Malignant cell receptor-ligand subtypes guide the prediction of prognosis and personalized immunotherapy of liver cancer

Junzheng Wu[1,*], Chuncheng Wu[1,*], Xianhui Cai[2,*], Peipei Li[1], Jianjun Lin[1], Fuqiang Wang[1]

[1]Xiamen Hospital of Traditional Chinese Medicine, Xiamen Hospital, Beijing University of Chinese Medicine, Xiamen, Fujian, China
[2]Xiamen Xianyue Hospital, Xiamen, Fujian, China
[*]Co-first authors

Correspondence to: Fuqiang Wang; **email:** wangfq@xmu.edu.cn

ABSTRACT

Objective: Liver cancer is a prevalent disease with a dismal prognosis. The aim of the research is to identify subgroups based on malignant cell receptor ligand gene from single-cell RNA, which might lead to customized immunotherapy for patients with liver cancer.
Methods: Based on scRNA-seq data, we identified the receptor-ligand genes associated with prognosis and classify patients into molecular subtypes by univariate Cox regression and consensus clustering. LASSO regression was performed to construct a prognostic model, which was validated in TCGA and ICGC datasets. Immune infiltration and prediction of immunotherapy response were analyzed using ssGSEA, ESTIMATE, TIDE, and TRS score calculation. Finally, qPCR and Western blot validation of key genes and protein levels in cell lines.
Results: A risk model using 16-gene expression levels predicted liver cancer patients' prognosis. The RiskScore associated significantly with tumor clinical characteristics and immunity, integrated with clinicopathological features for survival prediction. Differential expression of SRXN1 was verified in hepatocellular carcinoma and normal liver cells.
Conclusion: Our study utilizes single-cell analysis to investigate the communication between malignant cells and other cell types, identifying molecular subtypes based on malignant cell receptor ligand genes, offering new insights for the development of personalized immunotherapy and prognostic prediction models.

INTRODUCTION

Hepatocellular carcinoma (HCC), ranking sixth highest in incidence and third in mortality rates, constitutes a significant global public health concern [1]. Its prognosis is often bleak, as the disease is frequently detected in its advanced stages, contributing to a five-year survival rate that is disappointingly low [2]. The complex nature of HCC, along with the heterogeneity amongst patients, presents substantial challenges to effective treatment strategies, with no universally effective therapy currently available [3]. Hence, the exploration and investigation of HCC carries profound significance for the health of populations worldwide.

A variety of therapeutic strategies are utilized in the treatment of HCC, including surgical interventions [1], radiofrequency ablation (RFA) or microwave ablation (MWA), systemic chemotherapy, Radiation therapy, and targeted therapies [4]. Recent advancements have also been made in immunotherapies, notably immune checkpoint inhibitors such as nivolumab and pembro-lizumab [5]. Immune checkpoint inhibitors have been studied as a promising treatment option for liver cancer,

but finding reliable biomarkers to predict response remains challenging [6]. The utilization of these medicines is currently limited to patients with advanced HCC, and they show variable responses with some patients achieving substantial durable responses, while many others do not respond or develop resistance after an initial response.

Currently, the malignant cell receptor ligand genes refer to those genes that encode proteins serving as ligands for receptors found on HCC cells. These ligands can bind to their corresponding receptors, leading to the activation of signaling pathways involved in the progression and survival of malignant cells [7]. Ligand-receptor interactions in liver cancer profoundly affect the immune microenvironment through various mechanisms in various studies. One previous study revealed that tumor cells continuously produce kynurenic acid (Kyn), an endogenous ligand for the aryl hydrocarbon receptor (AHR), via tryptophan-2,3-dioxygenase (TDO). Consequently, Kyn, acting through AHR, suppresses anti-tumor immune responses and promotes tumor cell survival and motility [8]. Furthermore, the transforming growth factor-beta (TGFβ) signaling pathway, known for its critical role in tumor development, influences processes like cell invasion and immune modulation [9]. Interestingly, the genetic characteristics of tumors are correlated with immune cell infiltration and the presence of neoantigens [10].

The objective of this study is to identify molecular subtypes of liver cancer based on malignant cell receptor ligand genes, and to develop a prognostic model for patients with liver cancer. To improve personalized immunotherapy, molecular subtypes based on malignant cell receptor ligand genes have been proposed as a potential approach.

RESULTS

Single-cell RNA clustering

After rigorous quality control, we retained 24,329 cells, comprising eight cell subpopulations (Figure 1A). The number of UMIs and mRNAs showed a significant correlation, while the number of UMI/mRNAs was not significantly correlated with the content of mito-chondrial genes (Supplementary Figure 1A). Supplementary Figure 1B, 1C depict the violin plots before and after quality control, and Supplementary Figure 1D demonstrates significant elimination of batch effects between samples.

UMAP dimensionality reduction analysis was performed on 24,329 cells using RunUMAP function, revealing the presence of eight cell types: B cell,

Endothelial cells, Hepatocytes, Macrophages, Mono-cytes, NK cell, Fibroblasts, and T cells. Aneuploid cells were included in cell cluster marker identification for ten cell types. We further analyzed the 10 samples based on the proportion of the ten subpopulations and identified the location of aneuploid malignant cells on the UMAP (Figure 1B). The top-five marker genes with the most outstanding contribution in each subgroup are shown in Figure 1C, and the results of marker genes are provided in Supplementary Table 1. The differential gene enrichment analysis of each subpopulation of cells is shown in Figure 1D, and the aneuploid cells differential genes were mainly enriched in the complement and coagulation cascades, cholesterol metabolism, PPAR signaling pathway, and pentose and glucuronate interconversion.

Cellchat analysis and construction of cellular communication network

After rigorous quality control, we retained 49,701 cells, comprising 22,920 cancer cells and 26,781 normal liver tissue cells. Uniform manifold approximation and projection (UMI) and the number of mRNAs showed significant correlation, while there was no significant correlation between number of UMI/mRNAs and the content of mitochondrial genes, as seen in Supplementary Figure 1A. The violin plot before and after quality control can be seen in Supplementary Figure 1B–1D demonstrates significant elimination of batch effects between samples.

We used copykat to identify malignant cells and analyzed the intersubpopulation communication network using the CellChat (version 1.5.0) package. Results showed that the communication between malignant cells and macrophages had the highest number and intensity (Figure 2A, 2B). Additionally, visual analysis revealed that the first six signaling pathways (MIF, MHC-II, APP, MHC-I) of subgroup communication had signal outputs of malignant cells, with the MIF signaling pathway having the strongest output and Monocyte being the main signal receiver cell (Figure 2C, 2D).

Construction of molecular subtypes based on malignant cell receptor ligand genes

To analyze the role of malignant cell-associated receptors in Bulk RNA-seq data, we screened 72 malignant cell-associated receptor genes using the results of the single-cell dataset CellChat. We found that 7 of the 72 related genes were associated with prognosis ($p < 0.05$), and protective gene 3 and risk gene 4 are shown in Figure 3A. We observed a relatively stable clustering result for Cluster 2 in the cumulative distribution function (CDF) Delta area

curve (Figure 3B, 3C) and selected K = 2 to obtain two molecular subtypes (Figure 3D). Furthermore, analysis of the prognostic characteristics of these two molecular subtypes revealed a significant difference in prognosis (Figure 3E), with intersecting clust2 and clust1 subtypes having worse prognosis. The tcga. Subtype. CLI. Txt contains data for the TCGA dataset subtype. We performed the same clustering method on the independent ICGC-JP dataset and obtained similar results (Figure 3F). The icgc. Subtype. CLI table provides data on subtypes in the ICGC-JP dataset.

Mutation characteristics of molecular subtypes

We obtained SNV mutation data from TCGA using Mutect2, as shown in Figure 4A, and presented the top 15 genes with the most significant mutations in each subtype. Moreover, we surveyed "Homologous
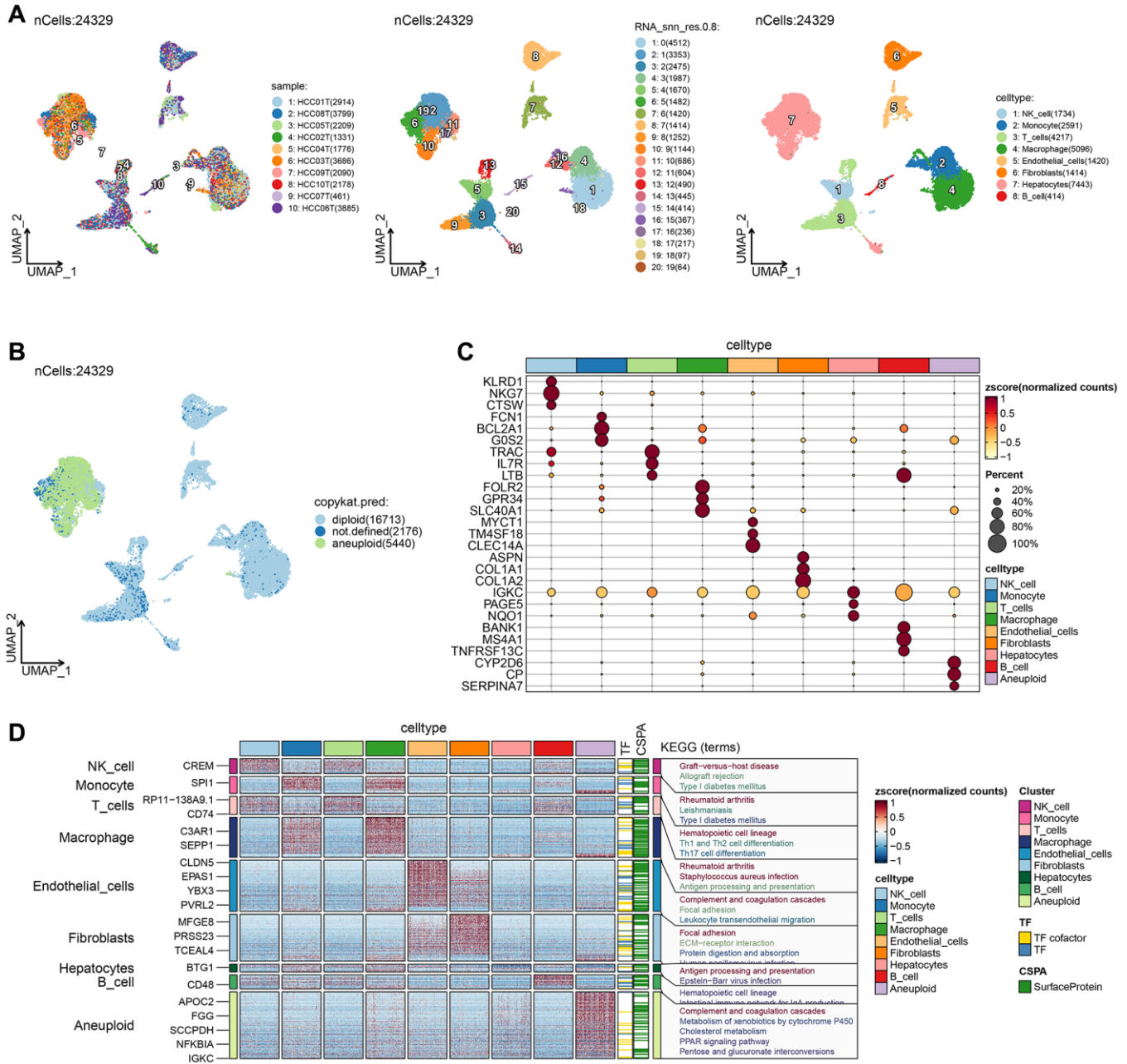

Figure 1. Single-cell RNA clustering. (**A**) UMAP plot of the distribution of 10 samples, UMAP plot of the distribution of 20 subpopulations, and UMAP plots of the subpopulations after cell annotation. (**B**) After CNV comment, the proportion and cell number of the subpopulation in each sample, and the UMAP map of malignant cells. (**C**) Dot plots of the expression of the first 5 marker genes of the subpopulations after CNV annotation. (**D**) The differential gene enrichment analysis of each subpopulation of cells.

