs overlapped in the above analyses were shown in Figure 10C and Supplementary Table 4. Among the nine TIC subsets, three of them correlated positively with ITK mRNA expression, including activated CD4[+] memory T cells, CD8[+] T cells, and M1 Macrophages; and six types of TICs correlated negatively with ITK expression, including plasma cells, resting natural killer (NK) cells, activated dendritic cells, activated NK cells, naïve CD4[+] T cells, and resting mast cells. We then carried out survival analysis for the nine TIC subsets to predict postoperative OS and RFS for patients with HCC (Supplementary Figures 6–7). Results showed that CD8[+]T cells was associated with better survival, while resting NK cells, and plasma cells were associated with poor survival ($P < 0.05$). These results further supported the view that ITK is involved in immune modulation of the TME and probably exerted anti-tumor activities.

DISCUSSION

In the present study, we identified ITK as a prognostic predictor and TME remodeling indicator for patients with HCC through comprehensive data mining based on the TCGA database. The ESTIMATE algorithm was applied to determine the predictive value of the stromal and immune scores in the TME of HCC and further screen out crucial molecules. We also depicted the landscape of the HCC TME using the CIBERSORT algorithm and revealed a correlation between the TIC composition and ITK expression. Most importantly, the results showed that ITK expression could reflect alterations in the TME of HCC and predicted the postoperative outcomes of patients with HCC.

The TME is a unique environment that emerges during tumor formation. It has been recognized that the TME, particularly its immune components, plays an important role in HCC progression and therapeutic management. Immunotherapy has been considered a novel treatment with great potential for patients with HCC. Nivolumab
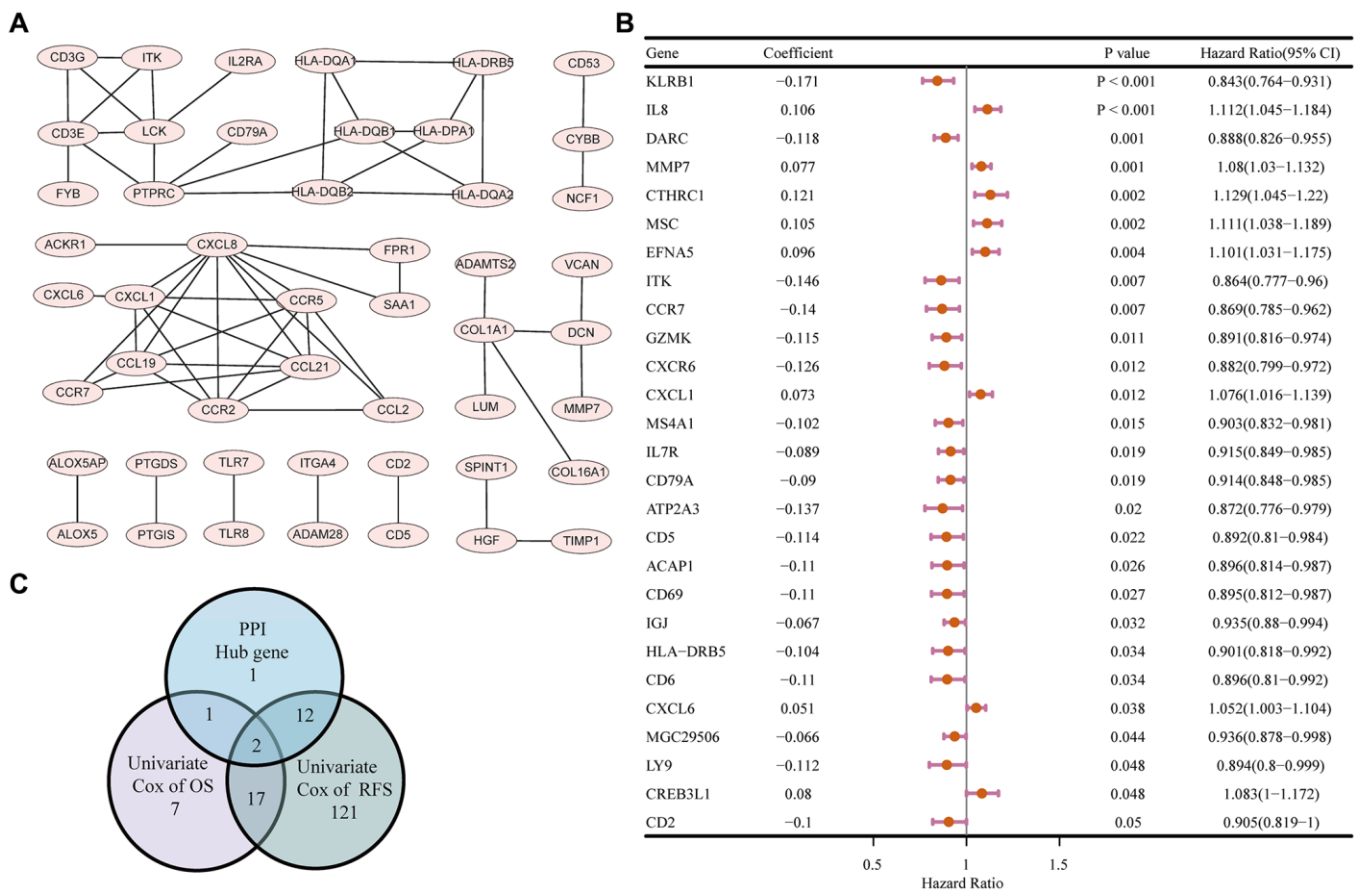

Figure 5. PPI network and univariate COX regression analysis. (**A**) PPI network of the nodes with combined score > 0.95. (**B**) Forest plot of the univariate COX regression analysis for OS. (**C**) Venn diagram of the factors commonly shared by hub genes in PPI and factors correlated with OS and RFS generated by univariate COX regression analysis.

(Opdivo) has been approved for second-line treatment in HCC [24]. Obstructions in detecting and analyzing the comprehensive immune microenvironment of each patient in a clinical context mean that there are still obstacles in therapeutic target identification and efficacy monitoring of immunotherapy for HCC. Although we have already identified several biomarkers to predict prognosis and evaluate the efficacy of treatment for HCC, such as alpha-fetoprotein (AFP) [25], Des-Carboxy Prothrombin (DCP) [26], osteo-pontin (OPN) [27], vascular endothelial growth factor (VEGF) [28], and Golgi protein 73 (Gp-73) [29, 30], representative and credible biomarkers that could precisely reflect the immune modulation of the TME are highly desirable.

Inspired by the above scenario, we carried out data mining of patients with HCC based on the TCGA database from an immunological point of view. The ESTIMATE algorithm was applied to evaluate the prognostic relevance of the stromal and immune components. Our results confirmed that the immune components in TME correlated significantly with 1-year RFS of patients with HCC. We then screened out DEGs generated from the low and high stromal and immune score subgroups. The results from GO, KEGG and

