

A Novel Immune-Related Gene-Based Model: Prognosis and Immunotherapy Response in Hepatocellular Carcinoma

Junkai Ren

Department of Hepatic Surgery, Liver Transplant Center, the Third Affiliated Hospital of Sun Yat-sen University, Guangzhou

Chao Yuan

Department of Hepatic Surgery, Liver Transplant Center, the Third Affiliated Hospital of Sun Yat-sen University, Guangzhou

Fanxin Zeng

Department of Hepatic Surgery, Liver Transplant Center, the Third Affiliated Hospital of Sun Yat-sen University, Guangzhou

Kaining Zeng

Department of Hepatic Surgery, Liver Transplant Center, the Third Affiliated Hospital of Sun Yat-sen University, Guangzhou

Tong Zhang (✉ zhjeff72@sina.com)

Department of Hepatic Surgery, Liver Transplant Center, the Third Affiliated Hospital of Sun Yat-sen University, Guangzhou

Research Article

Keywords:

Posted Date: May 12th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1630689/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Liver hepatocellular carcinoma (LIHC) remains a significant global health challenge. Immunotherapy for LIHC is a very promising treatment modality. However, some patients with LIHC have a poor response to tumor immunotherapy. Thus, it is imperative to fully understand liver cancer immune subtypes, by which individualized and optimized immunotherapy would be developed.

Methods: We downloaded transcriptome and clinical data of TCGA-LIHC from the Cancer Genome Atlas (TCGA) databases, GSE52018 and GSE72467 from the Gene Expression Omnibus (GEO). The immune cell infiltration of the tumor microenvironment was calculated. Based on the results of CIBERSORT, the LIHC patients were clustered into three subtypes. To establish the prognostic model, we employed LASSO-COX regression analysis. The differences in immune cell infiltration, tumor stemness, somatic mutations, and sensitivity to the immune checkpoint inhibitors in patients with different risk groups were systematically assessed.

Results: According to the results of CIBERSORT, LIHC samples were divided into three individual subtypes: Cluster-A, B, and C. These three subtypes have distinct characteristics with regard to immune cell infiltration, prognosis, and response to immune checkpoint inhibitors (ICIs) and chemotherapy drugs. Cluster-B belongs to hot tumors, and Cluster-A and Cluster-C belong to cold tumors. We trained a prognostic model based on differential genes between cold and hot tumors. We then attempted to explore the differences among different risk groups in the immune microenvironment, oncogenic biology, genetic variation, and prognosis. The characteristics of immune activation, enhanced immunotherapy response, and prolonged survival was observed in the low-risk group. Finally, we combined Riskscore and other clinicopathological features to build a nomogram to predict the 1-, 3-, and 5-year survival rates of LIHC patients.

Conclusion: In this study, an immune-related gene-based model of predicting the prognosis and immunotherapy response in hepatocellular carcinoma was established.

Introduction

Liver cancer is expected to exceed 1 million cases by 2025, which is one of the most malignant digestive tract cancers [1]. LIHC is the only cancer of the five deadliest cancers with an annual increase in incidence, which was characterized by high mortality, recurrence, and metastasis, about 90% of primary liver cancer [2]. The occurrence of LIHC is a multi-step and complex biological process. Many signaling cascades are changed in LIHC, resulting in a heterogeneous molecular profile that ultimately leads to tumorigenesis, progression, and metastasis. Although there are various treatment methods for LIHC, tumor recurrence is very common after the treatment, and the 5-year recurrence rate exceeds 70% [3]. For the recent few years, ICIs-based immunotherapy has been a very promising treatment in many advanced cancers, especially virus-induced cancers [4-6]. Some immune checkpoint molecules such as PD-1 and CTLA-4, are usually targets of immunotherapy. CTLA-4 is low expressed on activated T cells and

regulatory T cells (Tregs), but absent on naïve T cells [7]. CTLA-4 can negatively regulate T cell activation. PD-1 can regulate peripheral immune tolerance, and autoimmunity and inhibit T cells activation and proliferation. PD-1 is mainly expressed on CD8+ T cells, Treg, and myeloid-derived suppressor cells. [8, 9]. Current CTLA-4 and PD-1 studies in LIHC demonstrate manageable safety, optimistic response rates, and prolonged overall survival [10-12]. But in many patients, ICIs are ineffective or only partially effective, and response rates to current immunotherapy are low, at approximately 10-25% [13]. Another problem is acquired resistance in patients who initially respond, leading to tumor recurrence. The Tumor microenvironment (TME) takes an essential part in tumorigenesis, which consists of stromal cells and immune cells, such as CD8+, CD4+ T cells, Tregs, and dendritic cells. These tumor-associated immune cells may have tumor-antagonizing or tumor-promoting functions. Current studies have indicated that TME can affect the efficacy of ICIs [14, 15]. Therefore, exploring some immune-related genes that affect the abundance of immune cells in the TME may help to improve the therapeutic effect of ICIs.

Consequently, to understand the differences among diverse LIHC immune subtypes, and the impact on immunotherapy and prognosis. We divided LIHC samples into three subtypes based on the degree of immune cell abundance of LIHC. In addition, we utilized machine learning to construct the prognostic model to quantify risk between individuals. Our study integrates multi-omics analyses to better predict the prognosis of LIHC patients and the effectiveness of immunotherapy.

Material And Methods

Data Acquisition and Preprocessing

TCGA-LIHC transcriptome data with related clinical information and somatic mutation data were acquired from the TCGA database ([http:// cancer geno me. nih. gov/](http://cancer.genome.nih.gov/)). We downloaded GSE52018 and GSE72467 microarray data and corresponding clinical information on LIHC from the GEO database (<https://www.ncbi.nlm.nih.gov/geo/>). For more comparability with GEO microarray data, the TCGA data were transformed using the "ComBat" function in the "sva" R package to reduce batch effects due to non-biological experimental factors.

Consensus Clustering of LIHC Samples Based on CIBERSORT Results

We quantified the immune cell infiltration in each LIHC patient by the "CIBERSORT" algorithm. The TME consisting of Immune cells and Stromal cells were quantified by the "ESTIMATE" algorithm[16]. We correlation analyzed the results of CIBERSORT and ESTIMATE, then visualized them using the "corrplot" R package. Molecular subtypes were determined for each patient using consensus clustering analysis, relying on the results of CIBERSORT. Cases were classified according to k-means, using the "ConsensusClusterPlus" R package [17]. Next, We adopted the "survival" R package to understand

differences in the survival of different subtypes, and then the "survminer" R package was employed to determine the best cutoff.