Recombination Defects", "Fraction Altered", "Number of Segments", "Nonsilent Mutation Rate", and "Aneuploidy Score," as well as the distribution of the Silent Mutation Rate. We noticed that the Homologous Recombination Defects and Aneuploidy scores varied between subtypes, as demonstrated in Figure 4B (PMC5982584).

Pathway analysis of molecular subtypes

We involved performing differential analysis for TCGA gene expression profiles grouped by cluster subtypes. This analysis revealed that when comparing clust1 to clust2, the expression of 188 genes increased while 1083 genes exhibited decreased expression (Figure 5A). Using the Hallmark gene sets in the Misgbd database as the background set, we performed a difference analysis GSEA (Figure 5B). The analysis results revealed that the Peroxisome and Histidine metabolism were activated in the clust1 subtype, whereas *Salmonella* infection, Shigellosis, and Epithelial cell signaling in *Helicobacter pylori* infection were inhibited, as demonstrated in Figure 5C. Our next step utilized GSVA pathway scoring on KEGG pathway enrichment obtained from the Misgbd database. Our analysis showed that 33 pathways were up-regulated and 23 were down-regulated, as demonstrated in Figure 5D.
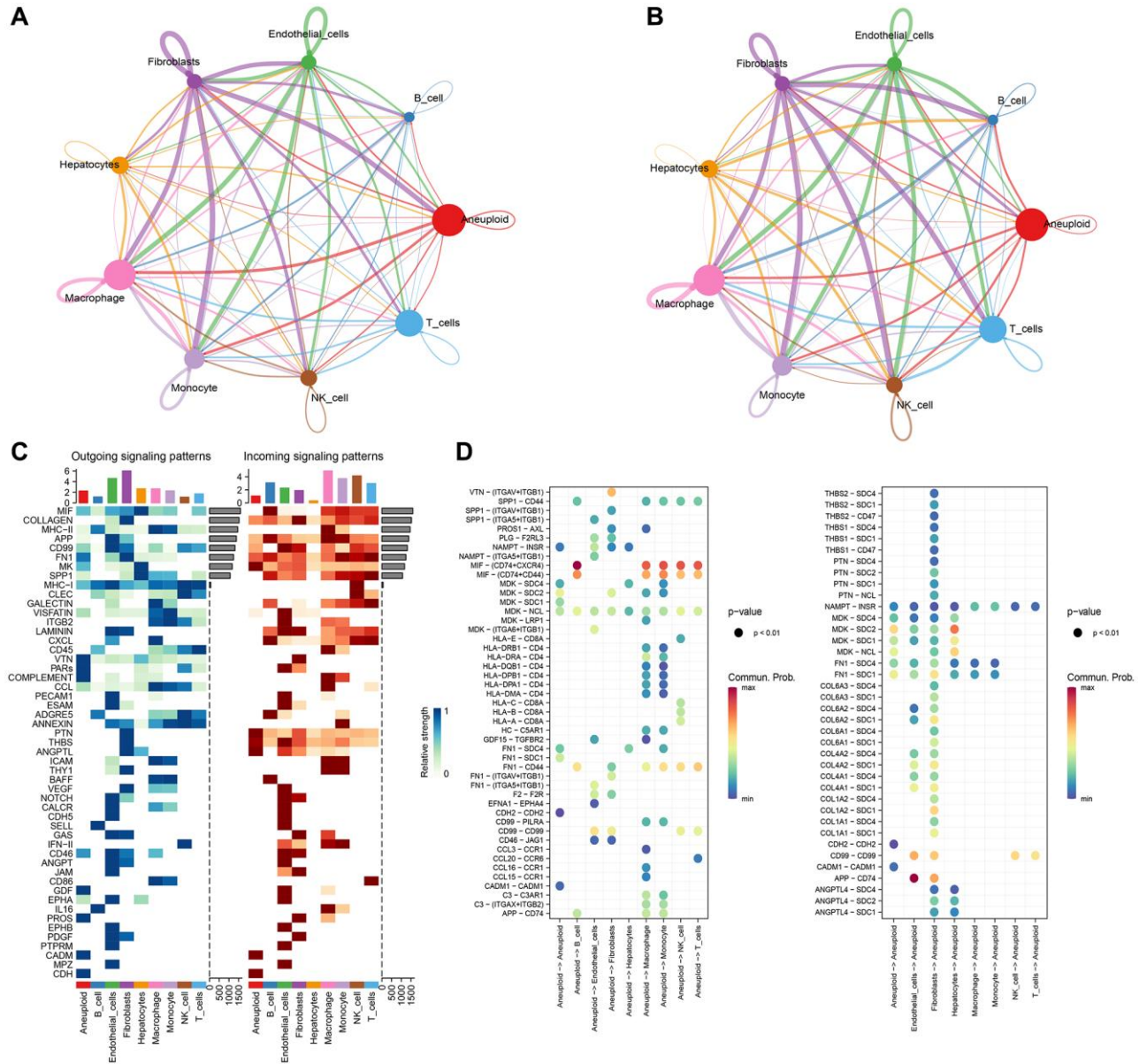

Figure 2. CellChat analysis and construction of cellular communication network. (**A**) Circle diagrams of that numb and strength of subpopulation interaction. (**B**) Number of communication interactions of each subgroup. (**C**) Communication and interaction strength of each subgroup. (**D**) Aneuploidy cell subpopulations and ligand receptor profiles among cell subpopulations.
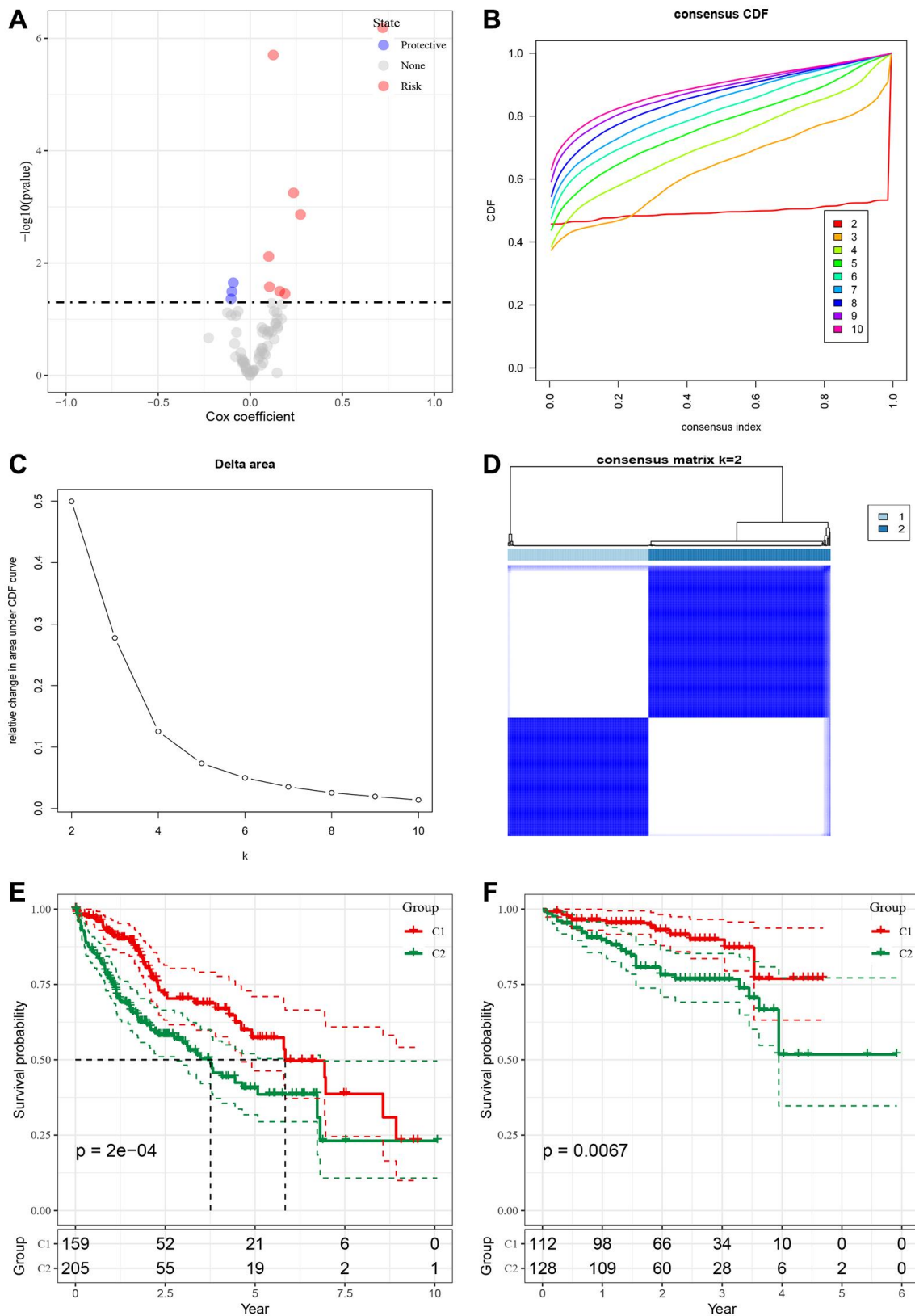
Figure 3. Construction of molecular subtypes based on malignant cell receptor ligand genes. (**A**) Single factor Cox analysis HR distribution map of malignant cell associated receptor gene set. (**B**) CDF curve of TCGA cohort sample. (**C**) CDF Delta area curve of TCGA cohort sample, the horizontal axis represents the number of clusters K, and the vertical axis represents the relative change of the area under the CDF curve. (**D**) Sample cluster heat map of consensus K = 2. (**E**) KM curve of prognosis relationship between two subtypes of TCGA. (**F**) KM curve of prognosis relationship between two subtypes of ICGC-JP.