Reactome enrichment analysis showed that the DEGs were mainly mapped to immune-related terms, such as lymphocyte differentiation, T cell activation, and T cell differentiation. This was consistent with our initial hypothesis, as well as previous research. Subsequent intersection analysis of the PPI network and univariate COX regression analysis showed that ITK and HLA-DRB5 might be the key DEGs for HCC prognosis. HLA-DRB5 was subsequently filtered out because of its poor predictive capacity for OS of patients with HCC. Thus, we identified ITK as an immunological biomarker that performed well in evaluating the clinical stage as well as predicting the postoperative RFS and OS of patients with HCC. To further evaluate the relationship between ITK and the immune TME of HCC, functional enrichment analysis was conducted. The results of GSEA revealed that genes in the high ITK subgroup were enriched in inflammatory-immunological terms, including IL-2/STAT5 signaling, interferon-alpha response, and IL-6/JAK/STAT3 signaling pathways. Furthermore, we depicted the landscape of the HCC immune TME using the CIBERSORT algorithm, using 22 subsets of immune cells. Three kinds of TICs (CD8$^+$T cells, activated CD4$^+$ memory T cells, and M1 Macrophages) that correlated positively with ITK expression in HCC lesions were identified. These
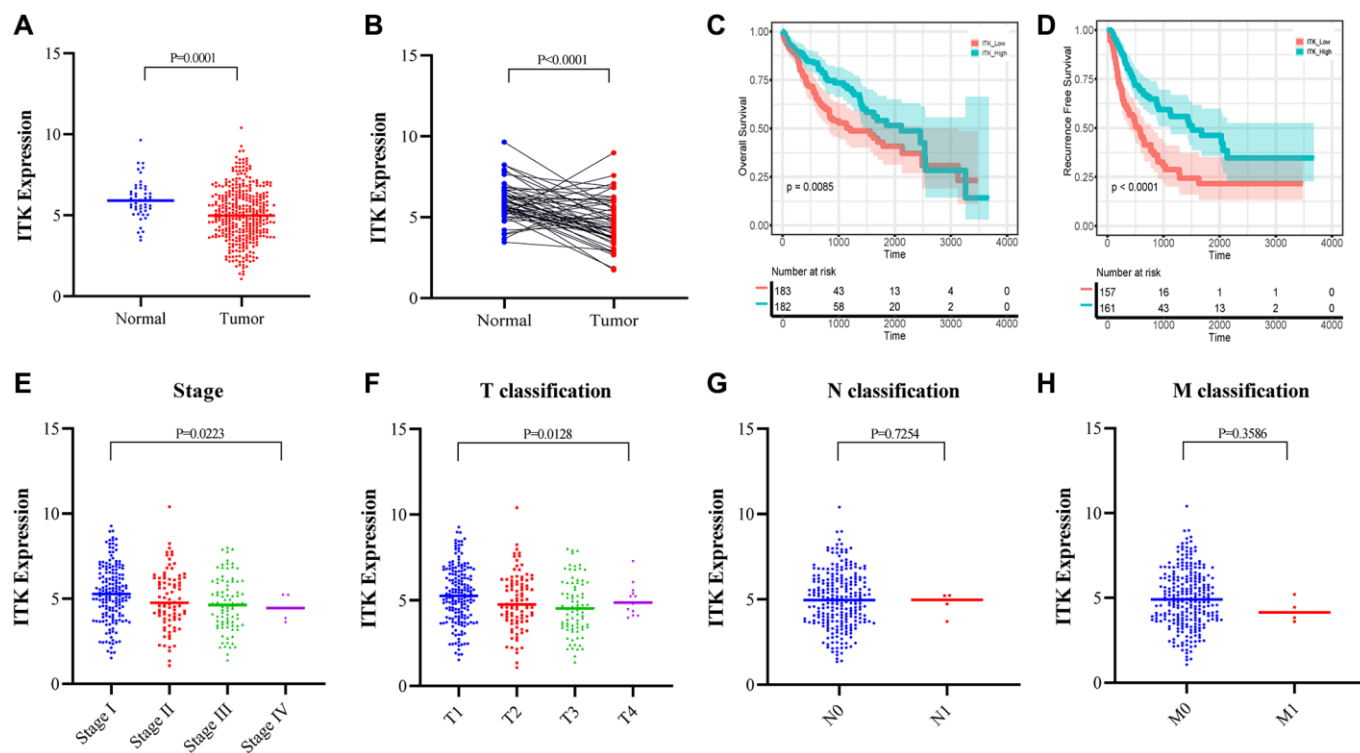
Figure 6. ITK expression in HCC tumors and paracarcinoma tissues, and the correlation between ITK and survival/clinical characteristics of patients with HCC (TCGA dataset). (**A**) ITK expression in HCC tumor tissues and paracarcinoma tissues ($p = 0.0001$). (**B**) ITK expression in paired HCC tumor tissues and paracarcinoma tissues derived from the same patient ($p < 0.0001$). (**C**) KM survival curves for the long-term OS of low/high ITK subgroups. (**D**) KM survival curves for the long-term RFS of low/high stromal score subgroups. (**E**–**H**) The correlation between ITK expression and clinicopathological stages.
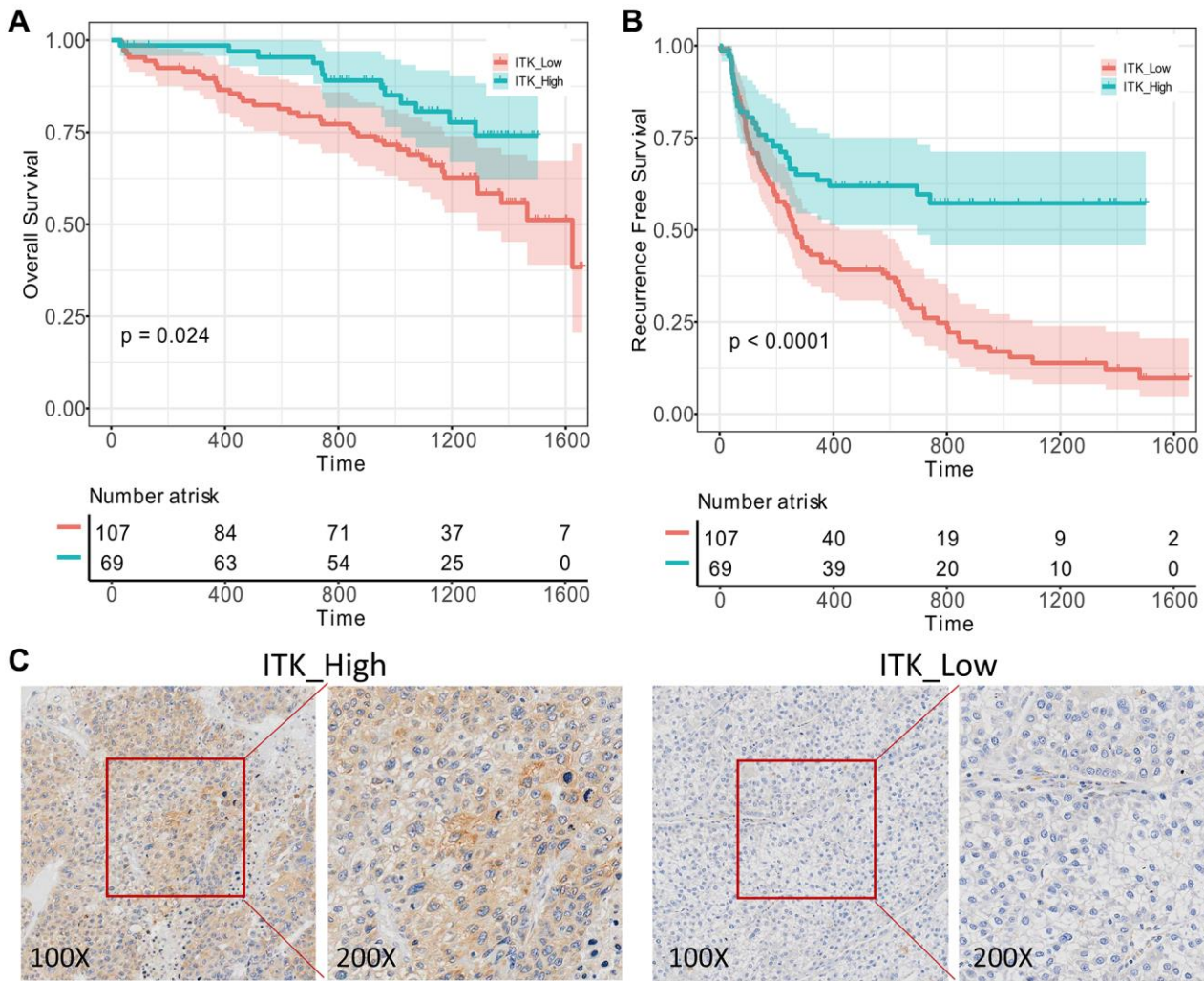
Figure 7. Validation of ITK's prognostic capacity. 176 pairs of HCC tumor and paracarcinoma tissues were obtained from our medical center and ITK expression levels were evaluated using IHC analysis. (**A**) KM survival curves for the post-operation OS of low/high ITK subgroups ($p = 0.024$). (**B**) KM survival curves for the post-operation RFS of low/high ITK subgroups ($p < 0.001$). (**C**) IHC of ITK expression.
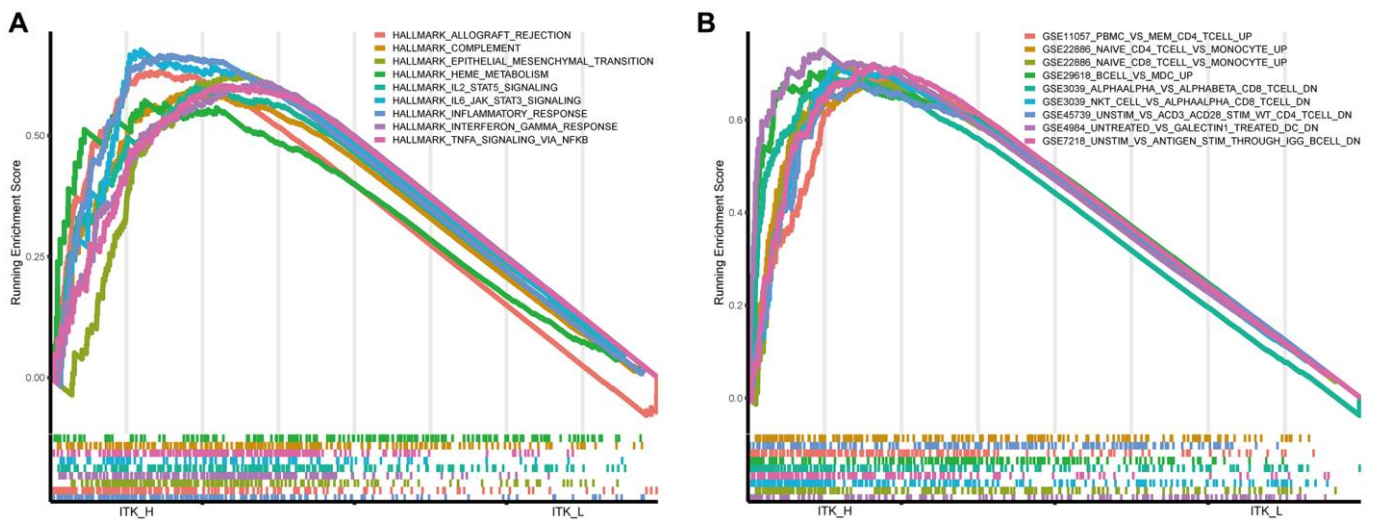
Figure 8. GSEA for HCC tumor samples. (**A**) Significantly enriched "hallmark gene sets" in the high ITK subgroup. (**B**) Significantly enriched "C7 gene sets" (the immunological gene sets) in the high ITK subgroup.

results were consistent with previous research findings. A recent study reported that the positive prognostic value of CD8[+] T cells was confirmed in more than 18,700 patients across 17 solid cancer types [31]. Similar conclusions for CD4[+] memory T cells and M1 Macrophages were published [32–36]. We also identified six TIC subpopulations, including plasma cells, resting NK cells, activated dendritic cells, activated NK cells, naïve CD4[+] T cells, and resting mast cells, which correlated negatively with ITK expression. Among these nine TIC subsets, CD8[+] T cells, resting NK cells, and plasma cells were related significantly to

the postoperative outcomes of patients with HCC. Our results indicated that the ITK expression level could dynamically reflect antitumor immune activities.