Comparison of Immune Infiltration

We visualized the results of CIBERSORT for three clusters using the "pheatmap" R package. Then we compared the Stromal Score, Immune Score, ESTIMATE Score (comprehensive scoring of stroma and immune), and Tumor Purity of the three clusters. We applied the Single-Sample Gene Set Enrichment Analysis (ssGSEA) function in the R package "GSVA", the 28 immune cell type abundance was estimated, and then compare between three clusters [18].

Evaluation of Sensitivity to Immunotherapy and Chemotherapy

First, we assessed the potential of LIHC patients to respond to immunotherapy, by comparing the expression of 13 immune checkpoint genes. We combined the Genomics of Drug Sensitivity in Cancer database (GDSC: <https://www.cancerrxgene.org>) and the "pRRophetic" R package to predict chemotherapy response for each LIHC sample [19]. The half-maximum inhibitory concentration (IC50) is used to measure the ability of the drug to induce apoptosis of tumor cells, that is, the stronger the induction ability, the lower the value is, indicating that the drug has better efficacy.

Functional Enrichment Analysis

Comparison of expression profiling data using the R package "limma" to identify the differential genes (DEGs) of Cluster-B vs Cluster-A or Cluster-C [20]. We set the threshold as $|\log_2\text{FoldChange}| > 1$ and P-value < 0.05 to obtain DEGs. We analyzed the connections between the proteins encoded by DEGs on the STRING website and mapped the protein interaction network [21]. Next, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis was adopted to understand the functions of 42 selected DEGs using the R package "clusterProfiler".

Construction of Immune Riskscore Based on DEGs

We constructed a prognostic model to assess immune-related risk factors in LIHC patients based on DEGs between three clusters. LASSO regression was carried out on 42 DEGs to reduce the number of predictors. The prognostic model was established using LASSO-COX regression analysis of selected prognostic DEGs. Finally, we calculated the Riskscore value of all LIHC samples. LIHC patients with Riskscore above the median were allocated in the high-risk group, and those with a Riskscore under the median were allocated in the low-risk group.

Validation of Hub Proteins and Verification of the Accuracy of the Model

First, to understand the survival differences in different groups, we performed the Kaplan-Meier survival analysis, then visualized using survival curves. Next, we employed the "heatmap" R package to draw a heatmap showing the presence of 5 hub genes in different risk groups. We downloaded immunohistochemical staining of partial hub gene proteins in LIHC and normal tissues from the Human Protein Atlas database (<https://www.proteinatlas.org/>). We included all predictors in univariate regression analysis followed by multivariate regression analysis. Forest maps were employed to present the results. Finally, The optimality of the prognostic model was verified by the receiver operating characteristic (ROC) curve.

Analysis of the Immune Cells Infiltrating, Somatic Mutation, and Stemness Indices

To further explore differences based on genomics among different risk groups, we [investigated](#) the differences in immune infiltration, somatic mutation, and tumor stemness index in the two risk groups. We used seven immune infiltration estimation algorithms in the TIMER2.0 database (<http://timer.cistrome.org/>) to quantify the proportion of immune cells in the two groups of samples [22-27]. Then, we visualized somatic mutations in both groups by the "maftools" R package. Tumor stemness index (mRNAsi) was generated for each sample using the one-class logistic regression (OCLR) to assess tumor dedifferentiation potential. [28]. Finally, using the "rms" R package, a nomogram was built up for the prediction of 1-, 3-, and 5-year survival rates of LIHC patients.

Results

Identification of Immune Cell Correlations and Patient Consensus Clustering

We performed correlation analysis on immune cells in LIHC samples. In **Fig. 1A**, T cells CD4 naive and Neutrophil cells with the highest positive correlation because the red sector has the largest area. NK cells activated and Macrophages M0 has the highest negative correlation. Then we identified three subtypes, Cluster-A, B, and C based on the results of CIBERSORT ($k=3$, **Fig. 1B, C**). The survival curve (**Fig. 1D**) shows that Cluster-B had the best survival, while Cluster-C had the most dismal prognosis ($p=0.032$).

Identification of Immune-Related Subtype

To further understand the difference between different subtypes, we made a heatmap to show the infiltration degree of different immune cells among each subtype (**Fig. 1E**). We found that Cluster-B had the most abundant enrichment of immune cells, involving CD8 T cells and NK cells which are associated with tumor killing. Others included T cells follicular helper and plasma cells. CD4 memory resting T cells

were significantly higher in Cluster-A, and Macrophages M0 cells were dramatically increased in Cluster-C. Then the result of ssGSEA shows that the infiltration degree of most immune cells was higher in Cluster-B too (**Fig. 1F**). Finally, by the result of ESTIMATE, we found no differences in the Stromal Scores of the three subtypes (**Fig. 1G**). Cluster-B had the highest Immune Score and ESTIMATE Score, and Cluster-C had the lowest score, which was consistent with the prognostic analysis. Cluster-C tumors have the highest purity and the worst prognosis, while Cluster-B tumors have the lowest purity and the best prognosis. Therefore, we believe that Cluster-A and Cluster-C fall into the cold tumor, while Cluster-B belongs to the hot tumor.

Response of Three Clusters to Immunotherapy and Chemotherapy

ICIs such as PD-1 and CTLA-4 are of great promise in LIHC therapy. We validated the expression of 13 immune checkpoint-related genes in 3 clusters. The results show that PD1 related (**Fig. 2A**), CTLA4 related (**Fig. 2B**), and other checkpoint genes (**Fig. 2C**) were highly expressed in Cluster-B.

In Cluster-A and Cluster-C, the expression of most immune checkpoint-related genes was not significantly different, but all were lower than in Cluster-B. Previous studies reported that cold tumors were insensitive to treatment with ICIs. In our opinion, the results of immune checkpoint genes validate our judgment on the three clusters of subtypes. We evaluated the response of chemotherapy drugs. Then, we found 9 drugs with potential predominance in Cluster-A (**Fig. 2D**), 53 in Cluster-B (**Fig. 2E**), and 9 in Cluster-C (**Fig. 2F**).