Immunological characteristics of molecular subtypes and differences in immunotherapy/chemotherapy

To determine the differences in immune micro-environment among molecular subtypes, we evaluated the infiltration level of immune cells using the expression levels of genes in immune cells in the TCGA cohort. We calculated the scores of immune cells using CIBERSORT, EPIC, Estlmate, MCPcounter, quanTlseg, TIMER, and xCell. Results indicated that the majority of immune cells had noteworthy differences between the two subtypes, with higher
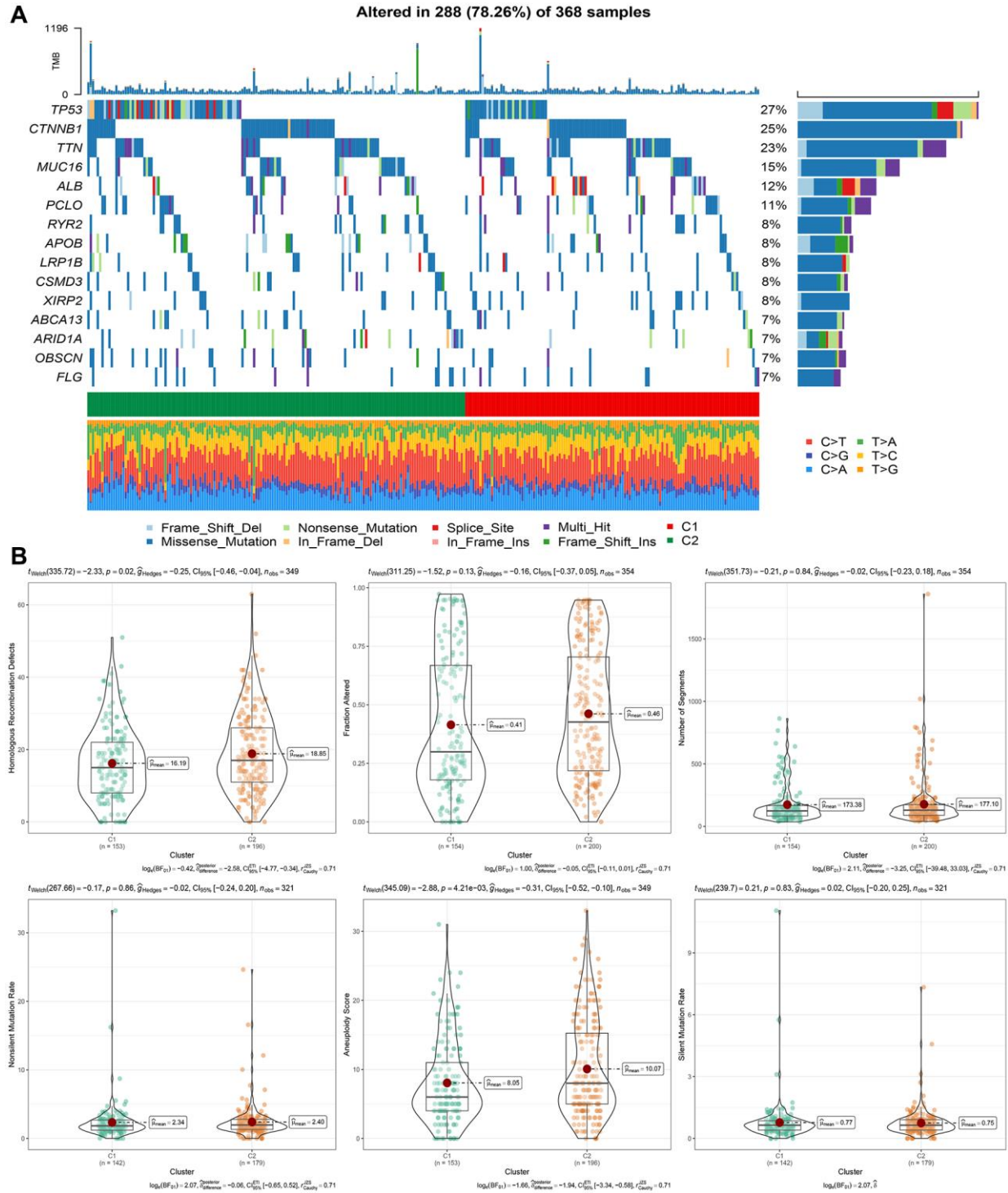

Figure 4. Genomic alterations in molecular subtypes of the TCGA cohort. (**A**) Somatic mutation analysis of different molecular subtypes in the TCGA cohort. (**B**) Comparison of "Homologous Recombination Defects", "Fraction Altered", "Number of Segments" "in different molecular subtypes of the TCGA cohort". "Nonsilent Mutation Rate", "Aneuploidy Score", "difference in Silent Mutation Rate".

scores observed in clust2 (Figure 6A), which indicates that clust2 (with poor prognosis) exhibited a high immune score. We also looked at the expression levels of immune checkpoint genes in the two subtypes and discovered that there were significant variations between the two subtypes, with clust1 having a higher expression level than clust2 (Figure 6B). Patients with high TIDE prediction scores are unlikely to benefit from immunotherapy, indicating a higher probability of immune escape. A Wilcox Test on Figure 6B showed that the TIDE score was highest in the TCGA cohort's clust2 subtype, suggesting that this subtype has an elevated likelihood of immune escape and is less likely to benefit from immunotherapy.

Construction of risk model

We conducted Cox regression analysis on the differentially expressed genes among the cluster subtypes and identified 285 prognostically relevant genes present in both the TCGA and ICGC datasets. Using Lasso regression, we compressed the 285 genes further in the TCGA dataset by reducing the number of genes present in the risk model. The change trajectory of each independent variable was analyzed to determine the optimal model based on the lambda parameter; the figure (Figure 7A) indicates that with the gradual increase of lambda, the number of independent variable coefficients tending to 0 also increases gradually. We used a 10-fold cross-validation technique to build the model and analyzed the confidence interval for each lambda parameter, as shown in Figure 7B. The optimal lambda value was found to be 0.04711966, and 16 genes were selected as target genes based on this value. Based on these 16 genes related to prognosis, nine genes (LPCAT1, SLC2A1, NEIL3, SRXN1, TRNP1, SLC7A11, STC2, ZNF239, CBX2, CDCA8, EPO, PON1, PBK, PFN2, ACOT12, ADH4) were identified
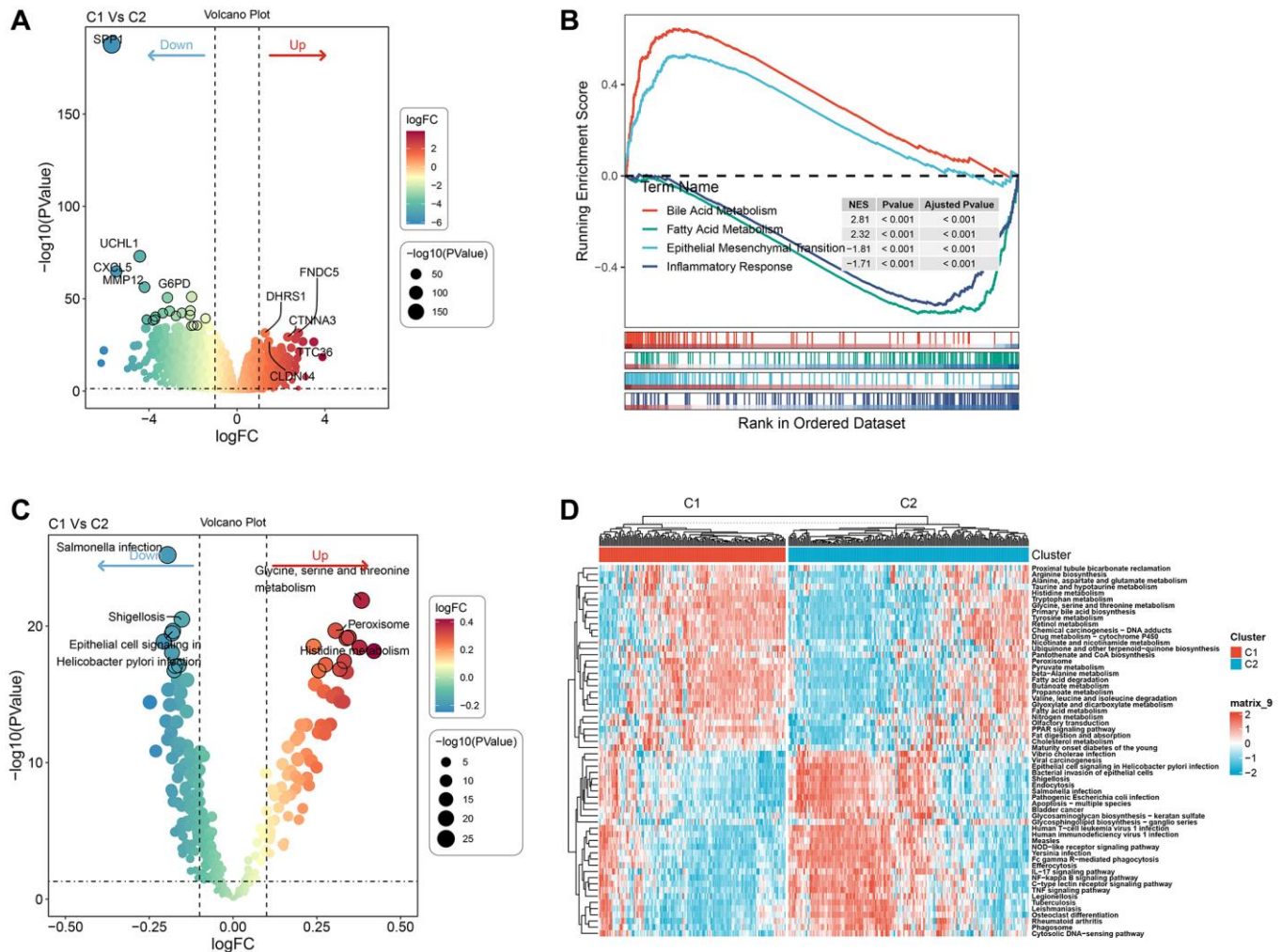

Figure 5. Pathway analysis of molecular subtypes. (**A**) Volcano map of differential analysis of genes between the two subtypes. (**B**) GSEA-GO analysis of differential genes between the two groups. (**C**) Bubble map of related pathways activated/inhibited in the comparison of the two subtypes. (**D**) Heat map of KEGG-related pathways with differences between the two subtypes.

using Lasso regression. We developed a final 16-gene signature with a Cox proportional hazards regression model with the RiskScore calculated as follows:

RiskScore = $-0.205 \times$ LPCAT1 $-0.068 \times$ SLC2A1 $-0.194 \times$ NEIL3 + $0.508 \times$ SRXN1 + $0.117 \times$ TRNP 1 + $0.29 \times$ SLC7A11 + $0.209 \times$ STC2 + $0.087 \times$ ZNF239 + $0.25 \times$ CBX2 + 0.111CDCA8 + 0.332EPO + 0.141PON1 + 0.377PBK + 0.186PFN2 + -0.175ACOT12 + $-0.052 \times$ ADH4.