Notably, ITK is an important member of the Tec family kinases and takes part in T cell receptor (TCR) signaling events driving processes including T cell development and Th2/Th9/Th17 responses, and mutations of ITK lead to T cell dysfunction [10, 23, 37]. Previous research found that T cell-mediated antitumor responses play an important role in natural HCC progression. For instance, tumor-associated antigen
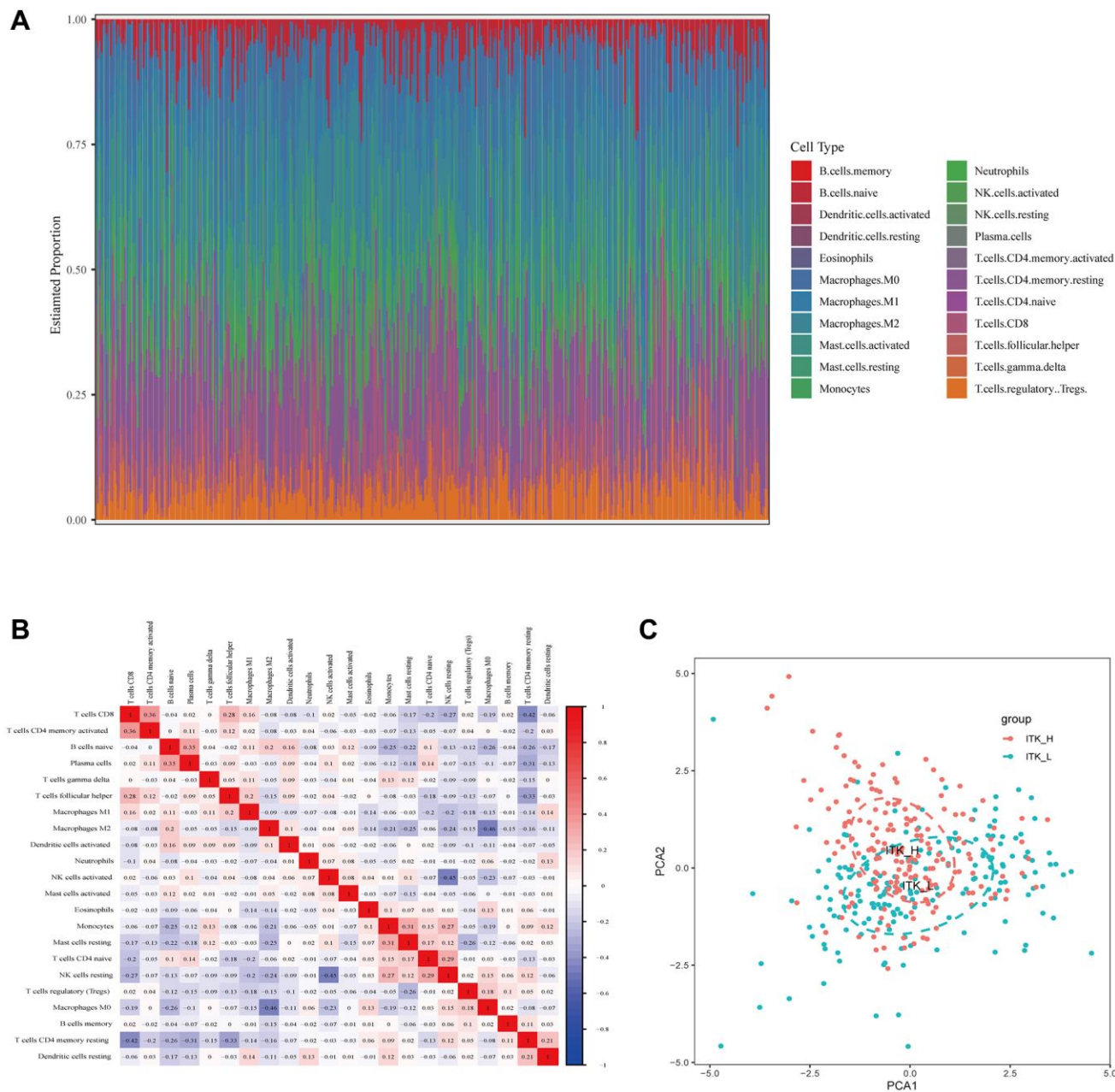
Figure 9. TIC profiling of HCC tumor tissues and correlation analysis. (**A**) The composition of 22 kinds of TICs in HCC tumor tissues is shown in a bar plot. (**B**) Heatmap showing the correlation between 22 kinds of TICs. (**C**) Principal component analysis of the HCC tumor tissues with high and low ITK expression.

(TAA)-specific CD8[+] T cells are significant components of HCC lesions and correlate with better RFS, indicating that the development of strategies aiming to enhance the total TAA-specific CD8[+] T cell response would unlock their full antitumor potential [38], thus contributing to better prognosis. Likewise, the upregulation of the other two identified TICs also leads to strengthened antitumor activities in the TME of patients with HCC. Among the reduced TIC subtypes in the high ITK group, dendritic cells (DCs) are important antigen-presenting cells (APCs) and are indispensable for T cell stimulation [39]. Our results showed that ITK expression correlated negatively with DCs density, indicating the dual role played by ITK in the immune
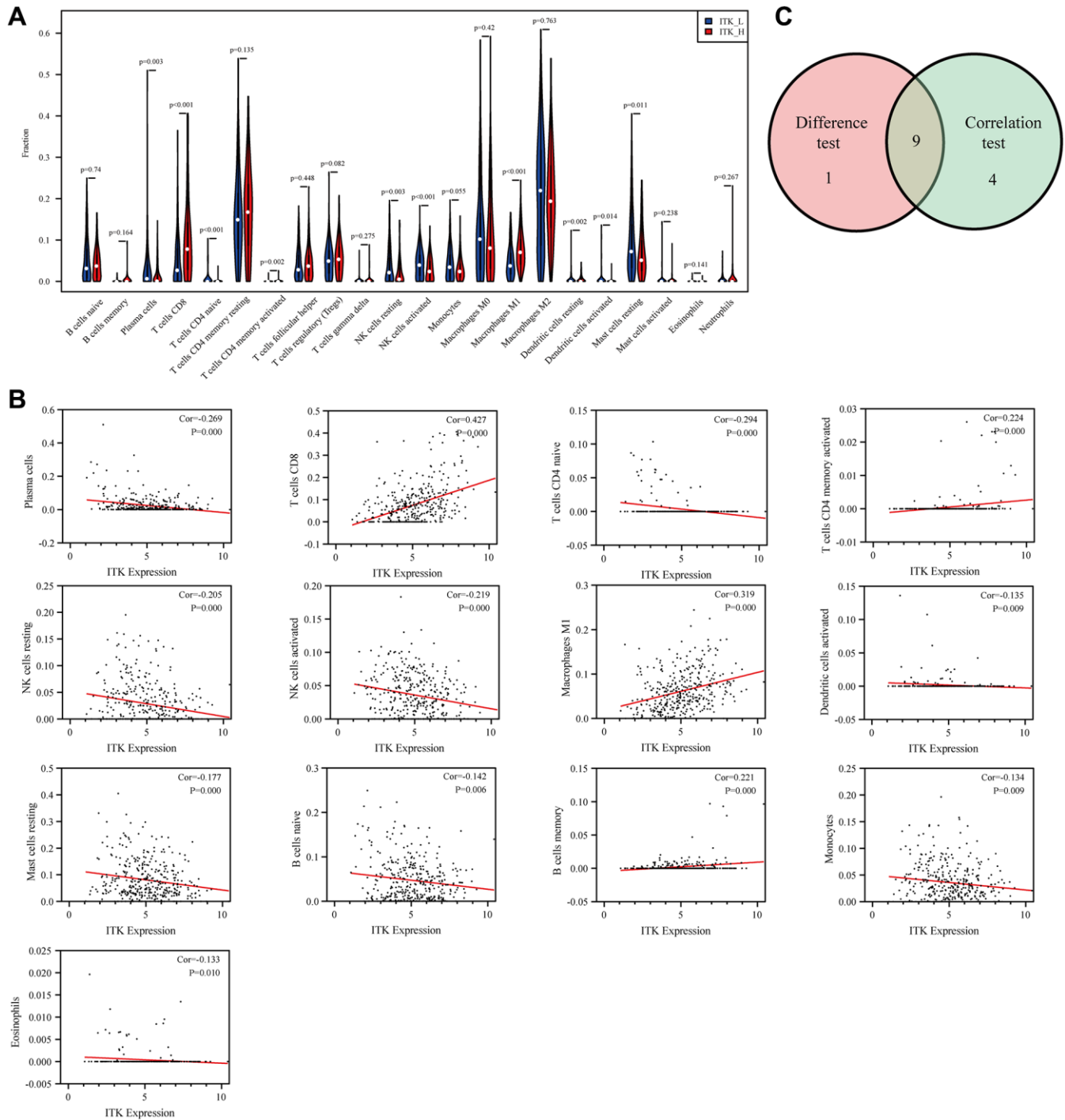
Figure 10. The correlation between ITK expression and TICs proportion. (**A**) The distributions of 22 kinds of TICs in low/high ITK subgroups are shown in a violin plot. (**B**) Scatter plots of the 13 kinds of TICs significantly correlated with ITK expression (*p* < 0.05). (**C**) Venn diagram showing nine common TICs shared, as assessed using differential analysis and correlation analysis.

remodeling of the TME in HCC. A previous study focusing on breast cancer stated that Ibrutinib, an ITK inhibitor, could suppress tumor development and metastasis by stimulating the development of mature DCs from myeloid-derived suppressor cells (MDSCs) and might be an effective therapeutic method to treat breast cancer [40]. This is consistent with the classic therapeutic function of Ibrutinib in hematological malignancies [41]. However, controversial results were shown in solid tumors. For instance, high ITK expression is related to better prognosis in lung adenocarcinomas (LUAD) [42]. In colon cell lines and rodent models, monotherapy using the ITK inhibitor, Ibrutinib, was reported to perform poorly in tumor suppression, whereas the combination of Ibrutinib and a programmed cell death 1 ligand 1 (PD-L1) inhibitor showed significant antitumor capacity [43]. This suggested that ITK participates in the immune remodeling of the TME in HCC through complex mechanisms.

In the present study, we observed that ITK could reflect the proportion of several important immune cells within the TME of HCC, acting as a credible biomarker for the remodeling status of the antitumor composition, and thus could act as a guide for immunotherapy applications [44]. The identification of molecular targets and the determination of their functions have revealed the therapeutic value of small-molecule inhibitors. To date, 43 small-molecule inhibitors have been approved by the FDA for oncology indications [45]. According to our study, ITK presents dual function in the immune TME of HCC, and further exploration should be carried out to determine the interaction between ITK and other immune cells, as well as cancer cells. Whether an ITK inhibitor or agonist is suitable for HCC treatment in different contexts needs more robust and direct experimental evidence.

Our study also has some limitations. Although we carried out data mining based on the authoritative TCGA database, further mechanistic studies and experimental validation are required.

In conclusion, we identified ITK as a novel biomarker of RFS and TME remodeling for HCC. As a crucial molecule in T cell differentiation, ITK upregulation in HCC lesions indicates the enhanced antitumor capacity of T cells in the TME and correlated positively with better prognosis. Increasing ITK expression contributes to TME remodeling of HCC via T cell activation. We hypothesized that ITK could serve as a sensitive biomarker for the prognosis of HCC and provides a potential therapeutic target for HCC treatment in the future.

MATERIALS AND METHODS

Data preparation

Transcriptome RNA-seq data of 373 patients with HCC and 50 normal liver samples and the corresponding clinical data were downloaded from the TCGA database (https://portal.gdc.cancer.gov/). Our study was performed in accordance with the publication guidelines provided by the TCGA.

Estimation of stromal and immune scores

ESTIMATE is an algorithm used to evaluate the tumor purity, which uses the gene expression of an immune signature (141 immune genes) and a stromal signature (141 stromal genes). [46] The ESTIMATE immune score and stromal score were calculated using R package estimate (v.1.0.13) [47] to analyze the infiltration levels of immune cells and stromal cells for each HCC sample, and patients with HCC were divided into high and low score groups based on the median value of immune or stromal score.

Survival analysis

Overall survival (OS) of 365 patients with HCC and recurrence-free survival (RFS) of 320 patients with HCC were measured using the R package survival (v.3.2–3) [48] and survminer (v.0.4.8) [49]. The Kaplan–Meier method was used to plot the survival curve, and the log-rank test was used to show statistical significance, in which $p < 0.05$ was considered significant.

Correlation analysis of the estimate score with clinical stages

The clinicopathological characteristics data of the patients with HCC were downloaded from the TCGA. SPSS (v24.0; IBM Corp., Armonk, NY, USA) was used to perform the analysis. The Mann–Whitney U test was used for comparison between scores and the N/M classification, and the Kruskal–Wallis rank-sum test was used to compare between scores and TNM stage/T classification, in which $p < 0.05$ was considered significant.