Functional Enrichment Analysis and PPI Network

Cluster B belongs to the hot tumor, while Cluster-A and Cluster-C are cold tumors. To further explore the genetic differences between the two tumors, we identified 42 DEGs in Cluster-B vs Cluster-A or Cluster-C. The results in GO (**Fig. 3A**) and KEGG (**Fig. 3B**) enrichment analysis suggest that the differences between hot and cold tumors may involve PD-L1 expression and PD-1 checkpoint pathway in cancer, and other immune pathways. In addition, there are proliferation, differentiation, and activation of various immune cells. Finally, we mapped the protein interaction network of 42 DEGs (**Fig. 3C**).

Establish the Prognostic Model

Considering the immune heterogeneity of the 3 Cluster, we have established a prognostic model based on this heterogeneity. We incorporated 42 DEGs into the Lasso regression analysis to reduce the number of predicted variables. Then, five hub genes were selected by the LASSO-COX regression analysis (**Fig. 3D**). The formula of Riskscore is as follows: . Exp is the expression value of the hub gene (Table S1). **Fig. 3E** and **3F** show the distribution of Riskscore and the survival status of LIHC patients. The K-M survival curve shows that the low-risk group has a better prognosis ($p < 0.001$) (**Fig. 3G**). **Fig. 3H** exhibits the distribution of various clinical features and hub gene expression levels in two risk groups. The expression of EOMES, GZMH, NKG7, and CD8A are higher in the low-risk group. The expression of ANKRD22 is higher in the high-risk group. Furthermore, immunohistochemical staining revealed differences in partial

hub genes between LIHC and normal tissues (**Fig. 3I**). **Fig. 3J** shows higher mortality in the high-risk group, and **Fig. 3K** indicates that there is a significant difference in Riskscore between dead and alive patients ($P < 0.001$).

We included age, gender, stage, and Riskscore in univariate (**Fig. 3L**) and multivariate (**Fig. 3M**) regression analysis. Results reveal that Riskscore is an independent prognostic factor for LIHC patients.

To assess the predictive accuracy of our models, we employed the ROC curve (**Fig. 3N**). The area under the curve (AUC) value (0.68) of Riskscore is higher than other clinicopathological characteristics, indicating that Riskscore has the best predictive ability.

Immune Infiltration and Sensitivity to Immunotherapy

As shown in **Fig. 4A**, the low-risk group had higher tumor infiltration of diverse anti-cancer immune cells, including CD8+ T cells, macrophages, Th1 cells, NK cells, and dendritic cells. However, M2 macrophages, which might have a suppressive TME, are more accumulated in the high-risk group. **Fig. 4B** shows that 21 immune cells were more strongly infiltrated in the low-risk group. The results in **Fig. 4C** indicate that the low-risk group had more abundant stromal components and immune infiltration. The higher tumor purity in high-risk groups may lead to a poor prognosis. The PD1 related (**Fig. 4D**), CTLA4 related (**Fig. 4E**), other immune checkpoint genes (**Fig. 4F**), and T cell activation agonists (**Fig. 4G**) are apparently higher in the low-risk group. In consideration of the positive correlation between the expression of the immune checkpoint genes and the response to immunotherapy, the low-risk group may benefit much more from immunotherapy. In conclusion, the low-risk group has stronger antitumor immunity, thereby displaying better immunotherapy response and has a more favorable prognosis.

Somatic mutation and stem cell index in two groups, and Nomogram construction

To further investigate the heterogeneity of the two risk groups, we analyzed the somatic mutation distribution and stem cell index. We found that in the TCGA-LIHC cohort, the high-risk group had a higher somatic mutation rate. (**Fig. 5A, B**). The mutation landscape showed the top 30 mutant genes. Then we compared the cancer stem cell index of the two groups and observed a higher stem cell index in the high-risk group (**Fig. 5C**). The Riskscore had a weaker positive correlation with the stem cell index (**Fig. 5D**). **Fig. 5E** is a nomogram predicting the overall survival of LIHC, combined with Riskscore and clinical features.

Discussion

Despite rapid advances in treatment and surgical techniques, the incidence and mortality of LIHC remain high. In the past few decades, immunotherapy for LIHC has made tremendous progress. Immunotherapy with ICIs improved outcomes in patients with LIHC. However, due to the immune complexity and the heterogeneity of LIHC, immunotherapy was only effective in 10-25% of patients. In recent years, the rapidly developing microarray technology has been widely used to predict disease progression, and

assess immunotherapy and prognosis. In the present study, LIHC samples with different immune cell infiltration were classified into 3 molecular subtypes by the results of transcriptomic data analysis. In this research, multi-omics data and clinical data were used to analyze the immune heterogeneity of LIHC patients, and to construct a prognostic model to evaluate patient prognosis and predict the effect of immunotherapy.

To obtain LIHC typing based on immune cell infiltration, we first calculated the degree of immune infiltration in all samples using CIBERSORT. Next, three immunophenotypes were comprehensively identified based on the consensus cluster analysis of the CIBERSORT results. The further study we found that the prognosis and the degree of immune infiltration were significantly different among the three clusters. Cluster-B has the highest degree of anti-tumor immune cells. We believed that the immune infiltration characteristics of Cluster-B help prolong the lifespan of patients. In contrast, Cluster-C has lower immune infiltration and the highest tumor purity and therefore the worst prognosis. Analysis of immune checkpoint-related genes showed that the expression in Cluster-B was significantly higher than in the other two clusters. Therefore, the immunotherapy effect of Cluster-B may be better. Currently, researchers are increasingly focusing on the immunophenotype of tumors. Tumors were divided into hot and cold tumors according to the T cell infiltration in the TME [29]. According to the infiltration of immune cells, tumors were divided into immune-inflamed type, immune-excluded type, and immune-desert type [29]. The immune-inflamed type contains a large number of CD8⁺ T cells infiltrate in the tumor, and PD-1/PD-L1 is highly expressed. The CD8⁺ T cells of immune-excluded tumors mainly infiltrate the stroma and cannot effectively invade the tumor. CD8⁺ T cells are lacking in the immune-desert type. Therefore, the immune-inflamed type is hot tumors. Besides weak infiltration of T cells, the immune-excluded type and immune-desert type also have the characteristics of low expression of PD-1/PD-L1, which belong to cold tumors [30]. A clinical trial shows that compared with cold tumors, the hot tumors may be more likely to benefit from immunotherapy because of the significantly higher PD-1/PD-L1 expression [31]. Our results show that Cluster-B has higher PD-1/PD-L1 expression than the other two clusters. At the same time, the CD8⁺ T cell infiltration of Cluster-B was also the highest. Based on our results, we believe that Cluster-B belongs to hot tumors, while Cluster-A and Cluster-C are cold tumors. Cluster-B may be more sensitive to immunotherapy.