We used the TCGA dataset as our training dataset and calculated the risk score of each sample through 16 gene expression levels to analyze the prognostic
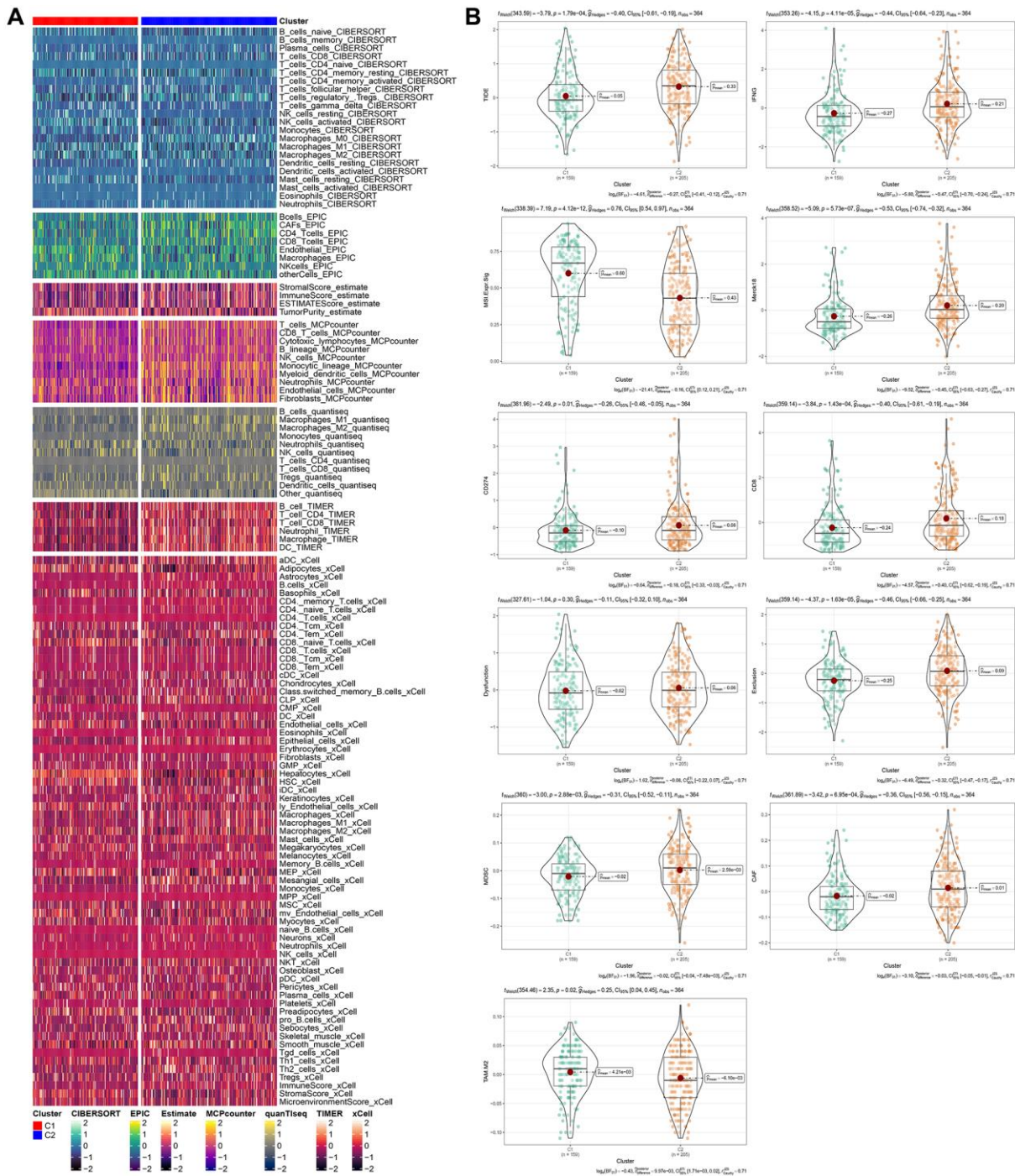

Figure 6. Immunological characteristics of molecular subtypes and differences in immunotherapy/chemotherapy. (**A**) The difference of 28 immune cell scores between different molecular subtypes in the TCGA cohort. (**B**) The difference of ESTIMATE immune infiltration between different molecular subtypes in the TCGA cohort. The immune checkpoint of differential expression between different subgroups in the TCGA cohort. Difference of TIDE analysis results between different groups in TCGA queue.

prediction efficiency classification at 1, 3 and 5 years. The AUC for each year was 0.7. The Riskscore was standardized by zscore, and samples with Riskscore greater than zero after zscore were divided into a high-risk group and a low-risk group.

The KM curve was drawn, and we found a significant difference between them ($p < 0.05$) (Figure 7C, 7D). To validate the accuracy of our model, we verified the ICGC dataset using the same method and found similar results (Figure 7E, 7F).
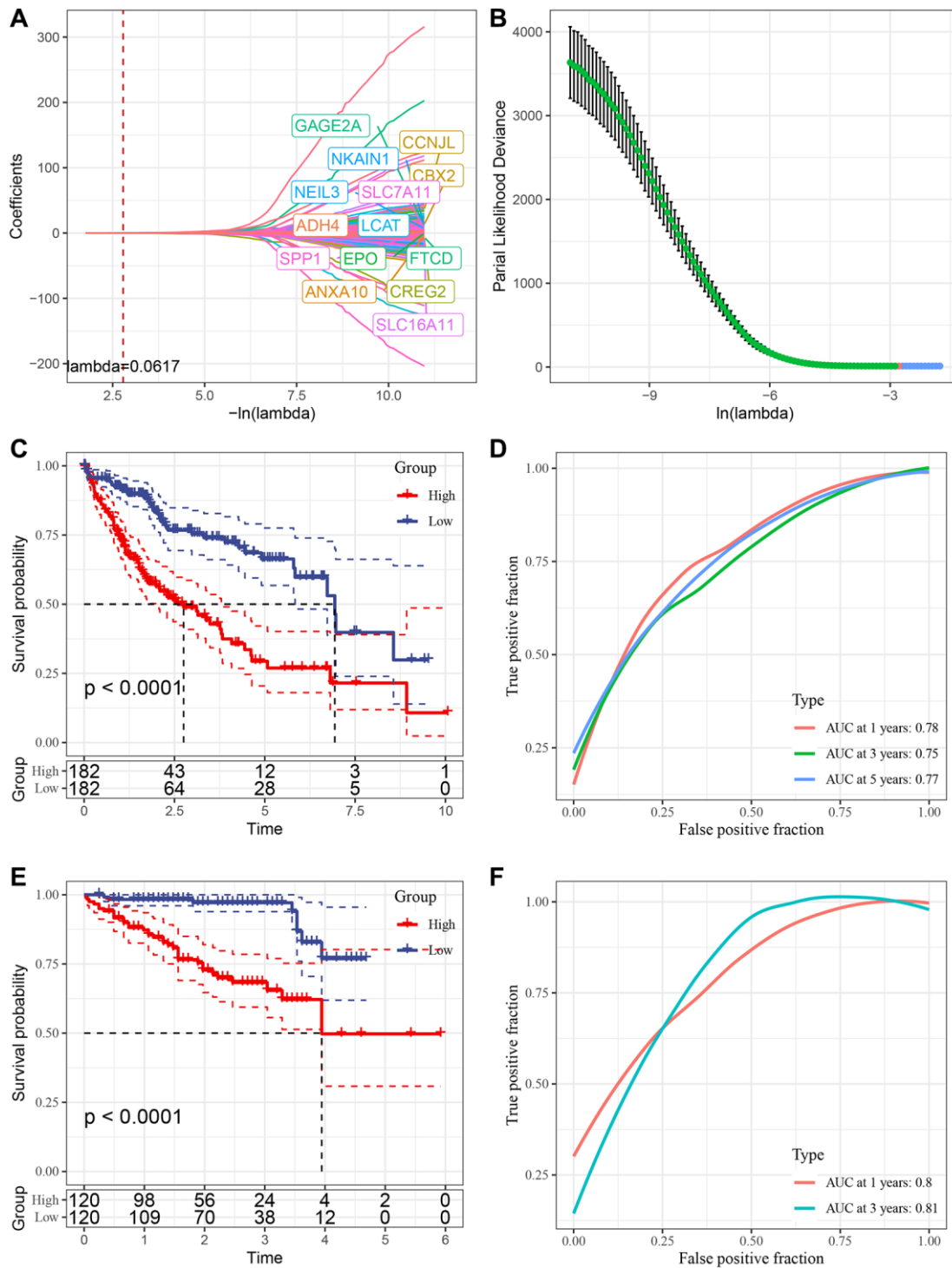

Figure 7. Construction of risk model. (**A**) Trajectory of each independent variable with lambda; (**B**) Confidence interval under lambda. (**C**) KM curve of high and low risk of risk model constructed by 16 genes in TCGA data set. (**D**) ROC curve of risk model constructed by 16 genes in TCGA data set. (**E**) KM curve of high and low risk of risk model built by 16 gene in ICGC data set. (**F**) ROC curve of risk models built by 16 gene in ICGC data set.

RiskScore in different clinicopathological features

We analyzed the relationship between RiskScore scores and tumor clinical characteristics by studying the differences in scores between various clinical phenotypes in the TCGA dataset (Figure 8A). The findings revealed that the risk score increased with clinicopathological characteristics such as clinical grade, as indicated in Figure 8A, 8B.

Relationship between RiskScore and immunity

To further investigate the relationship between RiskScore and immunity, we used ESTIMATE to predict the immunity score and, employing Spearman's method, found a significant positive correlation between the immunity score and RiskScore, as depicted in Figure 9A. Then we analyzed the scores of 28 immune cells to determine their association with high and low
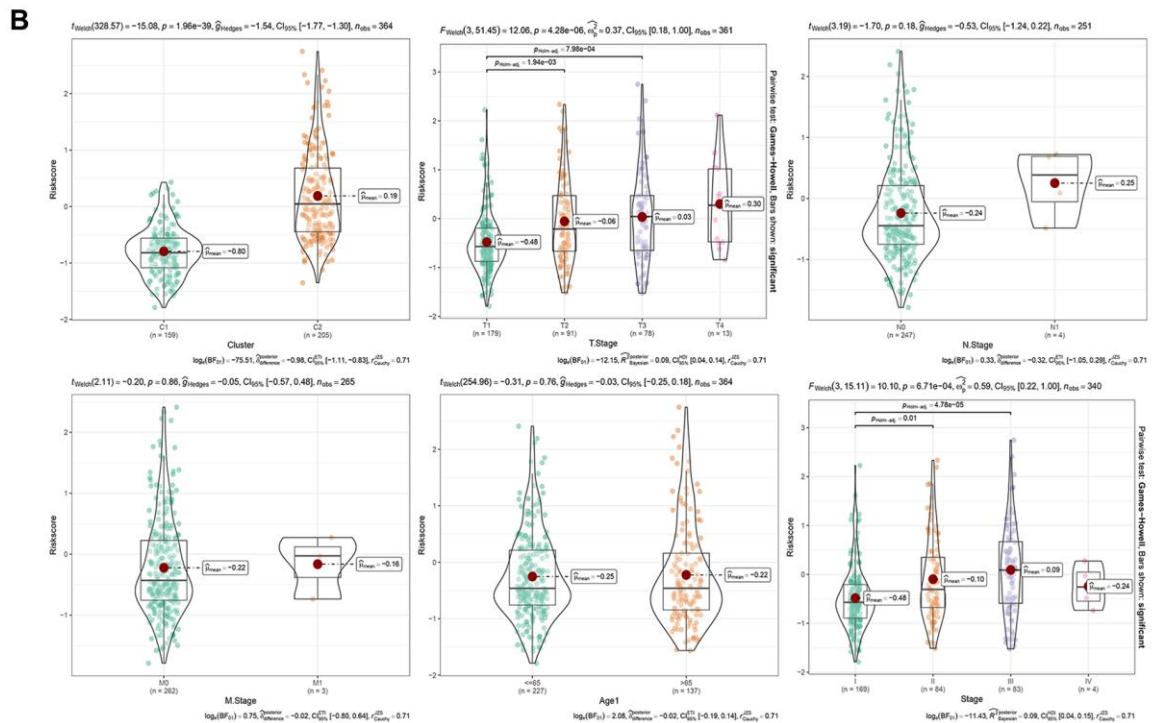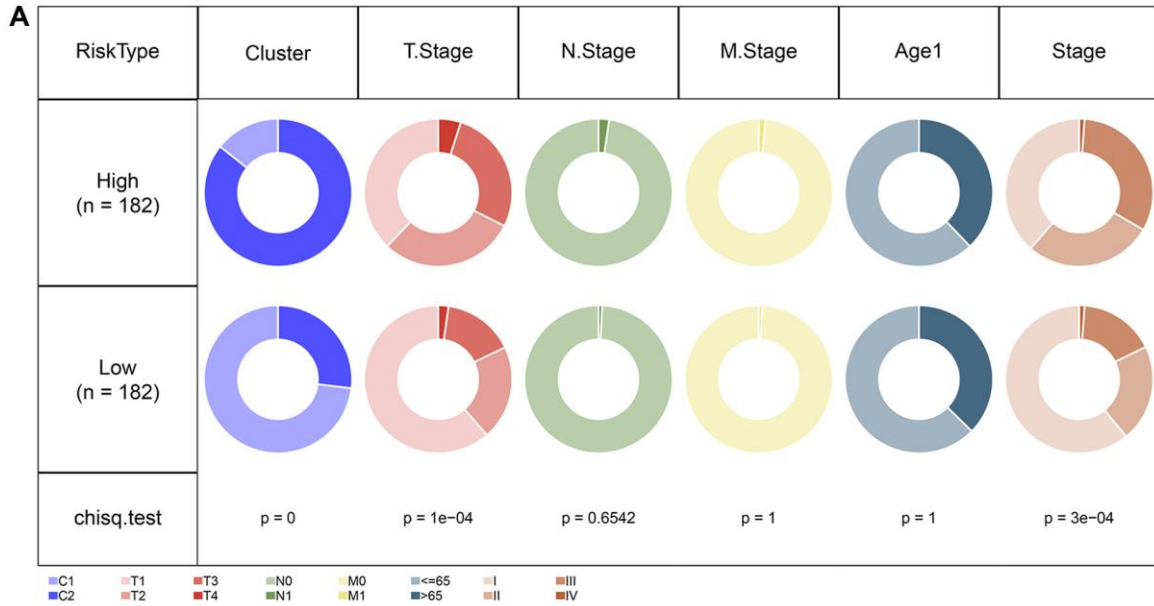