Differentially expressed gene analysis

Based on the median score of the immune and stromal scores, 373 patients with HCC were subdivided into two groups (high score group and low score group), respectively. DEGs analysis was performed using R Package limma [50] to identify DEGs between the high and low score groups. DEGs with an absolute log2

fold-change >1.5 (high score vs. low score group) and an adjusted *p*-value <0.05 were considered as statistically significant.

Functional enrichment analysis

The DEGs between the high score group and the low score group were used for GO, KEGG and Reactome enrichment analyses with the R package clusterProfiler [51] and ReactomePA [52]. Only terms with both an adjusted *p*-value and *q* value < 0.05 were considered as significantly enriched.

PPI network construction

DEGs were imported to the STRING database to evaluate their interactive relationships, and PPI pairs with a combined score > 0.95 were then reconstructed using Cytoscape version 3.7.1 [53]. Hub genes in the PPI network were identified using MCODE, a Cytoscape plug-in.

Gene set enrichment analysis

GSEA was performed using the R package clusterProfile with "Hallmark gene sets" and "c7 class: immunologic signature gene sets" v7.2 [54], and only gene sets with an adjusted *p*-value < 0.05 and *q*- < 0.05 were considered as significant.

TICs profile

The CIBERSORT [55] algorithm was used to evaluate the abundance of 22 types of TICs based on the expression data from 373 tumor samples from the TCGA.

IHC of primary samples of patients with HCC

Primary tumoral and paratumoral tissue samples of patients with HCC who underwent liver resection in our medical center between 2015.01.01 and 2017.12.31 (*n* = 176) were obtained (Ethical approval:2018–768). ITK immunohistochemistry was performed according to common protocols. The anti-ITK antibody was purchased from Abcam (Cambridge, MA, USA; No. ab32039).

Abbreviations

AFP: alpha-fetoprotein; APCs: antigen-presenting cells; DCP: Des-Carboxy Prothrombin; DEG: differentially expressed genes; EBV: Epstein-Barr-Virus; GO: Gene ontology; Gp-73: Golgi protein 73; GSEA: Gene Set Enrichment Analysis; HCC: hepatocellular carcinoma; HLA-DRB5: HLA Class II histocompatibility antigen DRβ5; IHC: Immunohistochemistry; ITK: interleukin-2 inducible T-cell kinase; KEGG: Kyoto Encyclopedia of Genes and Genomes; LUAD: lung adenocarcinomas; NK: natural killer; OPN: osteopontin; OS: overall survival; PCA: principal component analysis; PPI: protein-protein interaction; RFS: recurrence-free survival; TAA: Tumor-associated antigen; TCGA: The Cancer Genome Atlas; TCR: T cell receptor; Texs: exhausted T cells; TICs: tumor-infiltrating immune cells; TME: tumor microenvironment; Tregs: regulatory T cells; VEGF: vascular endothelial growth factor; WHO: World Health Organization.

AUTHOR CONTRIBUTIONS

All authors searched the literature, designed the study, interpreted the findings, and revised the manuscript. Binhua Pan and Modan Yang carried out data management and statistical analysis, and drafted the manuscript. Xuyong Wei, Wangyao Li, Mengfan Yang, Kun Wang, Di Lu, and Rui Wang helped with cohort identification and data management. Xiao Xu performed project administration.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest related to this study.

FUNDING

REFERENCES

1. Sun N, Lee YT, Zhang RY, Kao R, Teng PC, Yang Y, Yang P, Wang JJ, Smalley M, Chen PJ, Kim M, Chou SJ, Bao L, et al. Purification of HCC-specific extracellular vesicles on nanosubstrates for early HCC detection by digital scoring. Nat Commun. 2020; 11:4489. https://doi.org/10.1038/s41467-020-18311-0 PMID:32895384

2. World Health Organization. Projections of mortality and causes of death, 2016 to 2060. http://www.who.int/healthinfo/global_burden_disease/projections/en/.

3. Jin GZ, Zhang Y, Cong WM, Wu X, Wang X, Wu S, Wang S, Zhou W, Yuan S, Gao H, Yu G, Yang W. Phosphoglucomutase 1 inhibits hepatocellular carcinoma progression by regulating glucose trafficking. PLoS Biol. 2018; 16:e2006483. https://doi.org/10.1371/journal.pbio.2006483 PMID:30335765

4. Song M, He J, Pan QZ, Yang J, Zhao J, Zhang YJ, Huang Y, Tang Y, Wang Q, He J, Gu J, Li Y, Chen S, et al. Cancer-Associated Fibroblast-Mediated Cellular Crosstalk Supports Hepatocellular Carcinoma Progression. Hepatology. 2021; 73:1717–35. https://doi.org/10.1002/hep.31792 PMID:33682185

5. Poon RT, Fan ST, Ng IO, Lo CM, Liu CL, Wong J. Different risk factors and prognosis for early and late intrahepatic recurrence after resection of hepatocellular carcinoma. Cancer. 2000; 89:500–07. PMID:10931448

6. Singal AG, Rich NE, Mehta N, Branch A, Pillai A, Hoteit M, Volk M, Odewole M, Scaglione S, Guy J, Said A, Feld JJ, John BV, et al. Direct-Acting Antiviral Therapy Not Associated With Recurrence of Hepatocellular Carcinoma in a Multicenter North American Cohort Study. Gastroenterology. 2019; 156:1683–92.e1. https://doi.org/10.1053/j.gastro.2019.01.027 PMID:30660729

7. Yin Z, Jiang K, Li R, Dong C, Wang L. Multipotent mesenchymal stromal cells play critical roles in hepatocellular carcinoma initiation, progression and therapy. Mol Cancer. 2018; 17:178. https://doi.org/10.1186/s12943-018-0926-6 PMID:30593276

8. Villanueva A. Hepatocellular Carcinoma. N Engl J Med. 2019; 380:1450–62. https://doi.org/10.1056/NEJMra1713263 PMID:30970190

9. Ye Q, Ling S, Zheng S, Xu X. Liquid biopsy in hepatocellular carcinoma: circulating tumor cells and circulating tumor DNA. Mol Cancer. 2019; 18:114. https://doi.org/10.1186/s12943-019-1043-x PMID:31269959

10. Duan Q, Zhang H, Zheng J, Zhang L. Turning Cold into Hot: Firing up the Tumor Microenvironment. Trends Cancer. 2020; 6:605–18. https://doi.org/10.1016/j.trecan.2020.02.022 PMID:32610070

11. Azizi E, Carr AJ, Plitas G, Cornish AE, Konopacki C, Prabhakaran S, Nainys J, Wu K, Kiseliovas V, Setty M, Choi K, Fromme RM, Dao P, et al. Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment. Cell. 2018; 174:1293–308.e36. https://doi.org/10.1016/j.cell.2018.05.060 PMID:29961579

12. Guo X, Zhang Y, Zheng L, Zheng C, Song J, Zhang Q, Kang B, Liu Z, Jin L, Xing R, Gao R, Zhang L, Dong M, et al. Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. Nat Med. 2018; 24:978–85. https://doi.org/10.1038/s41591-018-0045-3 PMID:29942094

13. Savas P, Virassamy B, Ye C, Salim A, Mintoff CP, Caramia F, Salgado R, Byrne DJ, Teo ZL, Dushyanthen S, Byrne A, Wein L, Luen SJ, et al, and Kathleen Cuningham Foundation Consortium for Research into Familial Breast Cancer (kConFab). Publisher Correction: Single-cell profiling of breast cancer T cells reveals a tissue-resident memory subset associated with improved prognosis. Nat Med. 2018; 24:1941. https://doi.org/10.1038/s41591-018-0176-6 PMID:30135555

14. Wagner J, Rapsomaniki MA, Chevrier S, Anzeneder T, Langwieder C, Dykgers A, Rees M, Ramaswamy A, Muenst S, Soysal SD, Jacobs A, Windhager J, Silina K, et al. A Single-Cell Atlas of the Tumor and Immune Ecosystem of Human Breast Cancer. Cell. 2019; 177:1330–45.e18. https://doi.org/10.1016/j.cell.2019.03.005 PMID:30982598

15. Zheng C, Zheng L, Yoo JK, Guo H, Zhang Y, Guo X, Kang B, Hu R, Huang JY, Zhang Q, Liu Z, Dong M, Hu X, et al. Landscape of Infiltrating T Cells in Liver Cancer Revealed by Single-Cell Sequencing. Cell. 2017; 169:1342–56.e16. https://doi.org/10.1016/j.cell.2017.05.035 PMID:28622514

16. Savas P, Salgado R, Denkert C, Sotiriou C, Darcy PK, Smyth MJ, Loi S. Clinical relevance of host immunity in breast cancer: from TILs to the clinic. Nat Rev Clin Oncol. 2016; 13:228–41. https://doi.org/10.1038/nrclinonc.2015.215 PMID:26667975

17. Xie K, Xu L, Wu H, Liao H, Luo L, Liao M, Gong J, Deng Y, Yuan K, Wu H, Zeng Y. OX40 expression in hepatocellular carcinoma is associated with a distinct immune microenvironment, specific mutation signature, and poor prognosis. Oncoimmunology. 2018; 7:e1404214. https://doi.org/10.1080/2162402X.2017.1404214 PMID:29632718

18. Shen X, Yang F, Yang P, Yang M, Xu L, Zhuo J, Wang J, Lu D, Liu Z, Zheng SS, Niu T, Xu X. A Contrast-Enhanced Computed Tomography Based Radiomics Approach for Preoperative Differentiation of

Pancreatic Cystic Neoplasm Subtypes: A Feasibility Study. Front Oncol. 2020; 10:248.
https://doi.org/10.3389/fonc.2020.00248
PMID:32185129

19. Mano H. Tec family of protein-tyrosine kinases: an overview of their structure and function. Cytokine Growth Factor Rev. 1999; 10:267–80.
https://doi.org/10.1016/s1359-6101(99)00019-2
PMID:10647781

20. Tangye SG, Palendira U, Edwards ES. Human immunity against EBV-lessons from the clinic. J Exp Med. 2017; 214:269–83.
https://doi.org/10.1084/jem.20161846
PMID:28108590

21. Wang K, Wei X, Wei Q, Lu D, Li W, Pan B, Chen J, Xie H, Zheng S, Xu X. A two-circular RNA signature of donor circFOXN2 and circNECTIN3 predicts early allograft dysfunction after liver transplantation. Ann Transl Med. 2020; 8:94.
https://doi.org/10.21037/atm.2019.12.132
PMID:32175387

