Identification of fatty acid metabolism-related clusters and immune infiltration features in hepatocellular carcinoma

Zhixuan Ren[1,2], Duan Gao[3], Yue Luo[1,2], Zhenghui Song[1,2], Guojing Wu[1,2], Na Qi[1,3], Aimin Li[1,2,*], Xinhui Liu[1,2,*]

[1]Integrated Hospital of Traditional Chinese Medicine, Southern Medical University, Guangzhou 510315, China
[2]Cancer Center, Southern Medical University, Guangzhou 510315, China
[3]Department of Pharmacy, Guilin Medical University, Guilin 541004, China
[*]Equal contribution

Correspondence to: Xinhui Liu, Aimin Li; email: jkkl@smu.edu.cn, liaimin@smu.edu.cn

ABSTRACT

Hepatocellular Carcinoma (HCC) is a type of liver cancer which is characterized by inflammation-associated tumor. The unique characteristics of tumor immune microenvironment in HCC contribute to hepatocarcinogenesis. It was also clarified that aberrant fatty acid metabolism (FAM) might accelerate tumor growth and metastasis of HCC. In this study, we aimed to identify fatty acid metabolism-related clusters and establish a novel prognostic risk model in HCC. Gene expression and corresponding clinical data were searched from the Cancer Genome Atlas (TCGA) and the International Cancer Genome Consortium (ICGC) portal. From the TCGA database, by unsupervised clustering method, we determined three FAM clusters and two gene clusters with distinct clinicopathological and immune characteristics. Based on 79 prognostic genes identified from 190 differentially expressed genes (DEGs) among three FAM clusters, five prognostic DEGs (CCDC112, TRNP1, CFL1, CYB5D2, and SLC22A1) were determined to construct risk model by least absolute shrinkage and selection operator (LASSO) and multivariate cox regression analysis. Furthermore, the ICGC dataset was used to validate the model. In conclusion, the prognostic risk model constructed in this study exhibited excellent indicator performance of overall survival, clinical feature, and immune cell infiltration, which has the potential to be an effective biomarker for HCC immunotherapy.

INTRODUCTION

As a major health burden in the world, liver cancer is expected to affect more than one million people by 2025 [1]. The most common primary liver cancer, hepatocellular carcinoma (HCC), ranks fourth among all cancer-related deaths [1]. HCC patients in early stage can be cured by resection, transplantation, thermal ablation and TACE [2]. Early detection of HCC can increase the possibility of potentially curative treatment. Nevertheless, since early HCC diagnosis is challenging, the prognosis of HCC patients remains dismal. HCC patients with intermediate-stage can benefit from catheter-based locoregional treatment [3]. The multitargeted Tyrosine kinase inhibitors (TKI) sorafenib and lenvatinib were approved for the treatment of advanced-stage HCC [4]. A subset of patients treated with immune checkpoint inhibitors has demonstrated strong anti-tumor activity [5]. Identifying and validating predictive biomarkers is a major challenge for HCC immunotherapy. Thus, it is imperative to search novel molecular biomarkers to improve the diagnostic accuracy and guide therapies for HCC patients.

In HCC, cancer cells undergo considerable metabolic reprogramming when preparing to proliferate [6]. It has been clear that lipid metabolic rewiring is an influential metabolic alteration in cancer cells. Fatty acid is an integral component of lipid metabolism, it participates in membrane synthesis, storage of energy, and production of signaling molecules [7]. Over the past few years, there has been expanding understanding of the role of fatty acid metabolism (FAM) in tumor progression [8]. Cancer cells can obtain fatty acids from both intracellular and extracellular sources, and changes in fatty acid metabolism are characteristics of oncogenesis and metastasis [9]. By enhancing lipid synthesis, storage and degradation, aberrant fatty acid metabolism impacts the biology of cancer cells to drive tumorigenesis and disease progression [8]. A recent study has uncovered that fatty acid level influenced by cancer cell fatty acid metabolism can change CD8+ T cell activity [10]. It also found that tumor and immune cells compete for fatty acids, which promotes tumor growth [10]. According to a study, fatty acid chain lengthening has been determined as a distinguishing feature of lung cancer [11]. Increasing evidences indicated that fatty acids may contribute to the cancer initiation and development such as gastric cancer, colorectal cancer and breast cancer [12–14]. Deregulated fats can also affect the efficacy of chemotherapy and radiation therapy for cancer patients [15, 16], as well as the effectiveness of immunotherapy. Treatments that target deregulated fatty acids and the inhibition of immune checkpoints in cancer may augment each other's effects [17]. HCC prevention and treatment may benefit greatly from an understanding of fatty acid metabolism heterogeneity, nonetheless, few studies that investigate possible mechanism and prognostic value of fatty acid metabolism-related genes (FAMs) have been conducted in HCC.