Figure 8. RiskScore in different clinicopathological features. (**A**) Comparison of clinical phenotypes between RiskScore groups in the TCGA cohort. (**B**) Differences between RiskScores of different phenotypes in the TCGA cohort (Wilcox. Test, $^*P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; and $^{****}P < 0.0001$).

risk groups, discerning significant differences in the scores of 12 immune cells, as shown in Figure 9B. The results of TIDE prediction score demonstrated that the TIDE score was higher among those with higher risk scores, as indicated in Figure 9C. We also evaluated the expression of checkpoints CTLA4, PD-1, and PD-L1 in the high-risk and low-risk groups, finding that the expression of these immune checkpoint genes was higher in the high-risk group

(Wilcox. Test, Figure 9D). We calculated the TRS score using the ssGSEA method based on TRS-associated characteristic genes, discovering that the Risk-High group had a higher TRS score (Wilcox. Test, Figure 9E). Furthermore, we computed the interferon (IFN) score using ssGSEA method and identified the gamma score (Th1/IFNγ score) to be significantly higher in the Risk-high group than in the Risk-low group, as presented in Figure 9F.
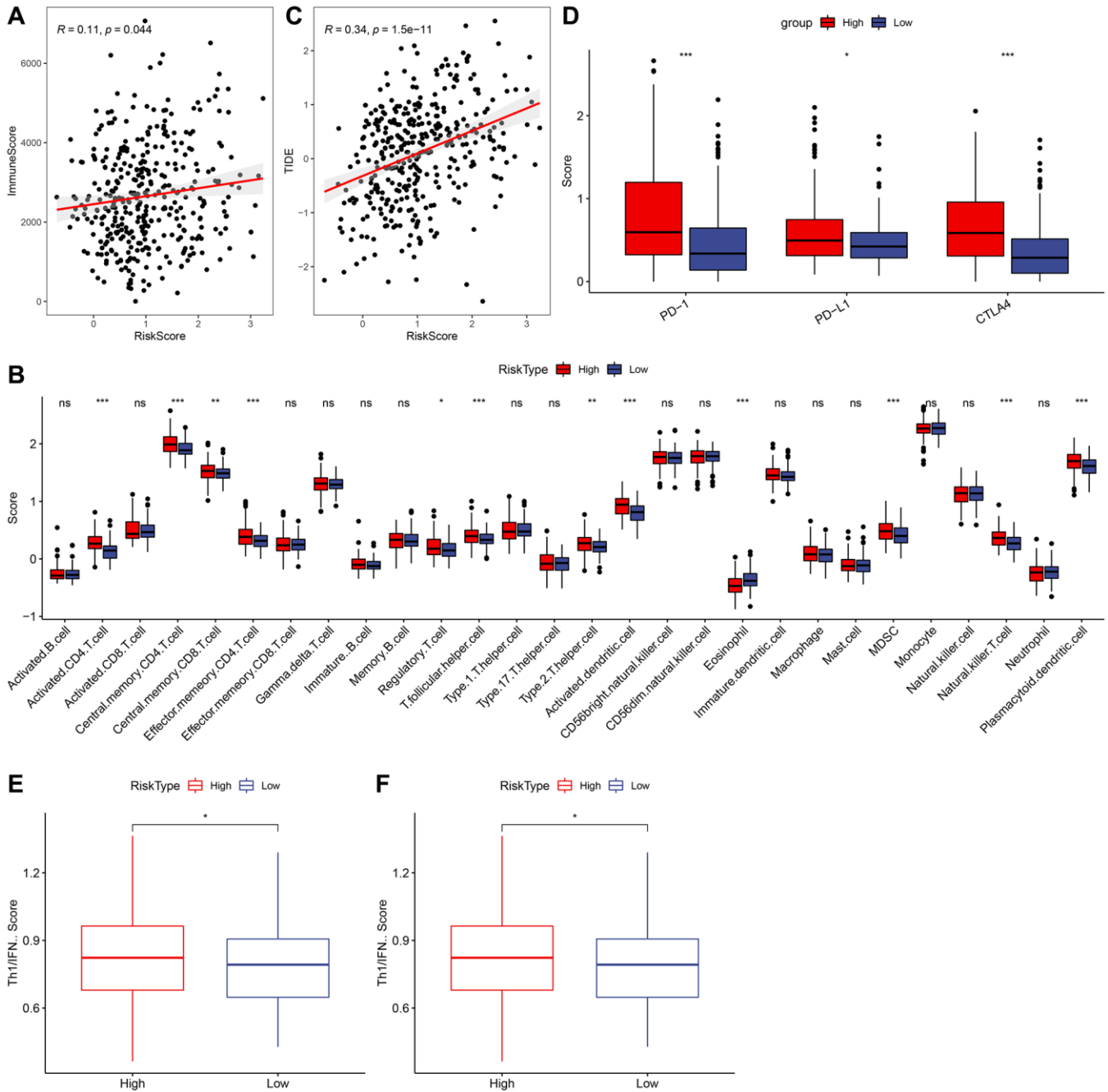

Figure 9. Relationship between RiskScore and immunity. (**A**) Correlation analysis of RiskScore and immune score. (**B**) Comparison of 28 immune cell scores in high and low risk groups. (**C**) Correlation analysis of RiskScore for TIDE. (**D**) Comparison of immune checkpoint expression in high and low risk groups. (**E**) Comparison of tumor reactivity scores in high and low risk groups. (**F**) Th1/IFNγ score comparison between high and low risk group.

RiskScore combined with clinicopathological features further improves the prognostic model and survival prediction

We demonstrated RiskScore as the most significant prognostic factor by performing univariate and multi-variate Cox regression analysis on RiskScore and clinical characteristics, as shown in Figure 10A, 10B. To estimate the risk assessment and survival probability of patients, we integrated RiskScore with other clinicopathological characteristics to design a nomogram presented in Figure 10C. The results of the model showed that RiskScore had the most substantial impact on survival

prediction. Moreover, we employed the Decision curve analysis (DCA) to evaluate the model's reliability and observed that both RiskScore and Nomogram benefits were significantly higher than the extreme curve compared to other clinicopathological features. Both the nomogram and RiskScore demonstrated significant survival predictive power, as presented in Figure 10D, 10F. Additionally, we assessed the prediction accuracy of the model using the Calibration curve, illustrated in Figure 10E, and observed that the calibration curve of the three calibration points at 1, 3, and 5 years was close to the standard curve. This indicates that the nomogram provided good prediction performance.
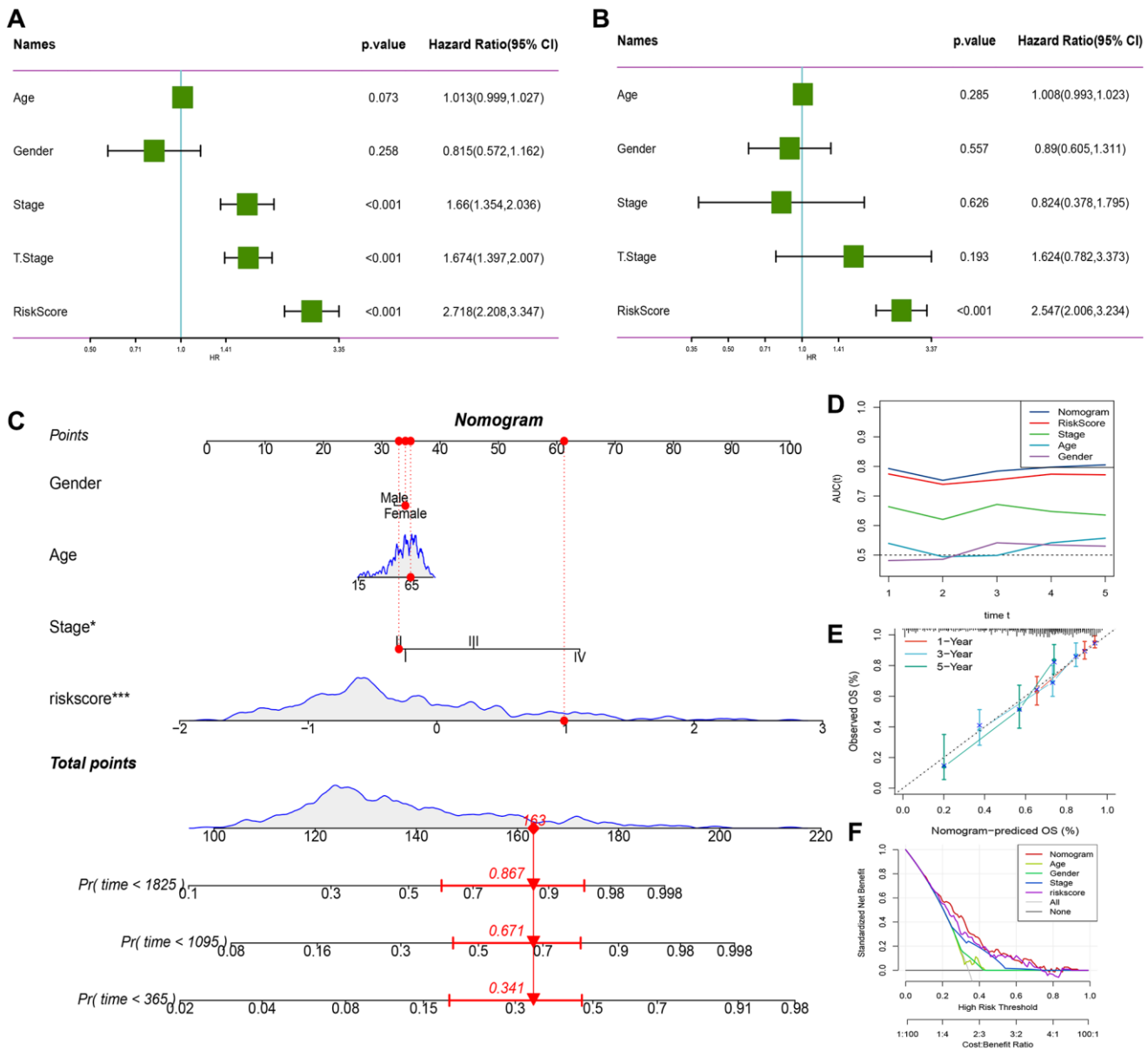

Figure 10. RiskScore combined with clinicopathological features further improves the prognostic model and survival prediction.
(**A**) Univariate Cox analysis of RiskScore and clinical characteristics. (**B**) Multivariate Cox analysis of RiskScore and clinical characteristics. (**C**) Nomogram model. (**D**) Calibration curve of nomogram at 1, 3 and 5 years. (**E**) Decision curve of nomogram. (**F**) AUC line chart, the horizontal axis is the time unit year, and the vertical axis is the model AUC value.