Next, to reveal the differences between hot and cold tumors, we obtained the DEGs of the Cluster-B vs Cluster-A or Cluster-C. Then we employed KEGG and GO analysis on the DEGs. The results of GO enrichment analysis show that these genes mainly take an important part in the proliferation, differentiation, activation of immune cells, and regulation of immune responses. For the KEGG pathway, DEGs were involved in immune-related processes, including PD-L1 expression and PD-1 checkpoint pathway in cancer.

We screened 5 hub genes (EMOES, ANKRD22, GZMH, NKG7, CD8A) by LASSO regression analysis on 42 DEGs. The predictive model was established based on hub genes to evaluate the sensitivity to immunotherapy and prognosis of LIHC patients. EOMES is highly expressed in CD8⁺ T cells expressing inhibitory receptors, which encodes transcription factors required for the differentiation of effector CD8⁺

T cells and is essential for CD8⁺ T cells which play an important part in the antitumor immune. [32]. ANKRD22, an oncogene in non-small cell lung cancer, promotes tumor advance mainly by upregulating E2F1 transcription and enhancing cell proliferation [33]. In breast cancer, ANKRD22 was found to enhance the malignant degree of breast cancer cells by regulating NuSAP1 expression to activate the Wnt/ β -catenin pathway [34]. GZMH is constitutively expressed in NK cells and has a significant role in the innate immune response against tumors and viruses. Studies have shown that low expression of GZMH in cytotoxic lymphocytes may predispose to HBV infection and hepatocellular carcinoma. [35]. NKG7 was upregulated on CD8⁺ T cells and NK cells for antitumor immunity and was required for T cell accumulation in the TME [36]. Wen et al. reported that NKG7 is an essential component of the cytotoxic function of CD8⁺ T cells and has been established as a T cell-intrinsic therapeutic target for enhanced cancer immunotherapy [37]. CD8A encodes the α chain of the CD8 antigen. The CD8 antigen played a crucial role in mediating cell-cell interactions within the immune system. which is a cell surface glycoprotein located on cytotoxic T lymphocytes. Combined with the current research, we believe that EMOES, GZMH, NKG7, and CD8A belong to tumor suppressor genes, which were highly expressed in the low-risk group. ANKRD22 is an oncogene with higher expression in high-risk groups.

There are profound clinical implications of Riskscore as we expected. First, Riskscore can efficiently forecast the prognosis of LIHC patients, and the low-risk patients have longer survival. Significantly, Riskscore is more precise than other clinical parameters. Second, tumors can be distinguished between cold or hot according to the Riskscore. Overall, the two-group classified by Riskscore showed a significant discrepancy in immune infiltration, immune score, and expression of immune checkpoint-related genes. In this study, patients in the low-risk group were classified as hot tumors, displaying the rich infiltration of activated CD8⁺ T cells and a higher expression of PD-1/PD-L1. In the low-risk group, not only for PD1, but the expression of the other immune checkpoints (CTLA4, LAG3, PRF1, TNF, CXCL9, GZMB, CD8A, GZMA) was higher than those in another group. Therefore, it is suggested that immunotherapy might be more effective in the low-risk group, whereas the high-risk group is the opposite. Third, there were more TP53 gene mutations in the high-risk group. The tumor suppressor gene TP53 mutations were the most common genetic changes in HCC, with an average of 30% [38]. In HBV-related HCC patients, the frequency of TP53 mutation can reach nearly 60% [39]. TP53 mutation is closely related to tumor differentiation, vascular invasion, serum alpha-fetoprotein (AFP) level, tumor stage, and the immune microenvironment in HCC [38]. TP53 mutation may affect patient prognosis, and HCC patients with this mutation have shorter overall survival and recurrence-free survival [38, 40]. CTNNB1 mutations were the highest mutation in the low-risk group, but lower than those in the high-risk group. Mutations in CTNNB1 can promote tumor advance by activating the Wnt/ β catenin pathway and are found in approximately 19%-26% of liver cancers [41, 42]. Intrinsic activation of the Wnt/ β -catenin pathway is associated with the lack of T cell infiltration within tumors and resistance to PDL1/CTLA4 immunotherapy [42-44]. Clinical studies have shown that Wnt/CTNNB1 mutations characterize the immune-exclusion category (cold tumors), and LIHC patients with these mutations are resistant to PD1 immunotherapy [45]. Fourth, we found that Riskscore had a weak positive correlation with tumor stemness index, and there were significant differences in stemness index between high- and low-risk groups. Cancer stem cells are the

crucial factors that contribute to cancer recurrence, metastasis, chemotherapy resistance, and patient prognosis in HCC [46]. Higher stemness index values correlated with active biological processes and greater tumor dedifferentiation in cancer stem cells, and metastatic tumor cells exhibit more dedifferentiation phenotypically, which may contribute to their aggressiveness. The high-risk group showed a higher stemness index, which may lead to a worse prognosis. Finally, we constructed a nomogram combining Riskscore and clinicopathological features to provide a novel approach to predicting overall survival in LIHC patients.

Conclusion

The present study constructed a valid, innovative, and reliable immune-related prognostic model (EOMES, ANKRD22, GZMH, NKG7, CD8A) to forecast the prognosis of LIHC patients and the sensitivity to immunotherapy. Our model is an independent risk factor for LIHC. Further, we constructed a nomogram for predicting OS in patients with LIHC.

Abbreviations

LIHC: Liver hepatocellular carcinoma; TCGA: The Cancer Genome Atlas; GEO: Gene Expression Omnibus; ICIs: immune checkpoint inhibitors; Tregs: regulatory T cells; MDSC: myeloid-derived suppressor cells; TME: Tumor microenvironment; ssGSEA: Single-Sample Gene Set Enrichment Analysis; IC50: half-maximum inhibitory concentration; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes;

Declarations

Acknowledgments

We thank the TCGA and GEO databases, and the authors who uploaded the relevant data.

Authors' contributions

RJK designed the overall study and wrote the original draft. YC and ZFX participated in the literature review and edited the manuscript. ZKN and ZT are responsible for proofreading. All authors read and approved the final manuscript.

Availability of data and materials

The datasets for this study can be found in the Cancer Genome Atlas (TCGA) database and the Gene Expression Omnibus (GEO) database.

Declarations

Ethics approval and consent to participate

The authors have no ethical conflicts to disclose.

Consent for publication

All authors agree with the paper's content.

Competing interests

There are no competing interests to declare.