In this study, we explored the fatty acid metabolism-related clusters and assessed the composition of tumor microenvironment (TME) in HCC. First, based on expression of 49 FAMs, we identified 3 FAM clusters with distinct biological pathways and immune characteristics. Then 2 gene clusters were determined according to 190 DEGs retrieved from 3 FAM clusters. Afterward, based on the prognostic value of 190 DEGs, we established a prognostic model. Finally, the reliability of the model and the immune landscape of HCC samples were determined.

MATERIALS AND METHODS

Data source

On TCGA website (https://portal.gdc.cancer.gov/), gene expression information (fragments per kilobase million,

FPKM) and clinical characteristics of 371 HCC patients were obtained. From ICGC database (https://icgc.org/), we acquired information of another 231 HCC patients, including RNA-seq data and clinical features [18]. Based on previous descriptions, we have transformed the LIHC (liver hepatocellular carcinoma) FPKM values into TPM (transcripts per kilobase million) values [19]. Through the GeneCards database, using "fatty acid metabolism" as a keyword, the fatty acid metabolism-related genes (FAMs) was searched and screened. Then, with a relevance score ≥ 50, 49 FAMs were retrieved for the next analyses and provided in Supplementary Table 1. In order to assess mutation states of FAM-related genes in HCC samples, mutation data was processed by "maftools" R package [20].

Consensus clustering for FAMs

As a result of consensus unsupervised clustering analysis, HCC patients were categorized into different clusters by the R package "ConsensusClusterPlus" based on the FAMs expression [21]. Using the R packages "survival" and "survminer", we tested whether there are any differences in survival time between distinct clusters using Kaplan-Meier curves. A heatmap plot of the clinical and pathological characteristics was created using R's "pheatmap" package. From the MSigDB (molecular signatures database) (https://www.gsea-msigdb.org/gsea/msigdb) we extracted the hallmark gene sets (c2.cp.kegg.v7.5.1) and performed gene set variation analysis (GSVA) to determine different biological processes between distinct clusters.

Immune landscape analysis

From previous literature, we gathered the gene sets for immune cells [22], and collected cancer-related gene signatures using the MSigDB. The level of immune cell infiltration and cancer-related gene signatures in the HCC tumor microenvironment were evaluated using single-sample gene set enrichment analysis (ssGSEA) [23]. As well, we compared expression levels of several immune checkpoints and HLA genes between different clusters.

Analysis of DEGs

With an adjusted *p*-value of 0.001, we identified DEGs among different FAM clusters by the "limma" package in R. By using the "clusterprofiler" R package, we performed functional enrichment analyses on the DEGs to determine their potential functions and pathways. Furthermore, by using a method of unsupervised clustering based prognostic DEGs expression, HCC samples were classified into different clusters (FAM gene cluster D and FAM gene cluster E) for deeper analysis.

Generation of the FAM-related prognostic model

Prognostic analysis of DEGs was conducted using univariate cox regression. Afterward, a prognostic model was established by lasso regression analysis and multivariate cox analysis. The risk score was calculated based on the follow formula: risk score = Σ (Expi × coefi). Expi means expression of each gene, and coefi represents the risk coefficient. HCC samples were categorized by the median risk score into high and low risk groups. Kaplan-Meier curves (K-M curves) were generated using the "survival" and "survminer" R packages in order to investigate the differences in survival between distinct groups. Based on clinical characteristics and risk scores, the "rms" package in R is used to plot the nomogram to predict survival outcomes [24]. Calibrating plot was used to determine the accuracy of the nomogram. In order to verify the model, we divided the ICGC set into high- and low-risk groups, and performed Kaplan–Meier curve and receiver operating characteristic (ROC) curve of ICGC set.

Statistical analyses

Our statistical analyses were performed using R version 4.1.2. Differential clinical characteristics among distinct groups were analyzed by the Chi-squared test. Cox regression analysis (univariate and multivariate) was conducted to identify the independent prognostic factors. Comparison between the two groups was performed using Wilcox rank sum test. The significance level was set at $p \times 0.05$, and two-tailed p values were applied.