Validation of the risk score in HCC

To validate the risk-scoring model in HCC, we first conducted a survival analysis to identify SRXN1 with prognostic significance in the TCGA cohort (Figure 11A). Next, we examined one of the previously identified SRXN1 mRNA that constructs predictive model in several tumor cells. The result of the Quantitative reverse transcriptase-polymerase chain reaction (qRT-PCR) reveals SRXN1 mRNA levels were increased significantly in all tumor cell lines (Figure 11B). Finally, we examined the protein level of SRXN1 in cells of THLE-2 and HEP3B. Our Western blot results show that the SRXN1 protein level was increased significantly in the tumor cell line (Figure 11C).

DISCUSSION

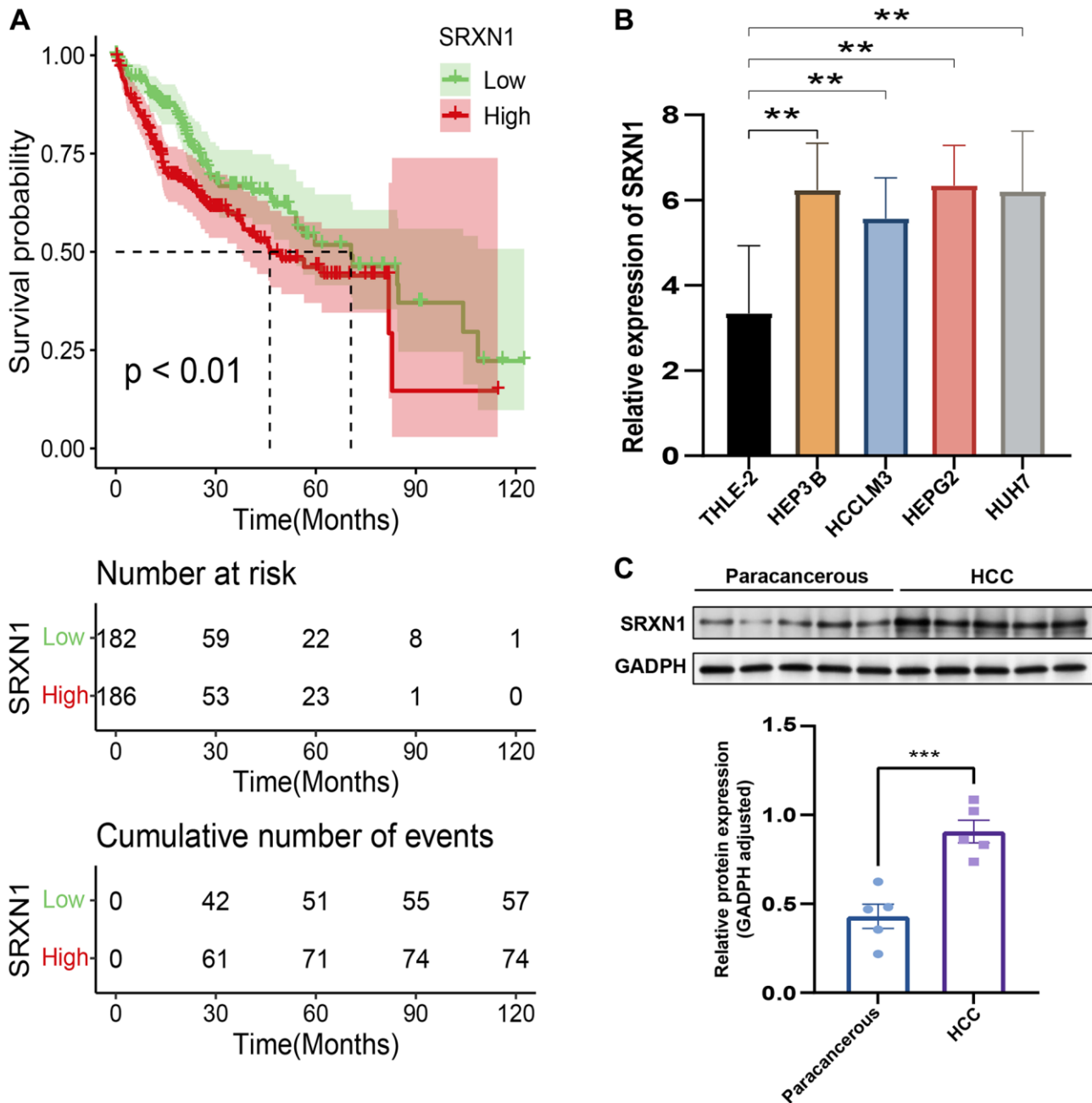In our study, we utilized single-cell genotyping to examine ligand-receptor interactions and identify the

Figure 11. Validation of the risk score in HCC. (**A**) Survival analysis of TCGA-LIHC based on SRXN1 levels. (**B**) qRT-PCR images of tumor cell line (HEP3B, HCCLM3, HEPG2 and HUH7). SRXN1 levels were analyzed by qRT-PCR; $^{**}P < 0.01$ versus liver cell line. (**C**) Representative images of Western blots of the liver cell line (THLE-2) and tumor cell line (HEP3B). SRXN1 levels were determined by Western blot (mean ± SEM; $n = 5$; $^{**}P < 0.001$ versus THLE-2).

key receptor-ligand genes in HCC. Based on these genes, we clustered the patients in two clusters with disguising TME characteristics. We then constructed a prognostic model which can creativity predict the prognosis of HCC with 16 genes (receptor-ligand genes). This approach allowed us to gain insight into the patterns of receptor-ligand genotyping within the tumor microenvironment both in intra-tumor and extra-tumor, which could have significant implications for HCC who can be benefits from the immunotherapy.

Ligand-receptor interactions, playing a crucial role for the signal transduction between cancer cells and non-malignant cells in the tumor microenvironment (TME), is the key to the progression of tumors and exhibit a high degree of heterogeneity, posing challenges in understanding tumor progression and treatment. This heterogeneity is manifested in the interactions within the tumor microenvironment, including the ligand-receptor signaling between tumor cells and non-malignant cells. To resolve it, the clustering method was used in the identifying the special subpopulation for HCC by ligand-receptor genes in our study. A previous study also conducted the comprehensive understanding of these interactions, particularly across 20 solid tumor types, can aid in better understanding the paracrine interactions of cancer cells and the tumor micro-environment [11]. According to previous studies, the intratumoral factors such as hypoxia [12] and metabolic abnormalities [13, 14] may lead to changes in the surface expression of ligands. For instance, in HCC, glycosylation regulator subtypes have been found to be associated with intratumoral factors [15, 16]. In addition, the evolution of hepatocellular carcinoma (HCC) is believed to be driven by both endogenous and exogenous factors, which may include genetic instability (endogenous factors) as well as factors from the tumor microenvironment, such as immune cells and cancer-associated fibroblasts (exogenous factors). Also, some receptors-ligand genes, such as PD-L1/PD1 both expressed in stromal cells or macrophage cells in HCC, could induce the immunosuppressive effect to TME [17, 18]. The aberrant activation of the Wnt signaling pathway, which is initiated by the binding of Wnt ligands to Frizzled receptors, is frequently implicated in stem cell cancer initiation and progression [19].

In our model, we next screened 16 ligand receptor genes in single cell tumor cells through prognosis, and extensive research has officially confirmed the role of these genes in hepatocellular carcinoma. For example, the increased expression of SRXN1 (Sulfiredoxin 1) may be helpful in resisting oxidative stress produced by tumor cells [20]. Overexpression of PBK (PDZ Binding Kinase) may lead to increased proliferation and survival ability of tumor cells [21]. EPO (Erythropoietin) can enhance the survival ability of tumor cells by inhibiting apoptosis and promoting angiogenesis. This may increase the risk of hepatocellular carcinoma [22]. SLC7A11 is involved in the uptake of sulfated amino acids by cells with increasing expression led to increased resistance to chemotherapy drugs in cancer [23].

However, our study is not without limitations. Although the clustering method was useful in identifying specific subpopulations in HCC, further research is needed to validate the clinical relevance of these clusters. In addition, the role of these 16 genes in HCC requires experimental validation to further establish their functional significance in the progression of HCC. Moreover, our prognostic model needs to be validated in larger, prospective studies and across different populations.

In conclusion, our study has shed light on the intricate landscape of ligand-receptor interactions within the TME of HCC. By identifying key receptor-ligand genes and developing a prognostic model, we provide new insights into HCC progression and potential therapeutic targets. Despite the limitations, our findings offer a promising step towards more personalized therapeutic strategies for HCC patients.

MATERIALS AND METHODS

Single cell data downloading and preprocessing

We utilized the Seurat (Version 4.1.1) R package to process single cell data obtained from the GEO database (https://www.ncbi.nlm.nih.gov/geo/) for a reliable liver cancer single cell expression profile dataset (GSE149614) [24]. The sequencing platform used was GPL24676 Illumina NovaSeq 6000 for Homo sapiens. We filtered the single-cell data by setting the criteria that each gene should be expressed in at least three cells and each cell should express at least 250 genes. We further ensured that the gene expressed by each cell was more than 500 and less than 4000, the unique molecular identifier (UMI) of each cell was less than 15000, mitochondrial gene content was less than 10%, ribosome expression was more than 3%, and red blood cell expression was less than 1%, using the Percentage Feature Set function. We obtained a total of 18 single-cell sequencing samples from the GSE149614 dataset, including 10 hepatocellular carcinoma tissues and 8 normal liver tissues.

TCGA database data and ICGC database data download

We retrieved clinical phenotype data of liver cancer from The Cancer Genome Atlas (TCGA) database and removed samples without survival time and survival

status. We ensured that all patients had a survival time of more than 0 days. Additionally, we downloaded TCGA expression profile data and obtained 368 tumor samples. Moreover, we downloaded the liver cancer data, expression profile data, and survival data of the ICGC-JP cohort from the International Cancer Genome Consortium (ICGC) database. We removed samples lacking survival time and survival status and ensured that all patients had a survival time of more than 0 days. Eventually, we obtained a total of 232 samples. We calculated snv mutation information of the gene of TCGA-LIHC using Mutect2 software through TCGAbiolinks (version 2.24.3) download [25]. Homologous Recombination Defects, Fraction Altered, Number of Segments, Nonsilent Mutation Rate, Aneuploidy Score, Silent Mutation Rate Data were obtained from a previous study [26].

Single-cell dimensionality reduction and annotation

To normalize the data of the 18 samples, we applied log-normalization separately. Then, we found hyper-variable genes using the FindVariable Features function, which identified variable features based on variance stabilization transformation (VST). We scaled all genes using the ScaleData function followed by PCA dimen-sionality reduction using RunPCA to find anchor points, choosing dim = 30. The cells were clustered through the FindNeighbors and FindClusters functions with a resolution of 0.45 to avoid batch effect of the samples, a Harmony package (version 0.1.0) was used to batch correct the samples. For cell type annotation, we used the HumanPrimary CellAtlasData dataset in the SingleR (version 1.10.0) R package. We also used the tools Copycat [27] (version 1.0.8). Subsequently, marker genes of each sub-population were screened using the FindAllMarkers function with logfc = 0.5 (fold difference) and Minpct = 0.5 (minimum expression ratio of differential genes), with a screening threshold of a corrected *p*-value less than 0.05.

Cell communication analysis

We used the R package CellChat (version 1.5.0) for the analysis of cell-to-cell communication [28]. The results of CellChat were used to screen out the receptor ligand genes associated with malignant cells.