Funding

This work is supported by the Natural Science Foundation of Guangdong Province (No.2019A1515011850 and No.2022A1515012224)

References

1. Llovet JM, Kelley RK, Villanueva A, Singal AG, Pikarsky E, Roayaie S, Lencioni R, Koike K, Zucman-Rossi J, Finn RS: **Hepatocellular carcinoma**. *Nat Rev Dis Primers* 2021, **7**:6.
2. Farazi PA, DePinho RA: **Hepatocellular carcinoma pathogenesis: from genes to environment**. *Nat Rev Cancer* 2006, **6**:674-687.
3. Anwanwan D, Singh SK, Singh S, Saikam V, Singh R: **Challenges in liver cancer and possible treatment approaches**. *Biochim Biophys Acta Rev Cancer* 2020, **1873**:188314.
4. Webster GJ, Reignat S, Maini MK, Whalley SA, Ogg GS, King A, Brown D, Amlot PL, Williams R, Vergani D, et al: **Incubation phase of acute hepatitis B in man: dynamic of cellular immune mechanisms**. *Hepatology* 2000, **32**:1117-1124.
5. Kudo M: **Scientific Rationale for Combination Immunotherapy of Hepatocellular Carcinoma with Anti-PD-1/PD-L1 and Anti-CTLA-4 Antibodies**. *Liver Cancer* 2019, **8**:413-426.
6. Jia L, Gao Y, He Y, Hooper JD, Yang P: **HBV induced hepatocellular carcinoma and related potential immunotherapy**. *Pharmacol Res* 2020, **159**:104992.
7. Chambers CA, Kuhns MS, Egen JG, Allison JP: **CTLA-4-mediated inhibition in regulation of T cell responses: mechanisms and manipulation in tumor immunotherapy**. *Annu Rev Immunol* 2001, **19**:565-594.
8. Dong H, Zhu G, Tamada K, Chen L: **B7-H1, a third member of the B7 family, co-stimulates T-cell proliferation and interleukin-10 secretion**. *Nat Med* 1999, **5**:1365-1369.
9. Freeman GJ, Long AJ, Iwai Y, Bourque K, Chernova T, Nishimura H, Fitz LJ, Malenkovich N, Okazaki T, Byrne MC, et al: **Engagement of the PD-1 immunoinhibitory receptor by a novel B7 family member leads to negative regulation of lymphocyte activation**. *J Exp Med* 2000, **192**:1027-1034.
10. Sangro B, Gomez-Martin C, de la Mata M, Inarrairaegui M, Garralda E, Barrera P, Riezu-Boj JI, Larrea E, Alfaro C, Sarobe P, et al: **A clinical trial of CTLA-4 blockade with tremelimumab in patients with hepatocellular carcinoma and chronic hepatitis C**. *J Hepatol* 2013, **59**:81-88.

11. El-Khoueiry AB, Sangro B, Yau T, Crocenzi TS, Kudo M, Hsu C, Kim TY, Choo SP, Trojan J, Welling THR, et al: **Nivolumab in patients with advanced hepatocellular carcinoma (CheckMate 040): an open-label, non-comparative, phase 1/2 dose escalation and expansion trial.** *Lancet* 2017, **389**:2492-2502.
12. Zhu AX, Finn RS, Edeline J, Cattan S, Ogasawara S, Palmer D, Verslype C, Zagonel V, Fartoux L, Vogel A, et al: **Pembrolizumab in patients with advanced hepatocellular carcinoma previously treated with sorafenib (KEYNOTE-224): a non-randomised, open-label phase 2 trial.** *Lancet Oncol* 2018, **19**:940-952.
13. Topalian SL, Drake CG, Pardoll DM: **Immune checkpoint blockade: a common denominator approach to cancer therapy.** *Cancer Cell* 2015, **27**:450-461.
14. Barry KC, Hsu J, Broz ML, Cueto FJ, Binnewies M, Combes AJ, Nelson AE, Loo K, Kumar R, Rosenblum MD, et al: **A natural killer-dendritic cell axis defines checkpoint therapy-responsive tumor microenvironments.** *Nat Med* 2018, **24**:1178-1191.
15. Lambrechts D, Wauters E, Boeckx B, Aibar S, Nittner D, Burton O, Bassez A, Decaluwe H, Pircher A, Van den Eynde K, et al: **Phenotype molding of stromal cells in the lung tumor microenvironment.** *Nat Med* 2018, **24**:1277-1289.
16. Yoshihara K, Shahmoradgoli M, Martinez E, Vegesna R, Kim H, Torres-Garcia W, Trevino V, Shen H, Laird PW, Levine DA, et al: **Inferring tumour purity and stromal and immune cell admixture from expression data.** *Nat Commun* 2013, **4**:2612.
17. Wilkerson MD, Hayes DN: **ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking.** *Bioinformatics* 2010, **26**:1572-1573.
18. Hanzelmann S, Castelo R, Guinney J: **GSEA: gene set variation analysis for microarray and RNA-seq data.** *BMC Bioinformatics* 2013, **14**:7.
19. Geeleher P, Cox N, Huang RS: **pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels.** *PLoS One* 2014, **9**:e107468.
20. 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.
21. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, et al: **STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets.** *Nucleic Acids Res* 2019, **47**:D607-D613.
22. Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, Khodadoust MS, Esfahani MS, Luca BA, Steiner D, et al: **Determining cell type abundance and expression from bulk tissues with digital cytometry.** *Nat Biotechnol* 2019, **37**:773-782.
23. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, Selves J, Laurent-Puig P, Sautès-Fridman C, Fridman WH, de Reyniès A: **Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression.** *Genome Biol* 2016, **17**:218.
24. Racle J, Gfeller D: **EPIC: A Tool to Estimate the Proportions of Different Cell Types from Bulk Gene Expression Data.** *Methods Mol Biol* 2020, **2120**:233-248.