RESULTS

Landscape of genetic variation and transcriptional alterations of FAMs in HCC

49 FAMs obtained from the Genecards website were included in this study. Based on analysis of somatic mutation incidence, the TCGA set of 49 FAMs displayed a relatively high rate of somatic mutations. FAMs mutations were detected in 119 (32.69%) of the 364 HCC samples (Figure 1A). Among these, ALB was found with the highest mutation frequency (13%), followed by APOB.

Afterward, we examined somatic copy number alterations (CNVs) in these 49 FAMs and found widespread alterations in all 49 FAMs. Among them, FASN, ACOX1 and MTR showed increased CNVs, while FABP3, ACADVL, HADH, FAAH, and ACADS showed decreases in CNVs (Figure 1B). The CNVs in the FAMs on their respective chromosomes were

showed in Figure 1C. Moreover, a comparison of mRNA levels of FAMs was made between HCC tumor and normal tissues, and as showed in Figure 1D, most FAMs expression levels were positively correlated with CNV gain or loss and significantly different in tumor tissues. Consequently, while CNVs can be the primary cause of FAM expression changes, they are not the only factor that regulates mRNA expression [25]. Gene expression can also be affected by transcription factors and DNA methylation [26, 27]. We found HCC and normal samples have remarkably different genetic landscapes and mRNA expression levels of FAMs, indicating that FAMs may play an undiscovered role in HCC. Furthermore, Supplementary Figure 1A shows that the association between each FAM was highly correlated. Similarly, the infiltration levels of immune cells were assessed by ssGSEA algorithm and they showed high correlation in HCC (Supplementary Figure 1B). In summary, the above results indicated that FAMs are strongly correlated with HCC.

Identification of FAM cluster in HCC

Through a FAMs network (Figure 2A), the full scope of FAMs interactions and their prognostic value in HCC patients was displayed. Next, consensus clustering analysis was used to investigate interactions between FAMs and HCC. Using a consensus clustering algorithm, HCC patients were categorized into different clusters (Supplementary Figure 2). Using k = 3, we were able to sort the entire cohort into cluster A (n = 197), B (n = 72) and C (n = 102) (Figure 2B). A principal component analysis (PCA) of the FAMs transcription profiles highlighted significant differences among the three clusters (Figure 2C). The Kaplan-Meier curves for the three FAM clusters indicated that cluster C had the most prominent survival advantage, while cluster A had the worst ($p < 0.05$) (Figure 2D). Moreover, as shown in heatmap of clinicopathological features and expression of FAMs in HCC patients, cluster A displayed the lowest level of FAMs expression (Figure 2E).

GSVA enrichment analysis and immune infiltration estimation in distinct clusters

GSVA enrichment analysis was conducted among different clusters in order to identify potential biological pathways in HCC. The top 20 pathways in each cluster were visualized (Figure 3A–3C). Cluster C was significantly enriched in fatty acid metabolism pathway and immune-related pathways, such as PPAR signaling pathway, Toll-like, B cell receptor signaling pathway, Fc-gamma-R-mediated phagocytosis pathway and Nod-like receptor signaling pathway (Figure 3C). Furthermore, in order to assess whether FAMs

contribute to TME of HCC, we used the ssGSEA algorithm to calculate connection between the three clusters and 23 kinds of immune cells of every HCC sample. Among the three clusters, there were significant differences in the infiltration of immune cells (Figure 3D). Besides, HCC patients in cluster A had the highest expression level of most immune checkpoints among three FAM clusters (Figure 3E), that implied an exhausted immune TME in cluster A patients.

Identified of gene clusters based on FAM cluster-related DEGs in HCC

In the previous steps, three clusters were determined, then significant DEGs with adjusted *p* value < 0.001 were identified by differential analyses between any two clusters. The Venn diagram (Figure 4A) illustrated the following intersections which resulted in 190 DEGs. A functional enrichment analysis was employed to research the potential biological behavior of 190 DEGs. According to GO (gene ontology) and KEGG

(the Kyoto encyclopedia of genes and genomes) analysis, these FAM cluster-related genes were significantly enriched in metabolism pathways (Supplementary Figure 3).