22. Zhao X, Hu D, Li J, Zhao G, Tang W, Cheng H. Database Mining of Genes of Prognostic Value for the Prostate Adenocarcinoma Microenvironment Using the Cancer Gene Atlas. Biomed Res Int. 2020; 2020:5019793.
https://doi.org/10.1155/2020/5019793
PMID:32509861

23. Lechner KS, Neurath MF, Weigmann B. Role of the IL-2 inducible tyrosine kinase ITK and its inhibitors in disease pathogenesis. J Mol Med (Berl). 2020; 98:1385–95.
https://doi.org/10.1007/s00109-020-01958-z
PMID:32808093

24. Oh S, Park Y, Lee HJ, Lee J, Lee SH, Baek YS, Chun SK, Lee SM, Kim M, Chon YE, Ha Y, Cho Y, Kim GJ, et al. A Disintegrin and Metalloproteinase 9 (ADAM9) in Advanced Hepatocellular Carcinoma and Their Role as a Biomarker During Hepatocellular Carcinoma Immunotherapy. Cancers (Basel). 2020; 12:745.
https://doi.org/10.3390/cancers12030745
PMID:32245188

25. Wong RJ, Ahmed A, Gish RG. Elevated alpha-fetoprotein: differential diagnosis - hepatocellular carcinoma and other disorders. Clin Liver Dis. 2015; 19:309–23.
https://doi.org/10.1016/j.cld.2015.01.005
PMID:25921665

26. Volk ML, Hernandez JC, Su GL, Lok AS, Marrero JA. Risk factors for hepatocellular carcinoma may impair the performance of biomarkers: a comparison of AFP, DCP, and AFP-L3. Cancer Biomark. 2007; 3:79–87.

https://doi.org/10.3233/cbm-2007-3202
PMID:17522429

27. Pan HW, Ou YH, Peng SY, Liu SH, Lai PL, Lee PH, Sheu JC, Chen CL, Hsu HC. Overexpression of osteopontin is associated with intrahepatic metastasis, early recurrence, and poorer prognosis of surgically resected hepatocellular carcinoma. Cancer. 2003; 98:119–27.
https://doi.org/10.1002/cncr.11487
PMID:12833464

28. Mazziotti G, Sorvillo F, Morisco F, Carbone A, Rotondi M, Stornaiuolo G, Precone DF, Cioffi M, Gaeta GB, Caporaso N, Carella C. Serum insulin-like growth factor I evaluation as a useful tool for predicting the risk of developing hepatocellular carcinoma in patients with hepatitis C virus-related cirrhosis: a prospective study. Cancer. 2002; 95:2539–45.
https://doi.org/10.1002/cncr.11002
PMID:12467068

29. Tremosini S, Forner A, Boix L, Vilana R, Bianchi L, Reig M, Rimola J, Rodríguez-Lope C, Ayuso C, Solé M, Bruix J. Prospective validation of an immunohistochemical panel (glypican 3, heat shock protein 70 and glutamine synthetase) in liver biopsies for diagnosis of very early hepatocellular carcinoma. Gut. 2012; 61:1481–87.
https://doi.org/10.1136/gutjnl-2011-301862
PMID:22287594

30. Sai WL, Yao M, Shen SJ, Zheng WJ, Sun JY, Wu MN, Wang L, Yao DF. Dynamic expression of hepatic GP73 mRNA and protein and circulating GP73 during hepatocytes malignant transformation. Hepatobiliary Pancreat Dis Int. 2020; 19:449–54.
https://doi.org/10.1016/j.hbpd.2020.02.009
PMID:32171652

31. Bruni D, Angell HK, Galon J. The immune contexture and Immunoscore in cancer prognosis and therapeutic efficacy. Nat Rev Cancer. 2020; 20:662–80.
https://doi.org/10.1038/s41568-020-0285-7
PMID:32753728

32. Galon J, Costes A, Sanchez-Cabo F, Kirilovsky A, Mlecnik B, Lagorce-Pagès C, Tosolini M, Camus M, Berger A, Wind P, Zinzindohoué F, Bruneval P, Cugnenc PH, et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. Science. 2006; 313:1960–64.
https://doi.org/10.1126/science.1129139
PMID:17008531

33. Fridman WH, Pagès F, Sautès-Fridman C, Galon J. The immune contexture in human tumours: impact on clinical outcome. Nat Rev Cancer. 2012; 12:298–306.
https://doi.org/10.1038/nrc3245
PMID:22419253

34. Aras S, Zaidi MR. TAMeless traitors: macrophages in cancer progression and metastasis. Br J Cancer. 2017; 117:1583–91.
https://doi.org/10.1038/bjc.2017.356
PMID:29065107

35. Fridman WH, Zitvogel L, Sautès-Fridman C, Kroemer G. The immune contexture in cancer prognosis and treatment. Nat Rev Clin Oncol. 2017; 14:717–34.
https://doi.org/10.1038/nrclinonc.2017.101
PMID:28741618

36. Hu Z, Gu X, Zhong R, Zhong H. Tumor-infiltrating CD45RO[+] memory cells correlate with favorable prognosis in patients with lung adenocarcinoma. J Thorac Dis. 2018; 10:2089–99.
https://doi.org/10.21037/jtd.2018.03.148
PMID:29850112

37. Wallace JG, Alosaimi MF, Khayat CD, Jaber F, Almutairi A, Beaussant-Cohen S, Pinkus G, Fleming M, Mehawej C, Chou J, Geha RS. ITK deficiency presenting as autoimmune lymphoproliferative syndrome. J Allergy Clin Immunol. 2021; 147:743–45.e1.
https://doi.org/10.1016/j.jaci.2020.06.019
PMID:32628964

38. Song B, Zhen S, Meng F. T cell inflammation profile after surgical resection may predict tumor recurrence in HBV-related hepatocellular carcinoma. Int Immunopharmacol. 2016; 41:35–41.
https://doi.org/10.1016/j.intimp.2016.10.015
PMID:27792920

39. Streba LA, Streba CT, Săndulescu L, Vere CC, Mitruţ P, Cotoi BV, Popescu LN, Ion DA. Dendritic cells and hepatocellular carcinoma. Rom J Morphol Embryol. 2014; 55:1287–93.
PMID:25611258

40. Varikuti S, Singh B, Volpedo G, Ahirwar DK, Jha BK, Saljoughian N, Viana AG, Verma C, Hamza O, Halsey G, Holcomb EA, Maryala RJ, Oghumu S, et al. Ibrutinib treatment inhibits breast cancer progression and metastasis by inducing conversion of myeloid-derived suppressor cells to dendritic cells. Br J Cancer. 2020; 122:1005–13.
https://doi.org/10.1038/s41416-020-0743-8
PMID:32025027

41. Latour S, Fischer A. Signaling pathways involved in the T-cell-mediated immunity against Epstein-Barr virus: Lessons from genetic diseases. Immunol Rev. 2019; 291:174–89.
https://doi.org/10.1111/imr.12791
PMID:31402499

42. Zhou L, Xie H, Chen X, Wan J, Xu S, Han Y, Chen D, Qiao Y, Zhou L, Zheng S, Wang H. Dimerization-induced self-assembly of a redox-responsive prodrug into nanoparticles for improved therapeutic index. Acta Biomater. 2020; 113:464–77.
https://doi.org/10.1016/j.actbio.2020.07.007
PMID:32652227

43. Sagiv-Barfi I, Kohrt HE, Czerwinski DK, Ng PP, Chang BY, Levy R. Therapeutic antitumor immunity by checkpoint blockade is enhanced by ibrutinib, an inhibitor of both BTK and ITK. Proc Natl Acad Sci U S A. 2015; 112:E966–72.
https://doi.org/10.1073/pnas.1500712112
PMID:25730880

44. Chen Y, E CY, Gong ZW, Liu S, Wang ZX, Yang YS, Zhang XW. Chimeric antigen receptor-engineered T-cell therapy for liver cancer. Hepatobiliary Pancreat Dis Int. 2018; 17:301–09.
https://doi.org/10.1016/j.hbpd.2018.05.005
PMID:29861325

45. Bedard PL, Hyman DM, Davids MS, Siu LL. Small molecules, big impact: 20 years of targeted therapy in oncology. Lancet. 2020; 395:1078–88.
https://doi.org/10.1016/S0140-6736(20)30164-1
PMID:32222192

46. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, Treviño V, Shen H, Laird PW, Levine DA, Carter SL, Getz G, Stemke-Hale K, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun. 2013; 4:2612.
https://doi.org/10.1038/ncomms3612
PMID:24113773

47. Kosuke Y, Hoon K, Roel GW. estimate: Estimate of Stromal and Immune Cells in Malignant Tumor Tissues from Expression Data. R package version 1.0.13/r21. https://R-Forge.R-project.org/projects/estimate/2016.

48. Therneau T. A Package for Survival Analysis in R. R package version 3.2–3, https://CRAN.R-project.org/package=survival. 2020.

49. Alboukadel K, Marcin K, Przemyslaw B. survminer: Drawing Survival Curves using 'ggplot2'. R package version 0.4.8., 2000. https://cran.r-project.org/package=survminer.

50. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015; 43:e47.
https://doi.org/10.1093/nar/gkv007
PMID:25605792

51. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012; 16:284–87.

https://doi.org/10.1089/omi.2011.0118
PMID:22455463

52. Yu G, He QY. ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. Mol Biosyst. 2016; 12:477–79.
https://doi.org/10.1039/c5mb00663e
PMID:26661513

53. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003; 13:2498–504.
https://doi.org/10.1101/gr.1239303
PMID:14597658

54. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005; 102:15545–50.
https://doi.org/10.1073/pnas.0506580102
PMID:16199517

55. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015; 12:453–57.
https://doi.org/10.1038/nmeth.3337
PMID:25822800
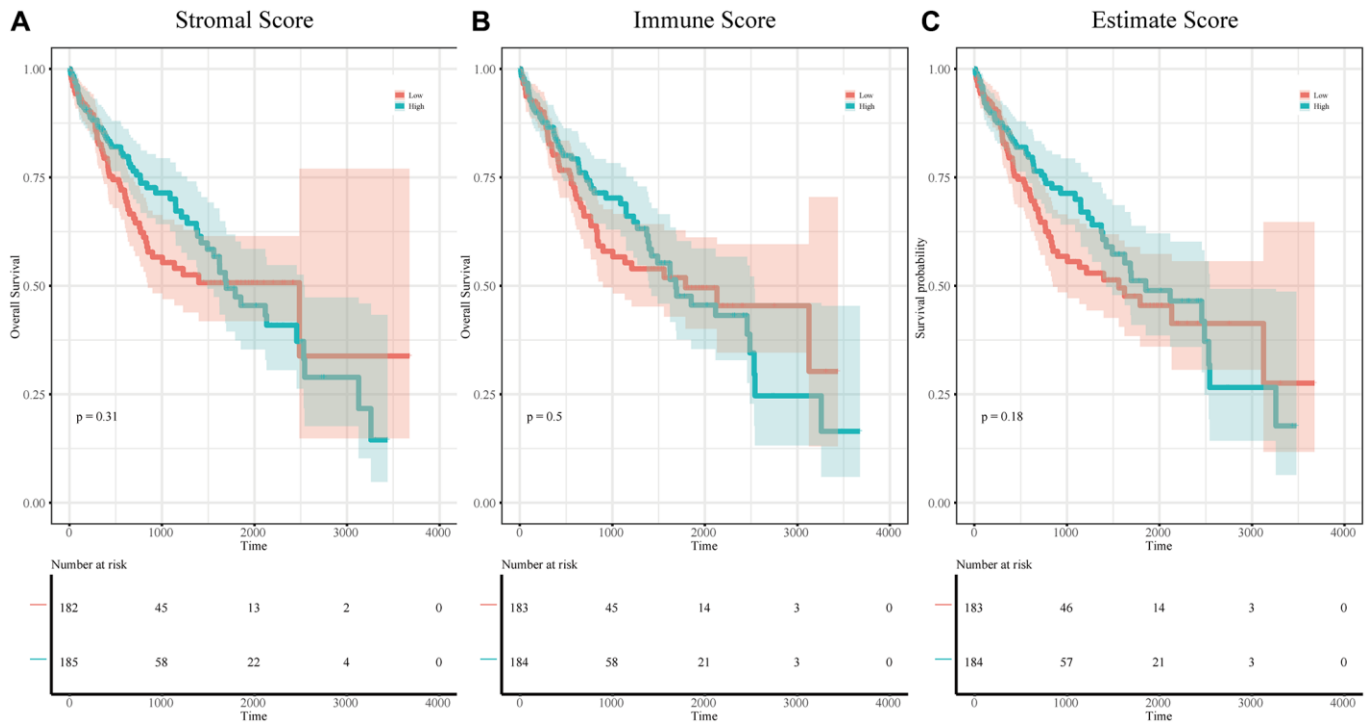
Supplementary Figure 1. The correlation between the scores and the long-term OS of patients with HCC. (**A**) KM survival curves for the long-term OS of low/high stromal score subgroups (p = 0.31). (**B**) KM survival curves for the long-term OS of low/high immune score subgroups (p = 0.5). (**C**) KM survival curves for the long-term OS of low/high estimate score subgroups (p = 0.18).

Supplementary Figure 2. Reactome enrichment analysis of the common DEGs.

| Gene	Coefficient	P value	Hazard Ratio(95% CI)
CXCR6	-0.255	P < 0.001	0.775(0.702-0.856)
IGJ	-0.154	P < 0.001	0.857(0.805-0.912)
CD79A	-0.192	P < 0.001	0.825(0.763-0.892)
CELF2	-0.312	P < 0.001	0.732(0.643-0.833)
ITK	-0.25	P < 0.001	0.779(0.702-0.865)
KLRB1	-0.227	P < 0.001	0.797(0.723-0.878)
CD48	-0.266	P < 0.001	0.766(0.684-0.859)
IKZF1	-0.299	P < 0.001	0.741(0.651-0.844)
ARHGAP15	-0.287	P < 0.001	0.751(0.662-0.851)
HLA-DPA1	-0.287	P < 0.001	0.751(0.662-0.851)
CD3E	-0.218	P < 0.001	0.804(0.73-0.885)
HLA-DOA	-0.269	P < 0.001	0.764(0.679-0.861)
DARC	-0.156	P < 0.001	0.855(0.798-0.917)
NAPSB	-0.194	P < 0.001	0.824(0.756-0.898)
ADAM6	-0.13	P < 0.001	0.878(0.829-0.931)
EVI2B	-0.316	P < 0.001	0.729(0.633-0.84)
CD5	-0.216	P < 0.001	0.806(0.732-0.888)
CCR5	-0.242	P < 0.001	0.785(0.704-0.875)
MS4A7	-0.332	P < 0.001	0.717(0.616-0.833)
MS4A1	-0.181	P < 0.001	0.835(0.769-0.906)
CD3G	-0.229	P < 0.001	0.796(0.718-0.882)
DOCK2	-0.241	P < 0.001	0.786(0.705-0.877)
PRKCB	-0.269	P < 0.001	0.764(0.676-0.863)
SLA	-0.284	P < 0.001	0.753(0.662-0.856)
LY9	-0.236	P < 0.001	0.789(0.709-0.88)
CD226	-0.249	P < 0.001	0.779(0.694-0.875)
LOC96610	-0.185	P < 0.001	0.831(0.762-0.905)
CD2	-0.208	P < 0.001	0.812(0.736-0.895)
GPR132	-0.26	P < 0.001	0.771(0.682-0.872)
CCR2	-0.204	P < 0.001	0.815(0.74-0.898)
CSF2RB	-0.289	P < 0.001	0.749(0.653-0.859)
GBP5	-0.208	P < 0.001	0.812(0.736-0.897)
SIGLEC8	-0.208	P < 0.001	0.812(0.735-0.898)
SLAMF7	-0.215	P < 0.001	0.807(0.728-0.895)
CD69	-0.205	P < 0.001	0.815(0.738-0.9)
ACAP1	-0.19	P < 0.001	0.827(0.754-0.907)
CCL22	-0.197	P < 0.001	0.821(0.746-0.904)
ATP2A3	-0.227	P < 0.001	0.797(0.713-0.89)
CD96	-0.208	P < 0.001	0.813(0.734-0.899)
NCKAP1L	-0.253	P < 0.001	0.776(0.686-0.879)
TAGAP	-0.235	P < 0.001	0.79(0.704-0.887)
CCL21	-0.105	P < 0.001	0.901(0.854-0.949)
FPR3	-0.279	P < 0.001	0.757(0.658-0.87)
CCR7	-0.199	P < 0.001	0.82(0.742-0.906)
BASP1	-0.213	P < 0.001	0.808(0.726-0.899)
CD53	-0.267	P < 0.001	0.766(0.669-0.876)
MS4A6A	-0.28	P < 0.001	0.756(0.656-0.871)
CD6	-0.192	P < 0.001	0.825(0.749-0.909)
HLA-DQA1	-0.208	P < 0.001	0.812(0.731-0.902)
CYTIP	-0.202	P < 0.001	0.817(0.737-0.906)
PLD4	-0.178	P < 0.001	0.837(0.764-0.916)
MGC29506	-0.126	P < 0.001	0.882(0.827-0.94)
LCK	-0.196	P < 0.001	0.822(0.743-0.908)
EPB41L3	-0.25	P < 0.001	0.779(0.685-0.885)
HLA-DQB1	-0.193	P < 0.001	0.825(0.747-0.911)
P2RY13	-0.242	P < 0.001	0.785(0.692-0.891)
PTPRC	-0.202	P < 0.001	0.817(0.735-0.909)
GZMK	-0.161	P < 0.001	0.851(0.782-0.926)
IL7R	-0.135	P < 0.001	0.874(0.814-0.938)
EMB	-0.212	P < 0.001	0.809(0.723-0.905)
WDFY4	-0.2	P < 0.001	0.818(0.736-0.91)
NCF1	-0.201	P < 0.001	0.818(0.734-0.911)
CD84	-0.213	P < 0.001	0.808(0.72-0.907)
FCRL5	-0.154	P < 0.001	0.857(0.788-0.931)
FYB	-0.224	P < 0.001	0.799(0.708-0.902)
CD163	-0.219	P < 0.001	0.803(0.712-0.905)
HAPLN3	-0.176	P < 0.001	0.839(0.761-0.924)
AOAH	-0.198	P < 0.001	0.82(0.735-0.915)
CPZ	-0.108	P < 0.001	0.898(0.845-0.953)
PLA2G2D	-0.128	P < 0.001	0.879(0.818-0.945)
EFEMP1	-0.092	P < 0.001	0.912(0.865-0.962)
IL21R	-0.162	P < 0.001	0.851(0.774-0.935)
AEBP1	-0.123	P < 0.001	0.884(0.822-0.951)
SIGLEC10	-0.196	0.001	0.822(0.731-0.924)
ITGA4	-0.208	0.001	0.812(0.717-0.92)
MNDA	-0.187	0.001	0.83(0.742-0.928)
HLA-DQA2	-0.146	0.001	0.864(0.79-0.944)
SMOC2	-0.134	0.001	0.875(0.806-0.949)
CD1C	-0.149	0.001	0.862(0.787-0.944)
MFAP4	-0.091	0.001	0.913(0.863-0.966)
CLEC7A	-0.196	0.002	0.822(0.728-0.928)
SHISA3	-0.115	0.002	0.891(0.83-0.957)
TFEC	-0.212	0.002	0.809(0.709-0.923)
HLA-DRB5	-0.153	0.002	0.858(0.78-0.944)
KLHL6	-0.209	0.002	0.811(0.712-0.925)
RUNX2	-0.17	0.002	0.843(0.757-0.94)
ADAMDEC1	-0.118	0.002	0.888(0.824-0.958)
CCL19	-0.083	0.002	0.92(0.873-0.97)
CYBB	-0.185	0.002	0.831(0.738-0.936)
DCN	-0.089	0.002	0.915(0.864-0.969)
ALOX5	-0.147	0.002	0.863(0.785-0.949)
GPR68	-0.178	0.002	0.837(0.746-0.939)
HLA-DQB2	-0.135	0.002	0.874(0.801-0.954)
NLRP3	-0.194	0.003	0.824(0.726-0.934)
IL2RA	-0.151	0.003	0.86(0.779-0.949)
GAPT	-0.172	0.003	0.842(0.753-0.942)
CD1E	-0.135	0.003	0.874(0.8-0.955)
CLEC4E	-0.149	0.003	0.862(0.781-0.95)
LILRB4	-0.174	0.003	0.841(0.749-0.943)
CR1	-0.147	0.003	0.863(0.783-0.952)
APOBEC3C	-0.173	0.004	0.841(0.748-0.946)
RSPO3	-0.117	0.004	0.889(0.821-0.964)
SAMSN1	-0.173	0.004	0.841(0.747-0.948)
DES	-0.114	0.005	0.892(0.824-0.966)
SIRPB2	-0.158	0.005	0.854(0.765-0.953)
TIMD4	-0.126	0.005	0.882(0.807-0.963)
WISP2	-0.097	0.005	0.908(0.848-0.972)
LOXL1	-0.109	0.006	0.897(0.83-0.968)
PRELP	-0.097	0.006	0.908(0.848-0.972)
GATA3	-0.145	0.006	0.865(0.781-0.959)
TLR7	-0.157	0.007	0.854(0.763-0.957)
TLR8	-0.138	0.007	0.871(0.789-0.963)
GPR183	-0.151	0.007	0.86(0.77-0.96)
ALOX5AP	-0.133	0.008	0.876(0.794-0.966)
GPBAR1	-0.136	0.008	0.873(0.788-0.966)
SVEP1	-0.086	0.009	0.917(0.86-0.978)
LUM	-0.067	0.009	0.935(0.889-0.983)
CCL2	-0.133	0.009	0.875(0.792-0.967)
TNFSF13B	-0.152	0.009	0.859(0.767-0.963)
HGF	-0.108	0.01	0.897(0.827-0.974)
FNDC1	-0.081	0.01	0.922(0.867-0.981)
PLAC8	-0.117	0.011	0.89(0.814-0.973)
CXCL14	-0.078	0.011	0.925(0.871-0.982)
PTGDS	-0.062	0.011	0.94(0.896-0.986)
GFPT2	-0.095	0.012	0.909(0.844-0.979)
CCDC80	-0.086	0.012	0.918(0.859-0.981)
PODN	-0.082	0.013	0.921(0.863-0.983)
TMEM119	-0.094	0.014	0.91(0.844-0.981)
F13A1	-0.108	0.015	0.898(0.823-0.979)
GLIPR2	-0.162	0.016	0.85(0.745-0.97)
ITGBL1	-0.075	0.016	0.928(0.873-0.986)
DPT	-0.068	0.017	0.934(0.883-0.988)
THBS2	-0.076	0.018	0.927(0.87-0.987)
EGR2	-0.109	0.021	0.897(0.817-0.983)
FPR1	-0.116	0.022	0.89(0.806-0.983)
DPEP1	-0.076	0.022	0.927(0.869-0.989)
ISLR	-0.071	0.022	0.931(0.877-0.99)
HAND2	-0.074	0.023	0.929(0.871-0.99)
ASPG	-0.049	0.024	0.952(0.912-0.993)
COL8A2	-0.098	0.025	0.907(0.832-0.988)
RGS1	-0.094	0.026	0.91(0.838-0.989)
MARCO	-0.071	0.026	0.931(0.875-0.992)
CADM3	-0.097	0.027	0.908(0.833-0.989)
GAS7	-0.139	0.032	0.871(0.767-0.988)
PDPN	-0.072	0.04	0.931(0.869-0.997)
ADAMTS2	-0.079	0.04	0.924(0.857-0.996)
ADRA2A	-0.074	0.041	0.929(0.865-0.997)
MOXD1	-0.067	0.043	0.936(0.877-0.998)
PDGFRA	-0.06	0.044	0.942(0.889-0.998)
FMOD	-0.078	0.044	0.925(0.857-0.998)
ADAM28	-0.089	0.045	0.915(0.839-0.998)
FCGR1A	-0.096	0.045	0.908(0.827-0.998)