25. Li T, Fu J, Zeng Z, Cohen D, Li J, Chen Q, Li B, Liu XS: **TIMER2.0 for analysis of tumor-infiltrating immune cells.** *Nucleic Acids Res* 2020, **48**:W509-W514.
26. Aran D, Hu Z, Butte AJ: **xCell: digitally portraying the tissue cellular heterogeneity landscape.** *Genome Biol* 2017, **18**:220.
27. Finotello F, Mayer C, Plattner C, Laschober G, Rieder D, Hackl H, Krogsdam A, Loncova Z, Posch W, Wilflingseder D, et al: **Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data.** *Genome Med* 2019, **11**:34.
28. Malta TM, Sokolov A, Gentles AJ, Burzykowski T, Poisson L, Weinstein JN, Kaminska B, Huelsken J, Omberg L, Gevaert O, et al: **Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation.** *Cell* 2018, **173**:338-354 e315.
29. Chen DS, Mellman I: **Elements of cancer immunity and the cancer-immune set point.** *Nature* 2017, **541**:321-330.
30. Galon J, Bruni D: **Approaches to treat immune hot, altered and cold tumours with combination immunotherapies.** *Nat Rev Drug Discov* 2019, **18**:197-218.
31. Idel C, Ribbat-Idel J, Klapper L, Krupar R, Bruchhage KL, Dreyer E, Rades D, Polasky C, Offermann A, Kirfel J, et al: **Spatial Distribution of Immune Cells in Head and Neck Squamous Cell Carcinomas.** *Front Oncol* 2021, **11**:712788.
32. Llaó-Cid L, Roessner PM, Chapaprieta V, Öztürk S, Roider T, Bordas M, Izcue A, Colomer D, Dietrich S, Stilgenbauer S, et al: **EOMES is essential for antitumor activity of CD8(+) T cells in chronic lymphocytic leukemia.** *Leukemia* 2021, **35**:3152-3162.
33. Yin J, Fu W, Dai L, Jiang Z, Liao H, Chen W, Pan L, Zhao J: **ANKRD22 promotes progression of non-small cell lung cancer through transcriptional up-regulation of E2F1.** *Sci Rep* 2017, **7**:4430.
34. Wu Y, Liu H, Gong Y, Zhang B, Chen W: **ANKRD22 enhances breast cancer cell malignancy by activating the Wnt/beta-catenin pathway via modulating NuSAP1 expression.** *Bosn J Basic Med Sci* 2021, **21**:294-304.
35. Tang H, Li C, Wang L, Zhang H, Fan Z: **Granzyme H of cytotoxic lymphocytes is required for clearance of the hepatitis B virus through cleavage of the hepatitis B virus X protein.** *J Immunol* 2012, **188**:824-831.
36. Li XY, Corvino D, Nowlan B, Aguilera AR, Ng SS, Braun M, Cillo AR, Bald T, Smyth MJ, Engwerda CR: **NKG7 Is Required for Optimal Antitumor T-cell Immunity.** *Cancer Immunol Res* 2022, **10**:154-161.
37. Wen T, Barham W, Li Y, Zhang H, Gicobi JK, Hirdler JB, Liu X, Ham H, Peterson Martinez KE, Lucien F, et al: **NKG7 Is a T-cell-Intrinsic Therapeutic Target for Improving Antitumor Cytotoxicity and Cancer Immunotherapy.** *Cancer Immunol Res* 2022, **10**:162-181.
38. Long J, Wang A, Bai Y, Lin J, Yang X, Wang D, Yang X, Jiang Y, Zhao H: **Development and validation of a TP53-associated immune prognostic model for hepatocellular carcinoma.** *EBioMedicine* 2019, **42**:363-374.
39. Gao Q, Zhu H, Dong L, Shi W, Chen R, Song Z, Huang C, Li J, Dong X, Zhou Y, et al: **Integrated Proteogenomic Characterization of HBV-Related Hepatocellular Carcinoma.** *Cell* 2019, **179**:1240.

40. Liu J, Ma Q, Zhang M, Wang X, Zhang D, Li W, Wang F, Wu E: **Alterations of TP53 are associated with a poor outcome for patients with hepatocellular carcinoma: evidence from a systematic review and meta-analysis.** *Eur J Cancer* 2012, **48**:2328-2338.
41. Kim S, Jeong S: **Mutation Hotspots in the β -Catenin Gene: Lessons from the Human Cancer Genome Databases.** *Mol Cells* 2019, **42**:8-16.
42. Rebouissou S, Franconi A, Calderaro J, Letouze E, Imbeaud S, Pilati C, Nault JC, Couchy G, Laurent A, Balabaud C, et al: **Genotype-phenotype correlation of CTNNB1 mutations reveals different ss-catenin activity associated with liver tumor progression.** *Hepatology* 2016, **64**:2047-2061.
43. Liu S, Ding G, Zhou Z, Feng C: **β -Catenin-driven adrenocortical carcinoma is characterized with immune exclusion.** *Onco Targets Ther* 2018, **11**:2029-2036.
44. Spranger S, Bao R, Gajewski TF: **Melanoma-intrinsic β -catenin signalling prevents anti-tumour immunity.** *Nature* 2015, **523**:231-235.
45. Pinyol R, Sia D, Llovet JM: **Immune Exclusion-Wnt/CTNNB1 Class Predicts Resistance to Immunotherapies in HCC.** *Clin Cancer Res* 2019, **25**:2021-2023.
46. Lin X, Zuo S, Luo R, Li Y, Yu G, Zou Y, Zhou Y, Liu Z, Liu Y, Hu Y, et al: **HBX-induced miR-5188 impairs FOXO1 to stimulate β -catenin nuclear translocation and promotes tumor stemness in hepatocellular carcinoma.** *Theranostics* 2019, **9**:7583-7598.