Furthermore, 190 genes were screened for prognostic value by univariate cox regression analysis, and among them, 79 genes were found to be associated with overall survival (OS) in HCC (Supplementary Table 2). According to 79 prognostic genes, two genomic clusters named gene clusters D and E were identified by consensus clustering algorithm (Figure 4B and Supplementary Figure 4). According to Kaplan-Meier curves, HCC patients in gene cluster D had poorer overall survival compared to those in gene cluster E (Figure 4C). Afterward, HCC patients in FAM gene cluster D were related with higher FAM gene expression, advanced stage, advanced grade, and higher dead risk (Figure 4D). The result of further expression analysis was consistent with that in heatmap (Figure 4E).
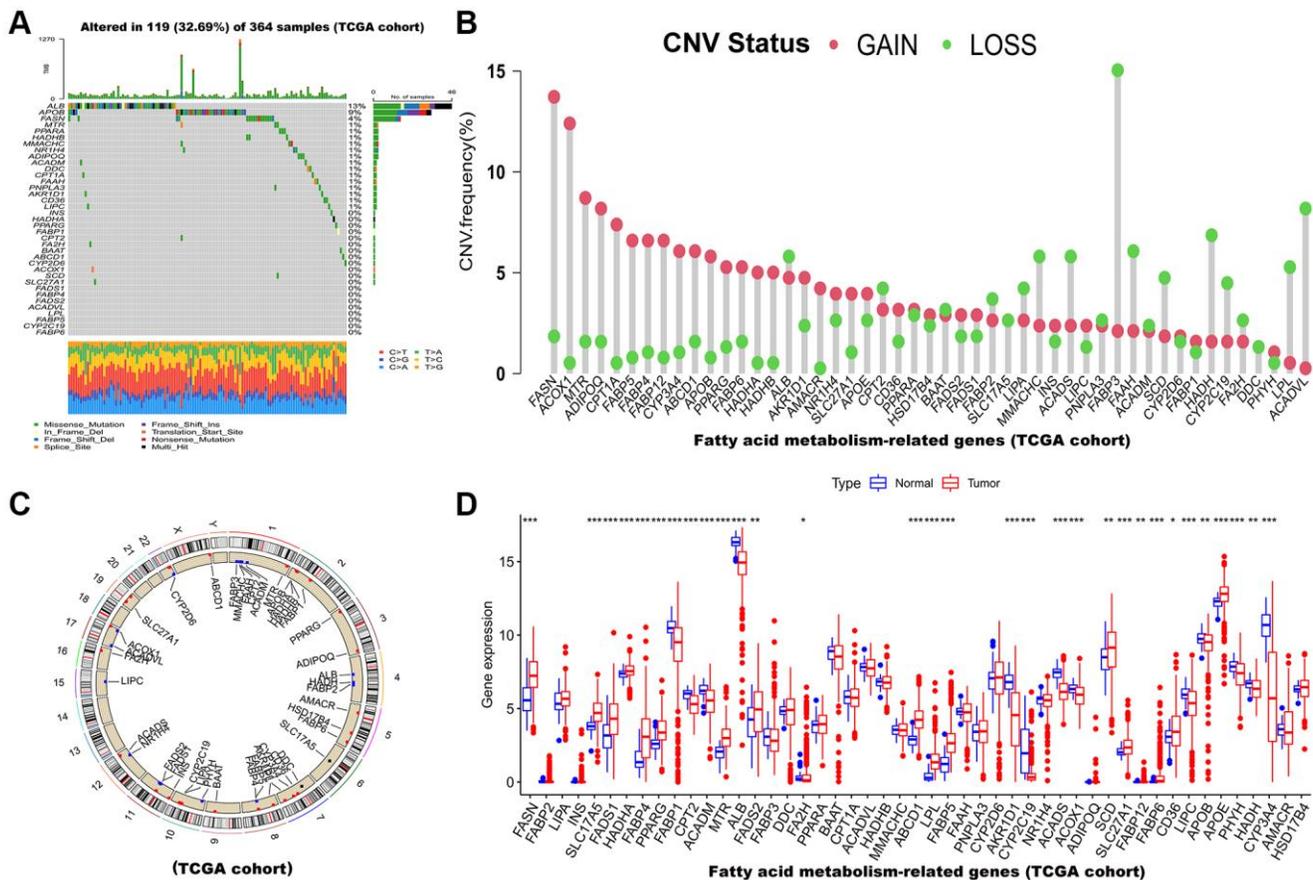
Figure 1. Multi-omics landscape of FAM-related genes in HCC based on TCGA cohort. (**A**) The mutation frequency of 49 FAMs in TCGA-LIHC cohort. Each column of the figure represents an individual patient. (**B**) The CNV variation frequency of FAMs (Red and green plots separately represent CNV gain and CNV loss). (**C**) Locations of CNV alterations in FAMs on 23 chromosomes. (**D**) The mRNA expression levels of 49 FAMs between HCC and normal tissues. Abbreviations: FAM: fatty acid metabolism; HCC: hepatocellular carcinoma; FAMs: fatty acid metabolism-related genes; TCGA: The Cancer Genome Atlas; LIHC: liver hepatocellular carcinoma; CNV: copy number variant. $^*p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