Consistent clustering for receptor-ligand genes

We performed univariate Cox regression analysis to identify genes associated with prognosis ($p < 0.05$), from the genes identified in CellChat. Based on the expression profile of these prognostically relevant genes, we classified the patients using the consensus clustering of tumor tissues in the TCGA dataset through

ConsensusClusterPlus (version 1.60.0) package in R programming language [29]. We used the HC algorithm with "Pearson" as the metric distance and performed 500 bootstraps with each bootstrap procedure including 80% of the training set of patients. The number of clusters was set from 2 to 10, and the optimal number of clusters was determined by calculating the consistency matrix and the consistency cumulative distribution function. The CDF was used to determine the best classification.

Path analysis and variance analysis

We obtained Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway genes from the msigdb (http://www.gsea-msigdb.org/) database. To calculate pathway scores for samples of different molecular subtypes, we utilized the GSVA package [30]. For differential analysis of the pathway scores between the two subtypes, we used the limma package (version 3.52.2) [31]. Here, we selected pathway visualization with |logFC| >0.15 and *P*-value < 0.05 criteria. Furthermore, we conducted differential analysis of gene expression by subtype using the limma package and identified differential genes by screening for |logFC| >log2 (1.2) and *P*-value < 0.05. We then employed gene set enrichment analysis (GSEA) using the Hallmark gene sets in the msigdb database as the background set, and we merged these results with the previous differential analysis to reveal further insights into the pathways and genes involved.

LASSO regression analysis and risk model construction

Genes significantly associated with survival outcomes were identified through univariate Cox regression models to construct a prognostic model. In addition, reliable predictors were selected using LASSO analysis [32], a compression estimator that constructs a penalty function to compress coefficients and set others to zero. This methodology retains the advantages of subset shrinkage and acts as a biased estimator that reduces multicollinearity problems encountered in regression analysis. We conducted LASSO Cox regression using the R software package glmnet (version 4.1-4). Risk scores for each patient in the TCGA and ICGC databases were calculated using the formula: Risk score = Σ coefficient mRNA × expression level mRNA. Finally, we analyzed the correlation between patient risk scores and prognosis.

Immune infiltration and prediction of immunotherapy

To assess the distribution of specific cellular components in the immune microenvironment, we

calculated the scores of 28 immune cells using ssGSEA based on characteristic genes obtained in a previous study [33]. Differences among subtypes were assessed using Kruskal test. In addition, we assessed immune cell infiltration using ESTIMATE [34] (version 1.0.13). We used TIDE to evaluate the potential clinical effects of immunotherapy across groups [35]. A previous study [36] showed that the TRS score could predict a patient's response to immunotherapy. CD8+ T cells in the tumor microenvironment can produce interferon-γ (IFNγ), which can up-regulate PD-1/PD-L1 and IDO1 gene expression [37, 38]. Up-regulated IDO1 expression is positively associated with tumor progression and poor prognosis [39]. We extracted Th1/IFNγ gene signatures from a previous study [32] and calculated IFNγ scores using ssGSEA for each patient.

Cell culture

Here, we purchased normal liver cell line (THLE-2, Shanghai Academy of Life Science), and tumor cell lines, including HEP3B (catalog No. ZQ0024), HEPG2 (catalog No. ZQ0022), HCCLM3 (catalog No. ZQ0023) and HUH7 (catalog No. ZQ0025). Culture dishes with Dulbecco's Modified Eagle's Medium (DMEM) containing 10% fetal bovine serum (FBS) were used to inoculate all cell lines to maintain growth at 37°C with5% CO_2. The cell culture medium was also supplemented with 1% penicillin/streptomycin. The culture dishes were purchased from Guangzhou Jet Biofiltration (Guangzhou, China). FBS, and penicillin/ streptomycin were purchased from BI (BI, Ridgefield, CT, USA).

RNA extraction and quantitative reverse transcription PCR

Using the total RNA isolation kit (Foregene, Chengdu, China) and the PrimeScript RT kit (Takara Biomedical Technology, Beijing, China), RNA was extracted from the cells and reverse transcribed into cDNA. TB Green Premix Ex Taq II (Takara Biomedical Technology, Beijing, China) and was tested by real-time PCR with $2^{-\Delta\Delta Ct}$. GAPDH was the reference gene in the process. The qPCR primers are listed as follows: SRXN1 Forward Primer CAGGGAGGTGACTACTTCTACTC; SRXN1 Reverse Primer CAGGTACACCCTTAGGTCTGA.

Western blot

RIPA lysis buffer was used to extract the protein, and the BCA Protein Assay Kit was used to calculate the protein concentration. 10% SDS-PAGE was used to separate equal amounts of protein, and the resultant membrane, polyvinylidene difluoride, was used. The membrane was incubated with the primary antibody overnight at 4°C after being blocked with 5% BSA in PBST for 1 hour at room temperature. It was then subjected to three PBST washes before being exposed to an appropriate horseradish peroxidase (HRP)-conjugated secondary antibody for one hour at room temperature. With the use of an ECL chemiluminescent reagent, the reaction was observed. Using Image Lab, the blot's intensity was measured.

Abbreviations

HCC: Hepatocellular carcinoma; RFA: Radiofrequency ablation; MWA: Microwave ablation; KYN: Kynurenic acid; AHR: Aryl hydrocarbon receptor; TDO: Tryptophan-2,3-dioxygenase; TGFβ: Transforming growth factor-beta; IFNγ: Interferon-γ; TCGA: The Cancer Genome Atlas; ICGC: International Cancer Genome Consortium; GSEA: Gene set enrichment analysis; TME: Tumor microenvironment.

AUTHOR CONTRIBUTIONS

FQ W designed the study. JZ W, CC W, and XH C performed data analysis. CC W collected clinical samples and JZ W validated Western blot on clinical samples, and in addition he validated qPCR on hepato-cellular carcinoma cell lines. JZ W, PP L and JJ L drafted the manuscript. FQ W revised the manuscript. All authors read and approved the final manuscript.

CONFLICTS OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

FUNDING

REFERENCES

1. Forner A, Reig M, Bruix J. Hepatocellular carcinoma. Lancet. 2018; 391:1301–14.
https://doi.org/10.1016/S0140-6736(18)30010-2
PMID:29307467

2. Yang JD, Hainaut P, Gores GJ, Amadou A, Plymoth A, Roberts LR. A global view of hepatocellular carcinoma: trends, risk, prevention and management. Nat Rev Gastroenterol Hepatol. 2019; 16:589–604.
https://doi.org/10.1038/s41575-019-0186-y
PMID:31439937

3. Llovet JM, Montal R, Sia D, Finn RS. Molecular therapies and precision medicine for hepatocellular carcinoma. Nat Rev Clin Oncol. 2018; 15:599–616.
https://doi.org/10.1038/s41571-018-0073-4
PMID:30061739

4. Omata M, Cheng AL, Kokudo N, Kudo M, Lee JM, Jia J, Tateishi R, Han KH, Chawla YK, Shiina S, Jafri W, Payawal DA, Ohki T, et al. Asia-Pacific clinical practice guidelines on the management of hepatocellular carcinoma: a 2017 update. Hepatol Int. 2017; 11:317–70.
https://doi.org/10.1007/s12072-017-9799-9
PMID:28620797

5. Llovet JM, Castet F, Heikenwalder M, Maini MK, Mazzaferro V, Pinato DJ, Pikarsky E, Zhu AX, Finn RS. Immunotherapies for hepatocellular carcinoma. Nat Rev Clin Oncol. 2022; 19:151–72.
https://doi.org/10.1038/s41571-021-00573-2
PMID:34764464

6. Sangro B, Sarobe P, Hervás-Stubbs S, Melero I. Advances in immunotherapy for hepatocellular carcinoma. Nat Rev Gastroenterol Hepatol. 2021; 18:525–43.
https://doi.org/10.1038/s41575-021-00438-0
PMID:33850328

7. Trenker R, Jura N. Receptor tyrosine kinase activation: From the ligand perspective. Curr Opin Cell Biol. 2020; 63:174–85.
https://doi.org/10.1016/j.ceb.2020.01.016
PMID:32114309

8. Opitz CA, Litzenburger UM, Sahm F, Ott M, Tritschler I, Trump S, Schumacher T, Jestaedt L, Schrenk D, Weller M, Jugold M, Guillemin GJ, Miller CL, et al. An endogenous tumour-promoting ligand of the human aryl hydrocarbon receptor. Nature. 2011; 478:197–203.
https://doi.org/10.1038/nature10491
PMID:21976023

9. Derynck R, Turley SJ, Akhurst RJ. TGFβ biology in cancer progression and immunotherapy. Nat Rev Clin Oncol. 2021; 18:9–34.
https://doi.org/10.1038/s41571-020-0403-1
PMID:32710082

10. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. Cell. 2015; 160:48–61.
https://doi.org/10.1016/j.cell.2014.12.033
PMID:25594174

11. Ghoshdastider U, Rohatgi N, Mojtabavi Naeini M, Baruah P, Revkov E, Guo YA, Rizzetto S, Wong AML, Solai S, Nguyen TT, Yeong JPS, Iqbal J, Tan PH, et al. Pan-Cancer Analysis of Ligand-Receptor Cross-talk in the Tumor Microenvironment. Cancer Res. 2021; 81:1802–12.
https://doi.org/10.1158/0008-5472.CAN-20-2352
PMID:33547160

12. Chen S, Gao Y, Wang Y, Daemen T. The combined signatures of hypoxia and cellular landscape provides a prognostic and therapeutic biomarker in hepatitis B virus-related hepatocellular carcinoma. Int J Cancer. 2022; 151:809–24.
https://doi.org/10.1002/ijc.34045
PMID:35467769

13. Fu J, Chen Q, Lai Z, Lin K, Fang G, Ding Z, Gao Y, Zeng Y. A novel preoperative inflammation score system established for postoperative prognosis predicting of intrahepatic cholangiocarcinoma. BMC Cancer. 2023; 23:188.
https://doi.org/10.1186/s12885-023-10668-x
PMID:36829168

14. Chen Q, Li F, Gao Y, Xu G, Liang L, Xu J. Identification of Energy Metabolism Genes for the Prediction of Survival in Hepatocellular Carcinoma. Front Oncol. 2020; 10:1210.
https://doi.org/10.3389/fonc.2020.01210
PMID:32903581

15. Zhao L, Guo Y, Liu Z, Ma J, Peng Y, Zhang D. Characterization of glycosylation regulator-mediated glycosylation modification patterns and tumor microenvironment infiltration in hepatocellular carcinoma. Front Genet. 2022; 13:1001901.
https://doi.org/10.3389/fgene.2022.1001901
PMID:36437920

16. Deng T, Ye Q, Jin C, Wu M, Chen K, Yang J, Chen Z, Yu X, Chen G, Wang Y. Identification and validation of a glycolysis-associated multiomics prognostic model for hepatocellular carcinoma. Aging (Albany NY). 2021; 13:7481–98.
https://doi.org/10.18632/aging.202613
PMID:33686959

17. Zhou G, Sprengers D, Boor PPC, Doukas M, Schutz H, Mancham S, Pedroza-Gonzalez A, Polak WG, de Jonge J, Gaspersz M, Dong H, Thielemans K, Pan Q, et al. Antibodies Against Immune Checkpoint Molecules Restore Functions of Tumor-Infiltrating T Cells in Hepatocellular Carcinomas. Gastroenterology. 2017; 153:1107–19.e10.
https://doi.org/10.1053/j.gastro.2017.06.017
PMID:28648905