Figures

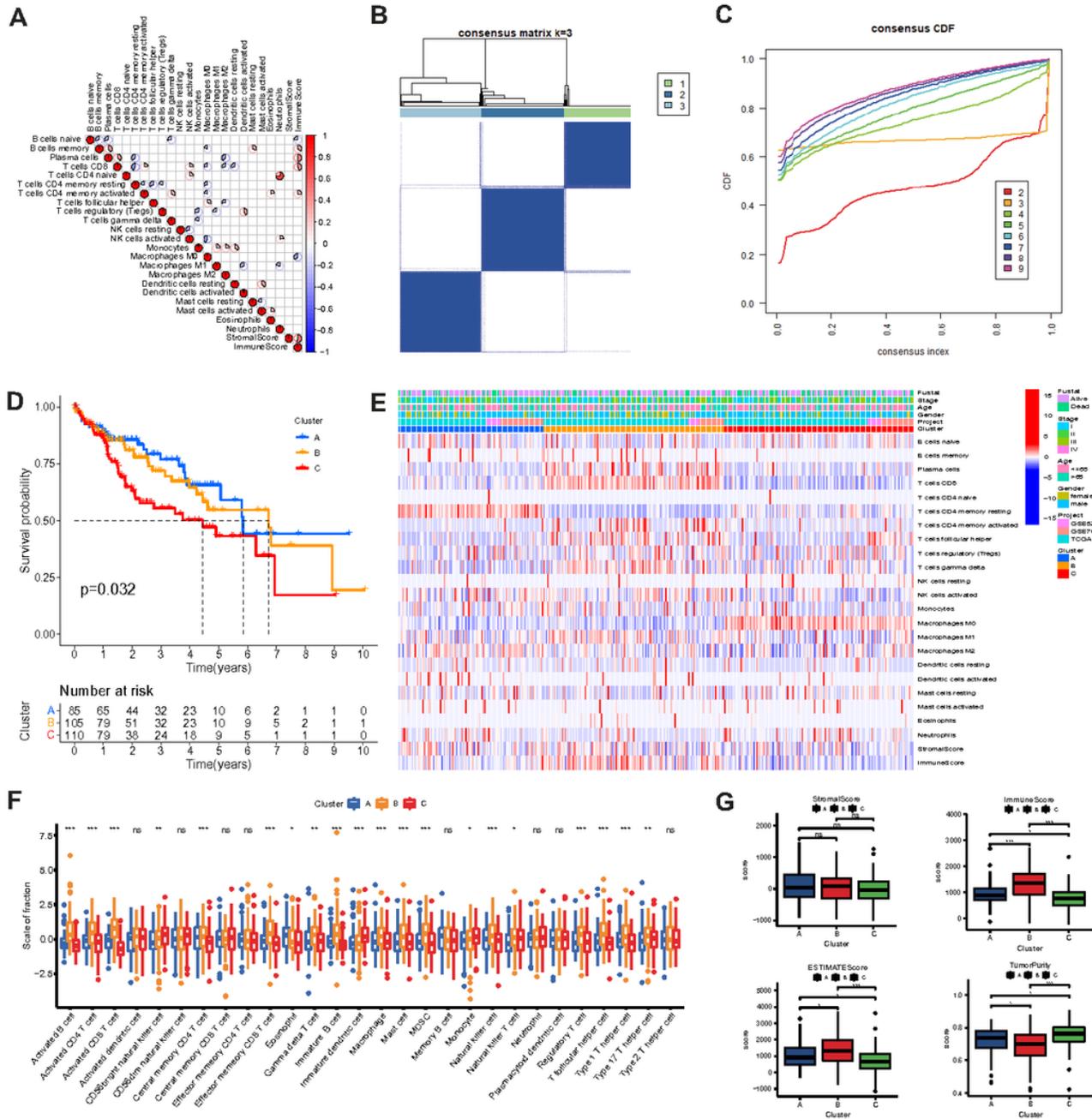


Figure 1

Consensus clusters and immunological signatures for each cluster. **(A)** Correlation analysis of 22 types of immune cells, Immune Score, and Stroma Score in LIHC samples. **(B)** The clustering heatmap corresponds to the consensus matrix when k=3 is obtained by consensus clustering. **(C)** CDF plot displays consensus distributions for k=2 to 9. **(D)** Kaplan-Meier curves with Log-rank test show survival differences for the three clusters. **(E)** Heatmap for immune cell infiltration and immune score based on CIBERSORT, ESTIMATE among three clusters. **(F)** Differences in immune cell infiltration among the three clusters, based on ssGSEA. **(G)** Differences in Stroma Score, Immune Score, ESTIMATE Score, and Tumor Purity

Purity among the three clusters. The P values are labeled using asterisks (ns, no significance, *P < 0.05, **P < 0.01, ***P < 0.001).

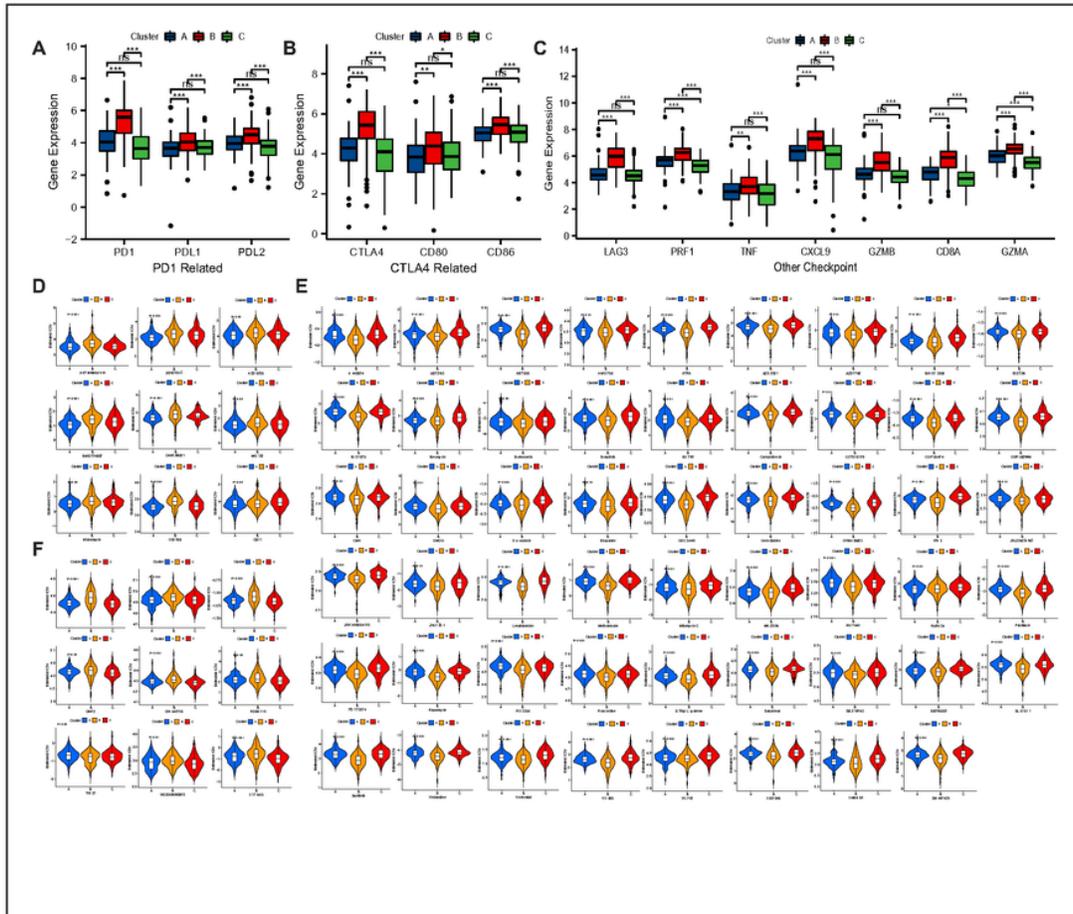


Figure 2

Immunotherapy and chemotherapeutic sensitivity of each cluster (**A-C**) Differences in the expression of immune checkpoint genes among three clusters. Sensitivity of commonly used chemotherapy drugs in Cluster-A (**D**), Cluster-B (**E**), Cluster-C (**F**). The statistical difference of clusters was compared using the Wilcoxon test. *P < 0.05, **P < 0.01; ***P < 0.001; ns, not significant.