In addition, the immune analysis between two gene clusters revealed that gene cluster D tend to have higher infiltration level of most immune cells such as activated B cell, activated CD8+ T cells and activated CD4+ T cells (Figure 5A). Consistent with this, patients in gene cluster D also had higher expression level of immune
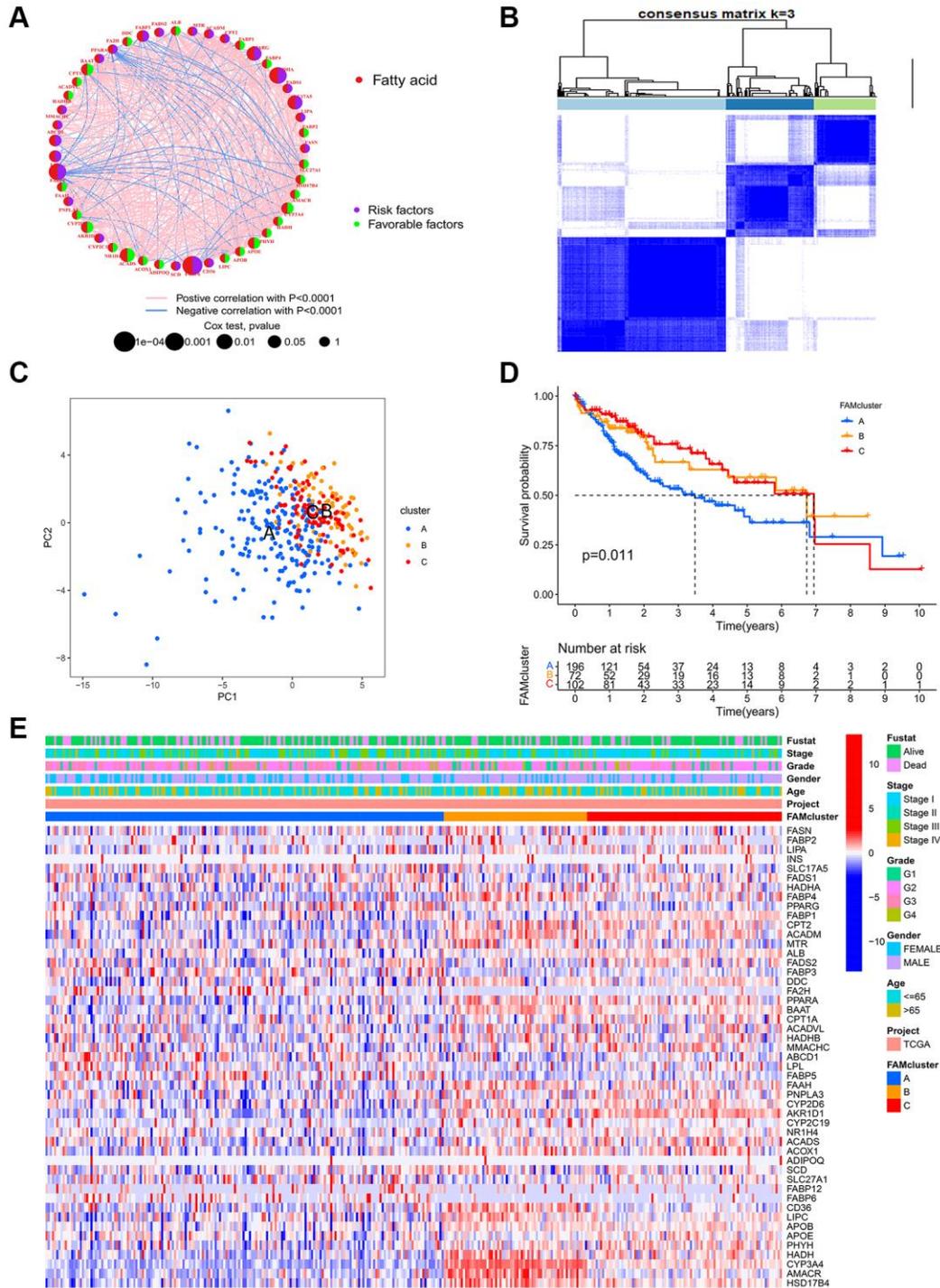
Figure 2. FAM clusters and relevant clinical features. (**A**) The interaction of expression on 49 FAMs in HCC. The line connecting the FAMs represents their interactions, with the line thickness indicating the strength of the association between FAMs. Red dots, fatty acid metabolism-related genes; Purple dots, risk factors for HCC; Green dots, favorable factors for HCC; Pink edges, positive correlation with $P < 0.0001$; Blue edges, negative correlation with $P < 0.0001$. (**B**) Consensus matrices of 49 FAMs in HCC for k = 3. (**C**) PCA analysis showing a remarkable difference in transcriptomes between the three FAM clusters in TCGA cohort. (**D**) K-M curve for the three FAM clusters. (**E**) The heatmap of clinical characteristics and expression levels of FAMs in different clusters. Abbreviations: FAM: fatty acid metabolism; FAMs: fatty acid metabolism-related genes; HCC: hepatocellular carcinoma; PCA: principal component analysis.