Hazard Ratio

Supplementary Figure 3. Forest plot of the univariate COX regression analysis for RFS.

Supplementary Figure 4. HLA-DRB5 expression in HCC tumors and paracarcinoma tissues, and the correlation between HLA-DRB5 and survival/clinical characteristics of patients with HCC. (**A**) HLA-DRB5 expression in HCC tumor tissues and tumor-adjacent normal tissues (p = 0.138). (**B**) HLA-DRB5 expression in paired HCC tumor tissues and tumor-adjacent normal tissues deriving from the same patient (p = 0.040). (**C**) Overall survival of HLA-DRB5 high and low expression subgroups (p = 0.12). (**D**) Recurrence-free survival of HLA-DRB5 high and low expression subgroups (p = 0.22). (**E–H**) The correlation of HLA-DRB5 expression with clinicopathological staging characteristics.
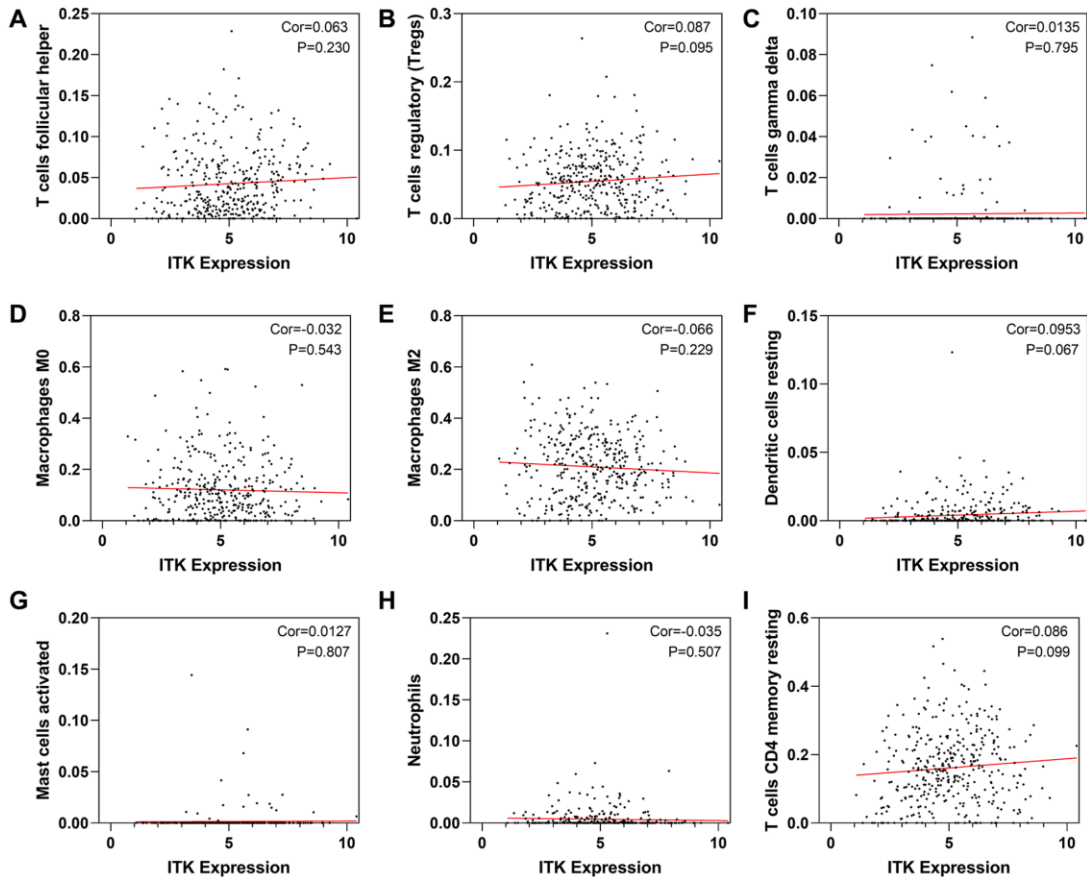
Supplementary Figure 5. Nine TICs not correlated with ITK expression.
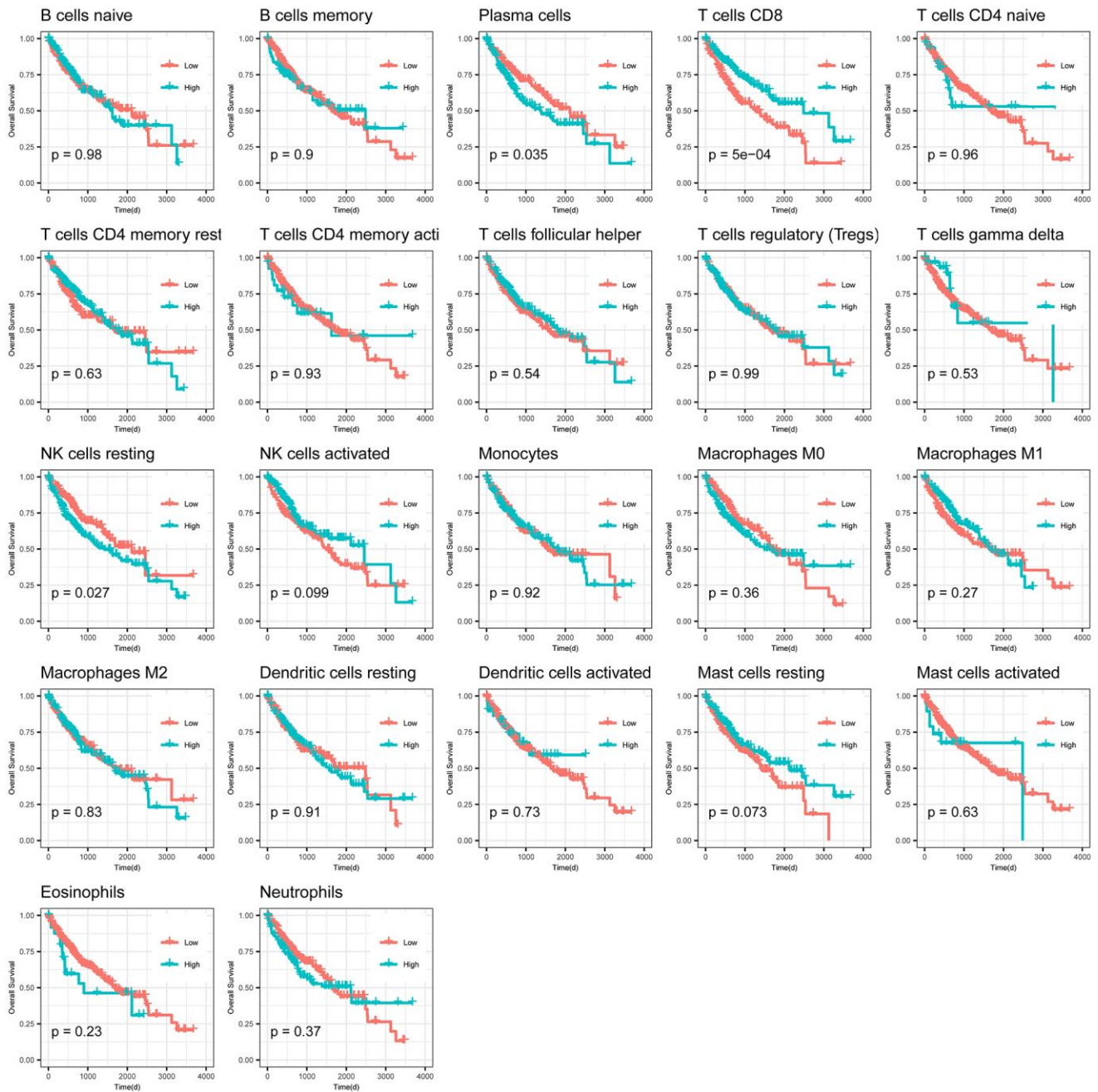
Supplementary Figure 6. KM survival curves for the long-term OS of low/high groups defined by 22 different immune cell subsets.
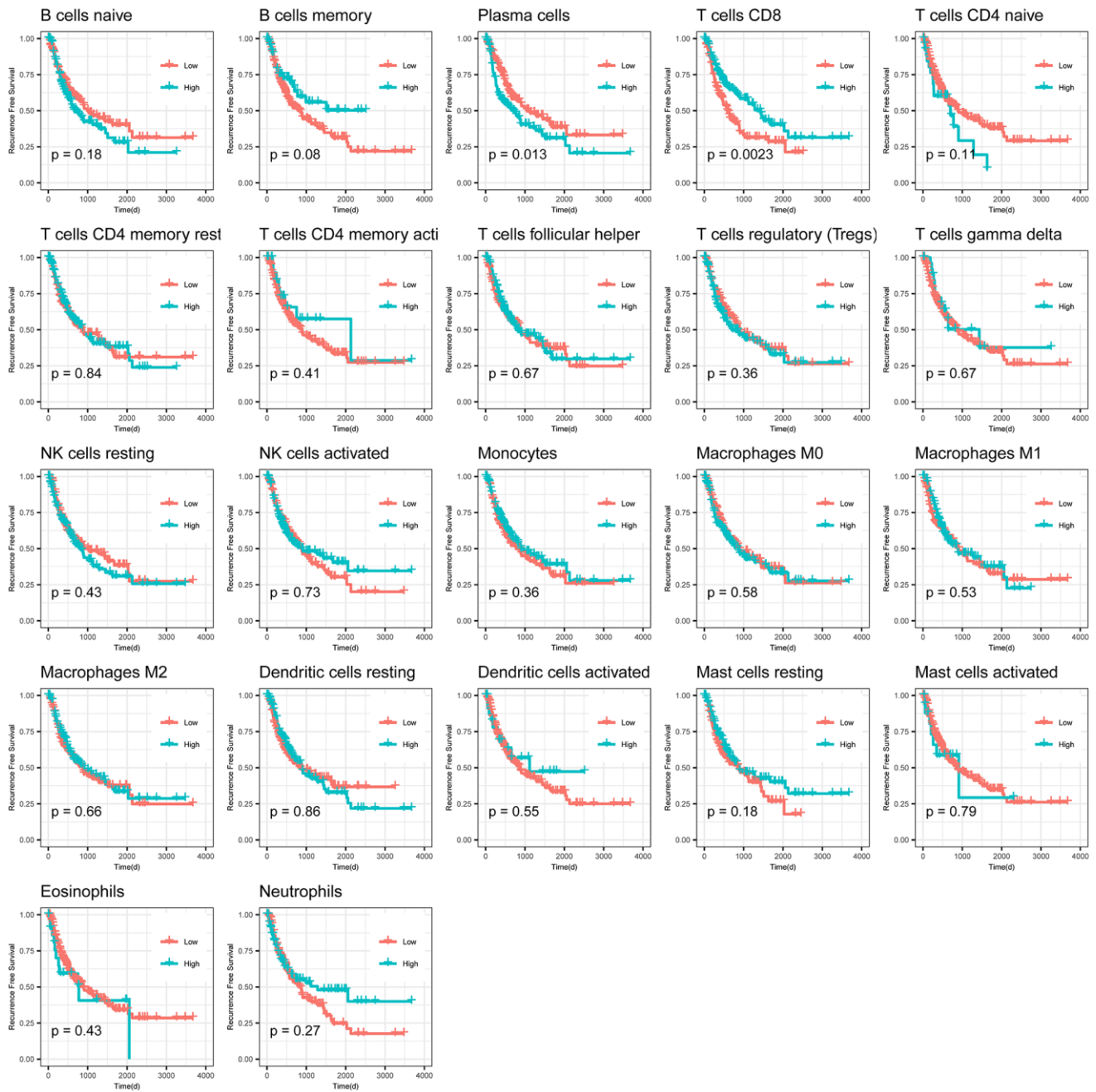
Supplementary Figure 7. KM survival curves for the RFS of low/high groups defined by 22 different immune cell subsets.