Figure 3

Construction of Riskscore scoring system. **(A)** GO enrichment analysis of DEGs. **(B)** KEGG enrichment analysis of DEGs. **(C)** PPI network of DEGs. **(D)** LASSO-COX regression model construction. **(E)** The risk score curve and **(F)** the distribution of survival status of LIHC patients. **(G)** Survival analyses for the low- and high-risk groups using Kaplan–Meier curve and Log-rank test. **(H)** The heatmap showed the expression of five hub genes in high- and low-risk groups. **(I)** Immunohistochemical staining of some hub genes in LIHC samples and normal samples **(J)** The proportion of survival status in high- and low-risk groups. **(K)** Riskscore differences in patients with different survival status. **(L)** The univariate and **(M)** multivariate independent prognostic analysis of different clinical traits. **(N)** The ROC curve verifies the reliability of the model.

Figure 4

Immune infiltration in high- and low-risk groups. **(A)** Heatmap for immune cell infiltration based on TIMER, CIBERSORT, CIBERSORT-ABS, quanTIseq, MCP-counter, xCell, and EPIC algorithms among two groups. **(B)** Expression of immune cells between the two groups, based on the ssGSEA algorithm. **(C)** Comparison of Stromal Score, Immune Score, ESTIMATE Score, Tumor Purity between two clusters. **(D-F)** The immune checkpoint genes and **(G)** T cell activations agonists expression between the two groups. The P values are labeled using asterisks (ns, no significance, *P < 0.05, **P < 0.01, ***P < 0.001).

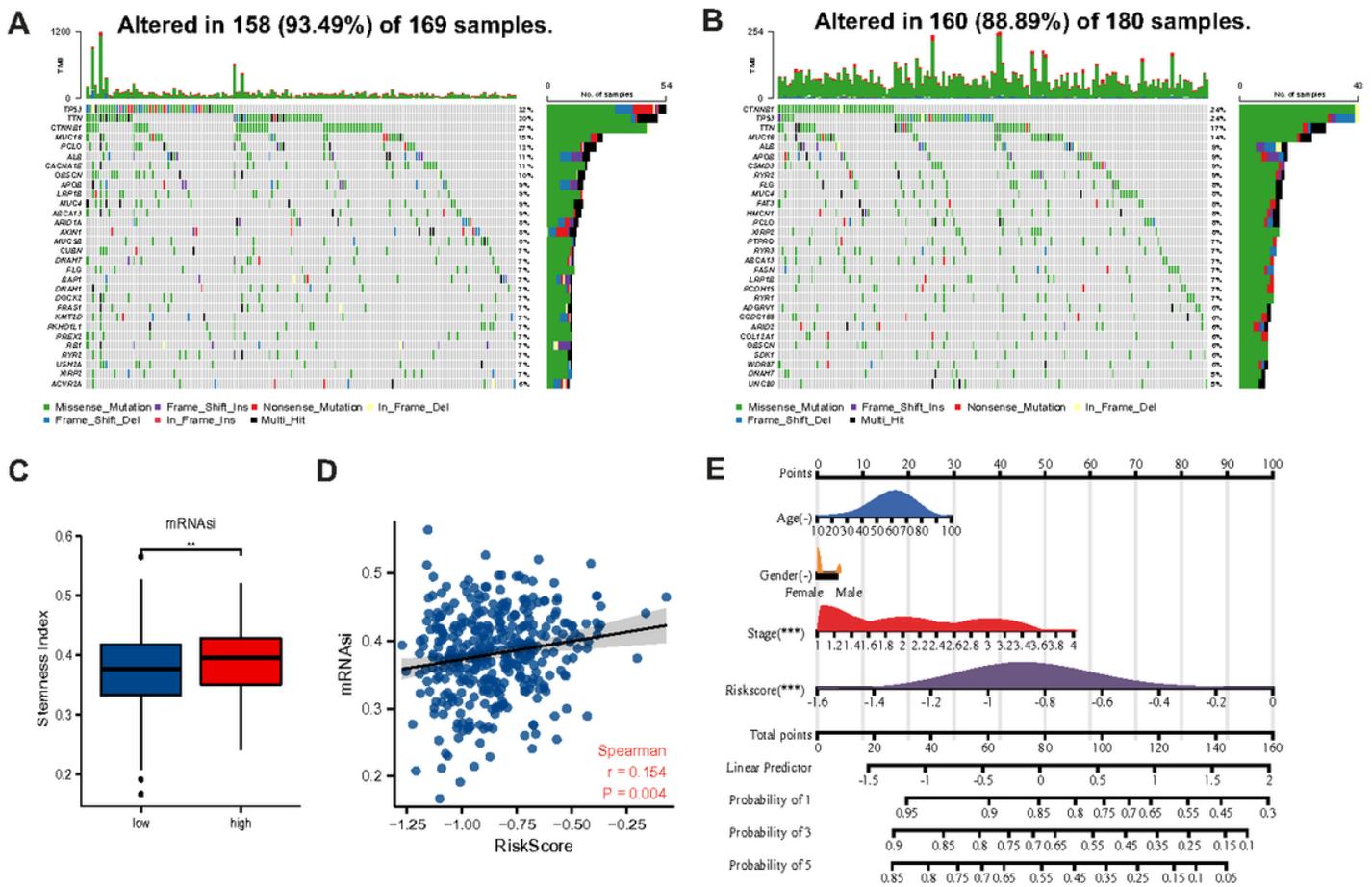


Figure 5

The oncoPrint plot shows tumor somatic mutations in the **(A)** high- and **(B)** low-risk groups. **(C)** Differences in tumor stemness index between high- and low-risk groups. **(D)** Correlation between tumor stemness index and RiskScore. **(E)** A prognostic nomogram predicting 1-, 2-, and 3-year overall survival of LIHC.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TableS1.pdf](#)