checkpoints (Figure 5B). Interestedly, we estimated the relative abundance of several important cancer-related signatures by ssGSEA algorithm in different gene clusters (Figure 5C). The results showed that HCC patients in gene cluster D had higher abundance levels of bad prognostic signatures, including EMT (epithelial-mesenchymal transition), poor survival, proliferation, vascular invasion, recurrent, metastasis signatures, and immune microenvironment signatures, such as innate immune response, pan-F-TBRS, co-inhibition antigen presenting cell (APC), co-stimulation APC, co-inhibition T cell, co-stimulation T cell, MHC-I HLA (major histocompatibility complex-I human leukocyte antigen), MHC-II HLA, antigen processing machinery, and immune checkpoint, compared to those in gene cluster E. Figure 5D showed that gene cluster D had higher expression levels of HLA genes.

Construction and verification of the prognostic risk model in HCC

By lasso regression analysis and multivariate cox analyses in 79 prognostic DEGs, we identified 5 genes including three risk factors (CCDC112, TRNP1, CFL1) and two protective factors (CYB5D2, SLC22A1) and created a prognostic model in HCC according to these five genes (Supplementary Figure 5 and Supplementary Table 3). The risk score of HCC patients was calculated as follows: risk score = $0.382912 \times$ TRNP1 + $0.65021 \times$ CCDC112 + $1.885657 \times$ CFL1 + $(-1.23099) \times$ CYB5D2 + $(-0.29032) \times$ SLC22A1. In TCGA-LIHC set, the median cut-off value was used to stratify the patients into two groups: high-risk score ($n = 182$) and low-risk score ($n = 183$). Figure 6A displayed the distribution of HCC patients across
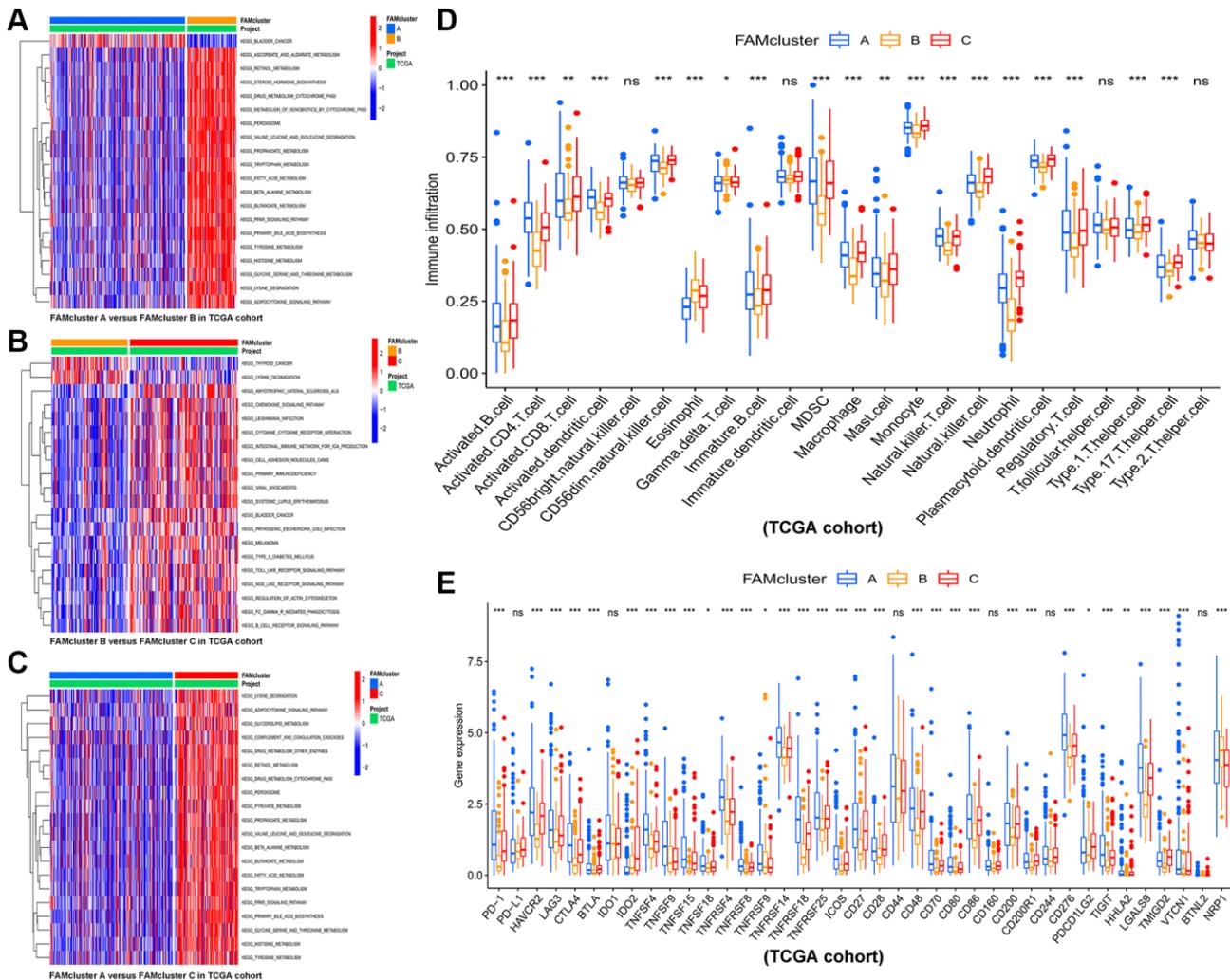