Supplementary Tables

Supplementary Table 1. Clinicopathological characteristics statistics of LIHC patients from TCGA.

Clinical characteristics		Total	%
		373	100
Age at diagnosis (y)	young age (< = 60)	177	47.5
	old age (>60)	195	52.3
Gender	Male	252	67.6
	Female	121	32.4
Stage	I	172	46.1
	II	87	23.3
	III	85	22.8
	IV	5	1.3
T classification	T1	182	48.8
	T2	95	25.5
	T3	80	21.4
	T4	13	3.5
	TX	2	0.5
M classification	M0	267	71.6
	M1	4	1.1
	MX	102	27.3
N classification	N0	253	67.8
	N1	4	1.1
	NX	116	31.1

Supplementary Table 2. Gene list of genes identified by MCODE and univariate cox analysis.

Table A: Gene list of hub genes identified by MCODE ($n = 16$)

CCL21	CXCL8	CCR5	CCL19	CCR2	CXCL1	CD3E	LCK
ITK	CD3G	HLA-DQA2	HLA-DQB2	HLA-DRB5	HLA-DQA1	HLA-DQB1	HLA-DPA1

Table B: Genes statistically significant in univariate COX analysis of RFS ($n = 152$)

CXCR6	IGJ	CD79A	CELF2	ITK	KLRB1	CD48	IKZF1
ARHGAP15	HLA-DPA1	CD3E	HLA-DOA	DARC	NAPSB	ADAM6	EVI2B
CD5	CCR5	MS4A7	MS4A1	CD3G	DOCK2	PRKCB	SLA
LY9	CD226	LOC96610	CD2	GPR132	CCR2	CSF2RB	GBP5
SIGLEC8	SLAMF7	CD69	ACAP1	CCL22	ATP2A3	CD96	NCKAP1L
TAGAP	CCL21	FPR3	CCR7	BASP1	CD53	MS4A6A	CD6
HLA-DQA1	CYTIP	PLD4	MGC29506	LCK	EPB41L3	HLA-DQB1	P2RY13
PTPRC	GZMK	IL7R	EMB	WDFY4	NCF1	CD84	FCRL5
FYB	CD163	HAPLN3	AOAH	CPZ	PLA2G2D	EFEMP1	IL21R
AEBP1	SIGLEC10	ITGA4	MNDA	HLA-DQA2	SMOC2	CD1C	MFAP4
CLEC7A	SHISA3	TFEC	HLA-DRB5	KLHL6	RUNX2	ADAMDEC1	CCL19
CYBB	DCN	ALOX5	GPR68	HLA-DQB2	NLRP3	IL2RA	GAPT
CD1E	CLEC4E	LILRB4	CR1	APOBEC3C	RSPO3	SAMSN1	DES
SIRPB2	TIMD4	WISP2	LOXL1	PRELP	GATA3	TLR7	TLR8

GPR183	ALOX5AP	GPBAR1	SVEP1	LUM	CCL2	TNFSF13B	HGF
FNDC1	PLAC8	CXCL14	PTGDS	GFPT2	CCDC80	PODN	TMEM119
F13A1	GLIPR2	ITGBL1	DPT	THBS2	EGR2	FPR1	DPEP1
ISLR	HAND2	ASPG	COL8A2	RGS1	MARCO	CADM3	GAS7
PDPN	ADAMTS2	ADRA2A	MOXD1	PDGFRA	FMOD	ADAM28	FCGR1A

Table C: Genes statistically significant in univariate COX analysis of OS ($n = 27$)

KLRB1	IL8	DARC	MMP7	CTHRC1	MSC	EFNA5	ITK
CCR7	GZMK	CXCR6	CXCL1	MS4A1	IL7R	CD79A	ATP2A3
CD5	ACAP1	CD69	IGJ	HLA-DRB5	CD6	CXCL6	MGC29506
LY9	CREB3L1	CD2					

Supplementary Table 3. Enriched gene sets.

MSigDB collection	Description	NES	adj.p.value	q.value
h.all.v7.2.symbols.gmt				
ITK high expression	HALLMARK_INFLAMMATORY_RESPONSE	1.5482	0.0005034	0.0003921
	HALLMARK_ALLOGRAFT_REJECTION	1.4715	0.005095	0.0039687
	HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION	1.4536	0.0055532	0.0043256
	HALLMARK_INTERFERON_GAMMA_RESPONSE	1.4195	0.0105352	0.0082064
	HALLMARK_IL2_STAT5_SIGNALING	1.3947	0.016093	0.0125356
	HALLMARK_IL6_JAK_STAT3_SIGNALING	1.4786	0.016093	0.0125356
	HALLMARK_TNFA_SIGNALING_VIA_NFKB	1.4019	0.019148	0.0149153
	HALLMARK_COMPLEMENT	1.3644	0.0281044	0.0218918
	HALLMARK_HEME_METABOLISM	1.328	0.0376875	0.0293566
c7.all.v7.2.symbols.gmt				
ITK high expression	GSE4984_UNTREATED_VS_GALECTIN1_TREATED_DC_DN	1.7377	2.37E-06	1.76E-06
	GSE22886_NAIVE_CD8_TCELL_VS_MONOCYTE_UP	1.678	9.87E-06	7.33E-06
	GSE3039_NKT_CELL_VS_ALPHAALPHA_CD8_TCELL_DN	1.6849	1.18E-05	8.75E-06
	GSE7218_UNSTIM_VS_ANTIGEN_STIM_THROUGH_IGG_BCELL_DN	1.6628	7.78E-05	5.77E-05
	GSE3039_ALPHAALPHA_VS_ALPHABETA_CD8_TCELL_DN	1.6203	0.0001806	0.0001341
	GSE29618_BCELL_VS_MDC_UP	1.6334	0.0002267	0.0001683
	GSE11057_PBMC_VS_MEM_CD4_TCELL_UP	1.6131	0.0002267	0.0001683
	GSE45739_UNSTIM_VS_ACD3_ACD28_STIM_WT_CD4_TCELL_DN	1.5972	0.0007672	0.0005696

Abbreviation: NES: normalized enrichment score; p.adjust: adjusted p.value; Gene sets with p.adjust less than 0.05 and q.value less than 0.05 were considered as statistical significance. Only several leading sets enriched in ITK high expression both in C7 were listed here due to a large number of enriched gene sets.

Supplementary Table 4. TICs co-determined by difference test and correlation test.

TICs	Pearson r	*P* (two-tailed)	Difference test (*p*-value)
Plasma cells	−0.269	<0.0001	0.003
T cells CD8	0.4274	<0.0001	<0.001
T cells CD4 naive	−0.2935	<0.0001	<0.001
T cells CD4 memory activated	0.2237	<0.0001	0.002
NK cells resting	−0.2056	<0.0001	0.003
NK cells activated	−0.2193	<0.0001	<0.001
Macrophages M1	0.3197	<0.0001	<0.001
Dendritic cells activated	−0.1351	0.0092	0.014
Mast cells resting	−0.1769	0.0006	0.011