18. Juneja VR, McGuire KA, Manguso RT, LaFleur MW, Collins N, Haining WN, Freeman GJ, Sharpe AH. PD-L1 on tumor cells is sufficient for immune evasion in immunogenic tumors and inhibits CD8 T cell cytotoxicity. J Exp Med. 2017; 214:895–904.

https://doi.org/10.1084/jem.20160801
PMID:28302645

19. Mao XD, Wei X, Xu T, Li TP, Liu KS. Research progress in breast cancer stem cells: characterization and future perspectives. Am J Cancer Res. 2022; 12:3208–22.
PMID:35968346

20. Lv X, Yu H, Zhang Q, Huang Q, Hong X, Yu T, Lan H, Mei C, Zhang W, Luo H, Pang P, Shan H. SRXN1 stimulates hepatocellular carcinoma tumorigenesis and metastasis through modulating ROS/p65/BTG2 signalling. J Cell Mol Med. 2020; 24:10714–29.
https://doi.org/10.1111/jcmm.15693
PMID:32746503

21. Liu Z, Pu Y, Bao Y, He S. Investigation of Potential Molecular Biomarkers for Diagnosis and Prognosis of AFP-Negative HCC. Int J Gen Med. 2021; 14:4369–80.
https://doi.org/10.2147/IJGM.S323868
PMID:34408477

22. Wang T, Dai L, Shen S, Yang Y, Yang M, Yang X, Qiu Y, Wang W. 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. 2022; 13:843408.
https://doi.org/10.3389/fimmu.2022.843408
PMID:35693827

23. Chen Q, Zheng W, Guan J, Liu H, Dan Y, Zhu L, Song Y, Zhou Y, Zhao X, Zhang Y, Bai Y, Pan Y, Zhang J, Shao C. SOCS2-enhanced ubiquitination of SLC7A11 promotes ferroptosis and radiosensitization in hepatocellular carcinoma. Cell Death Differ. 2023; 30:137–51.
https://doi.org/10.1038/s41418-022-01051-7
PMID:35995846

24. Hao Y, Hao S, Andersen-Nissen E, Mauck WM 3rd, Zheng S, Butler A, Lee MJ, Wilk AJ, Darby C, Zager M, Hoffman P, Stoeckius M, Papalexi E, et al. Integrated analysis of multimodal single-cell data. Cell. 2021; 184:3573–87.e29.
https://doi.org/10.1016/j.cell.2021.04.048
PMID:34062119

25. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, Sabedot TS, Malta TM, Pagnotta SM, Castiglioni I, Ceccarelli M, Bontempi G, Noushmehr H. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. Nucleic Acids Res. 2016; 44:e71.
https://doi.org/10.1093/nar/gkv1507
PMID:26704973

26. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, Porta-Pardo E, Gao GF, Plaisier CL, Eddy

JA, Ziv E, Culhane AC, Paull EO, et al, and Cancer Genome Atlas Research Network. The Immune Landscape of Cancer. Immunity. 2018; 48:812–30.e14.
https://doi.org/10.1016/j.immuni.2018.03.023
PMID:29628290

27. Gao R, Bai S, Henderson YC, Lin Y, Schalck A, Yan Y, Kumar T, Hu M, Sei E, Davis A, Wang F, Shaitelman SF, Wang JR, et al. Delineating copy number and clonal substructure in human tumors from single-cell transcriptomes. Nat Biotechnol. 2021; 39:599–608.
https://doi.org/10.1038/s41587-020-00795-2
PMID:33462507

28. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, Myung P, Plikus MV, Nie Q. Inference and analysis of cell-cell communication using CellChat. Nat Commun. 2021; 12:1088.
https://doi.org/10.1038/s41467-021-21246-9
PMID:33597522

29. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics. 2010; 26:1572–3.
https://doi.org/10.1093/bioinformatics/btq170
PMID:20427518

30. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics. 2013; 14:7.
https://doi.org/10.1186/1471-2105-14-7
PMID:23323831

31. 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

32. Danilova L, Ho WJ, Zhu Q, Vithayathil T, De Jesus-Acosta A, Azad NS, Laheru DA, Fertig EJ, Anders R, Jaffee EM, Yarchoan M. Programmed Cell Death Ligand-1 (PD-L1) and CD8 Expression Profiling Identify an Immunologic Subtype of Pancreatic Ductal Adenocarcinomas with Favorable Survival. Cancer Immunol Res. 2019; 7:886–95.
https://doi.org/10.1158/2326-6066.CIR-18-0822
PMID:31043417

33. 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

34. 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

35. Yan M, Hu J, Ping Y, Xu L, Liao G, Jiang Z, Pang B, Sun S, Zhang Y, Xiao Y, Li X. Single-Cell Transcriptomic Analysis Reveals a Tumor-Reactive T Cell Signature Associated With Clinical Outcome and Immunotherapy Response In Melanoma. Front Immunol. 2021; 12:758288.
https://doi.org/10.3389/fimmu.2021.758288
PMID:34804045

36. Takikawa O, Tagawa Y, Iwakura Y, Yoshida R, Truscott RJ. Interferon-gamma-dependent/independent expression of indoleamine 2,3-dioxygenase. Studies with interferon-gamma-knockout mice. Adv Exp Med Biol. 1999; 467:553–7.
https://doi.org/10.1007/978-1-4615-4709-9_68
PMID:10721099

37. Garcia-Diaz A, Shin DS, Moreno BH, Saco J, Escuin-Ordinas H, Rodriguez GA, Zaretsky JM, Sun L, Hugo W, Wang X, Parisi G, Saus CP, Torrejon DY, et al. Interferon Receptor Signaling Pathways Regulating PD-L1 and PD-L2 Expression. Cell Rep. 2017; 19:1189–201.
https://doi.org/10.1016/j.celrep.2017.04.031
PMID:28494868

38. Zhang R, Liu H, Li F, Li H, Yu J, Ren X. The correlation between the subsets of tumor infiltrating memory T cells and the expression of indoleamine 2,3-dioxygenase in gastric cancer. Dig Dis Sci. 2013; 58:3494–502.
https://doi.org/10.1007/s10620-013-2837-0
PMID:23979437

39. Chen JY, Li CF, Kuo CC, Tsai KK, Hou MF, Hung WC. Cancer/stroma interplay via cyclooxygenase-2 and indoleamine 2,3-dioxygenase promotes breast cancer progression. Breast Cancer Res. 2014; 16:410.
https://doi.org/10.1186/s13058-014-0410-1
PMID:25060643
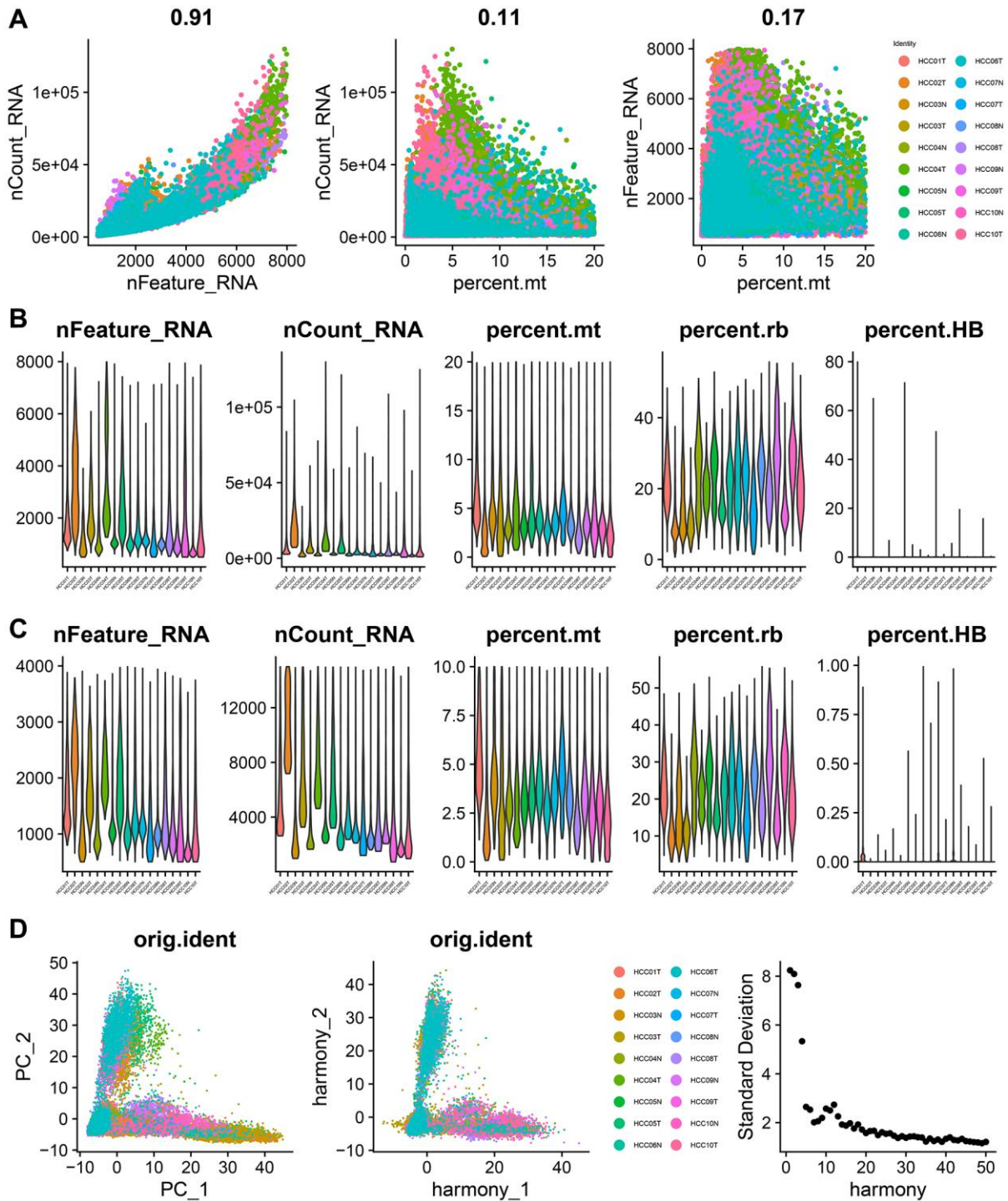
Supplementary Figure 1. (**A**) Relationship between mitochondrial gene and UMI/mRNA quantity, relationship between UMI and mRNA quantity. (**B**) Relationship between mRNA/UMI/mitochondrial content/rRNA content of each sample before filtration. (**C**) Relationship between mRNA/UMI/mitochondrial content/rRNA content of each sample after filtration. (**D**) Sample distribution map of PCA dimension reduction and anchor map of PCA.

Supplementary Table

Supplementary Table 1. Raw data for western blots and qRT-PCR.

	WB		
	GADPH	**SRXN1**	**GAPDH/SRXN1**
1	99187.844	35326.844	0.356161023
2	84853.137	18479.309	0.217779915
3	88187.844	41430.673	0.469800271
4	94555.966	59051.551	0.62451428
5	95978.966	46234.602	0.481715983
6	96021.43	104183.673	1.085004389
7	107641.258	79371.137	0.737367237
8	101272.258	87097.844	0.860036556
9	86776.894	72204.016	0.832064996
10	85029.53	86875.187	1.021706071

qPCR				
THLE-2	**HEP3B**	**HCCLM3**	**HEPG2**	**HUH7**
3.082	5.122	5.103917211	7.839581737	5.285772093
4.053	7.378	7.032982417	6.364222221	4.958378712
2.902	7.542	5.468909092	6.374691752	5.459431619
6.146	5.861	5.56924803	5.185866545	8.572624443
2.054	6.522	6.063934306	6.763145958	7.275472718
1.806	4.987	4.164303583	5.502712486	5.656782364