Figure 3. The results of GSVA and immune infiltration analysis in three clusters. GSVA results of biological pathways between: (**A**) cluster A vs. cluster B, (**B**) cluster B vs. cluster C, and (**C**) cluster A vs. cluster C, red and blue represent activated and inhibited pathways, respectively. (**D**) The infiltration levels of 23 immune cells in the three FAM clusters. (**E**) Significant differences in expression of immune checkpoint in the three FAM clusters. Abbreviations: GSVA: gene set variation analysis; FAM: fatty acid metabolism. $^*p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

three FAM clusters, two gene clusters, and two risk score groups. There was a significant risk score difference between FAM clusters and gene clusters. The risk score of cluster B was the lowest, while that of cluster A was the highest (Figure 6B). Cluster D had a higher risk score than cluster E (Figure 6C). In
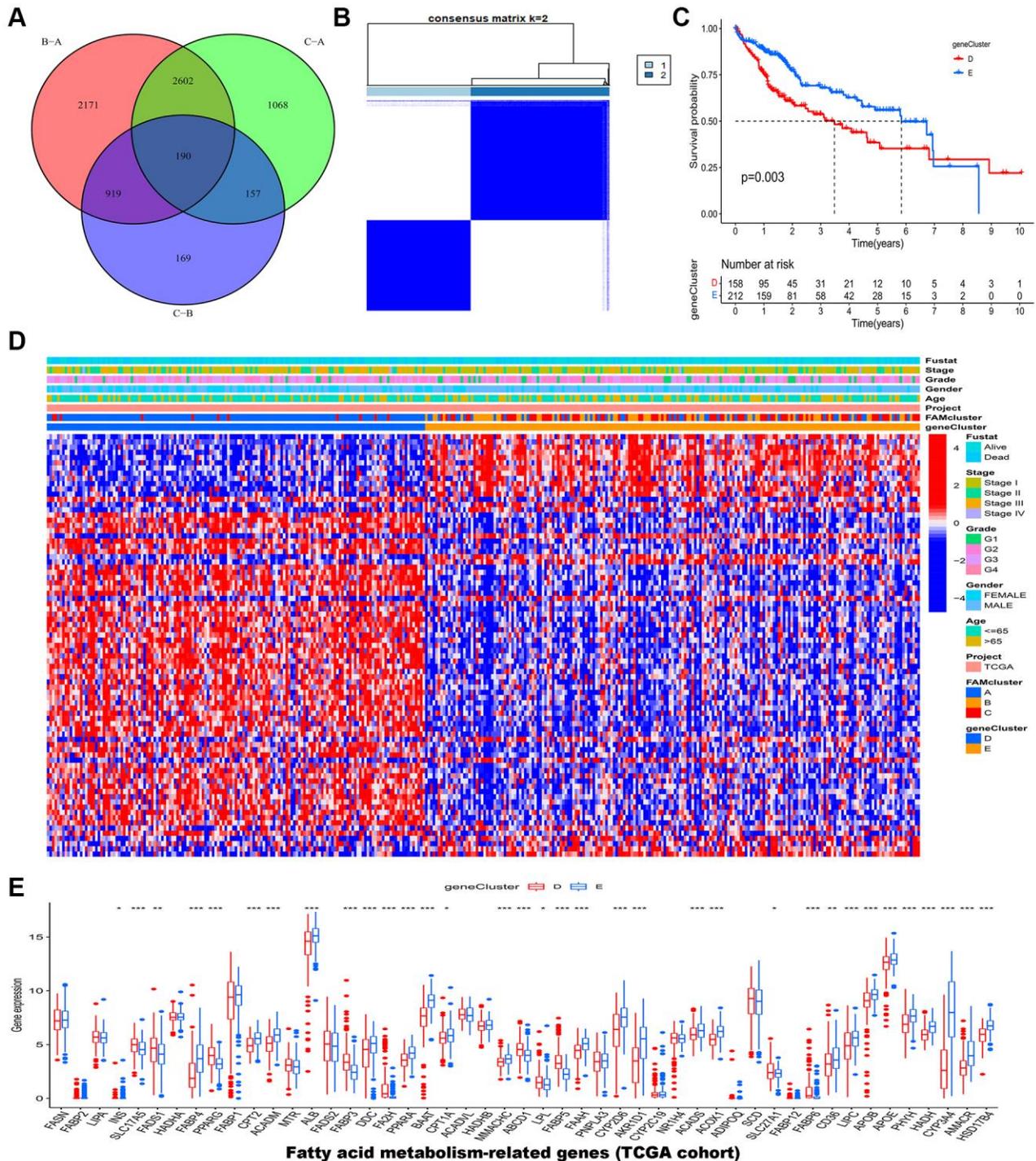
Figure 4. Identification of gene clusters based on DEGs in the TCGA-LIHC cohort. (**A**) Venn diagram showed the DEGs among the three FAM clusters. (**B**) HCC samples were divided into two clusters based on the consensus clustering (k = 2). (**C**) The OS analysis of HCC samples between gene cluster D and E. (**D**) The heatmap of clinical characteristics of HCC patients in different clusters. (**E**) The mRNA expression levels of 49 FAMs between gene cluster D and E. Abbreviations: DEGs: different expressed genes; TCGA: the Cancer Genome Atlas; LIHC: liver hepatocellular carcinoma; FAM: fatty acid metabolism; HCC: hepatocellular carcinoma; OS: overall survival; FAMs: fatty acid metabolism-related genes. $^*p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

TCGA-LIHC set, high-risk patients had a worse outcome than low-risk patients, and AUC (Area under curve) values of 0.708, 0.682, and 0.650 respectively represent 1-, 2-, and 3-year survival rates of risk scores (Figure 6D–6F).

As an external validation cohort, patients in ICGC-JP (ICGC-Japan) cohort were categorized, by the median risk score, into high- and low-risk groups. Consistently, in ICGC-JP cohort, high-risk patients had worse outcomes than low-risk patients, and the corresponding AUC values of 1-, 2-, and 3-year survival rates were 0.777, 0.718, 0.695, respectively, which indicated a good efficiency (Figure 6G–6I). Furthermore, cox regression analysis, both univariate (Figure 6J) and multivariate (Figure 6K), revealed the prognostic risk model is a reliable independent prognostic factor of HCC patients.

We have done more exploration of five genes on other databases, such as TCGA (Supplementary Figure 6A–6D), ICGC (Supplementary Figure 6E–6G) and GEO database (GSE25097, GSE112790, GSE102079, GSE45267, GSE39791 datasets) (Supplementary Figure 7). Moreover, we verified IHC on HPA database (Supplementary Figure 8A) and protein expression levels on CPTAC database (Supplementary Figure 8B). Interestingly, all results are consistent with our study, which TRNP1, CCDC112, CFL1 were risk factors (compared to normal tissues, there was a significant upregulation of TRNP1, CCDC112, CFL1 expression in HCC tissues. K-M curves showed that upregulated TRNP1, CCDC112, CFL1 were associated with poor OS) and CYB5D2, SLC22A1 were protective genes (expression of CYB5D2, SLC22A1 were decreased in tumor tissues, and higher expression of CYB5D2, SLC22A1 was associated with good OS).
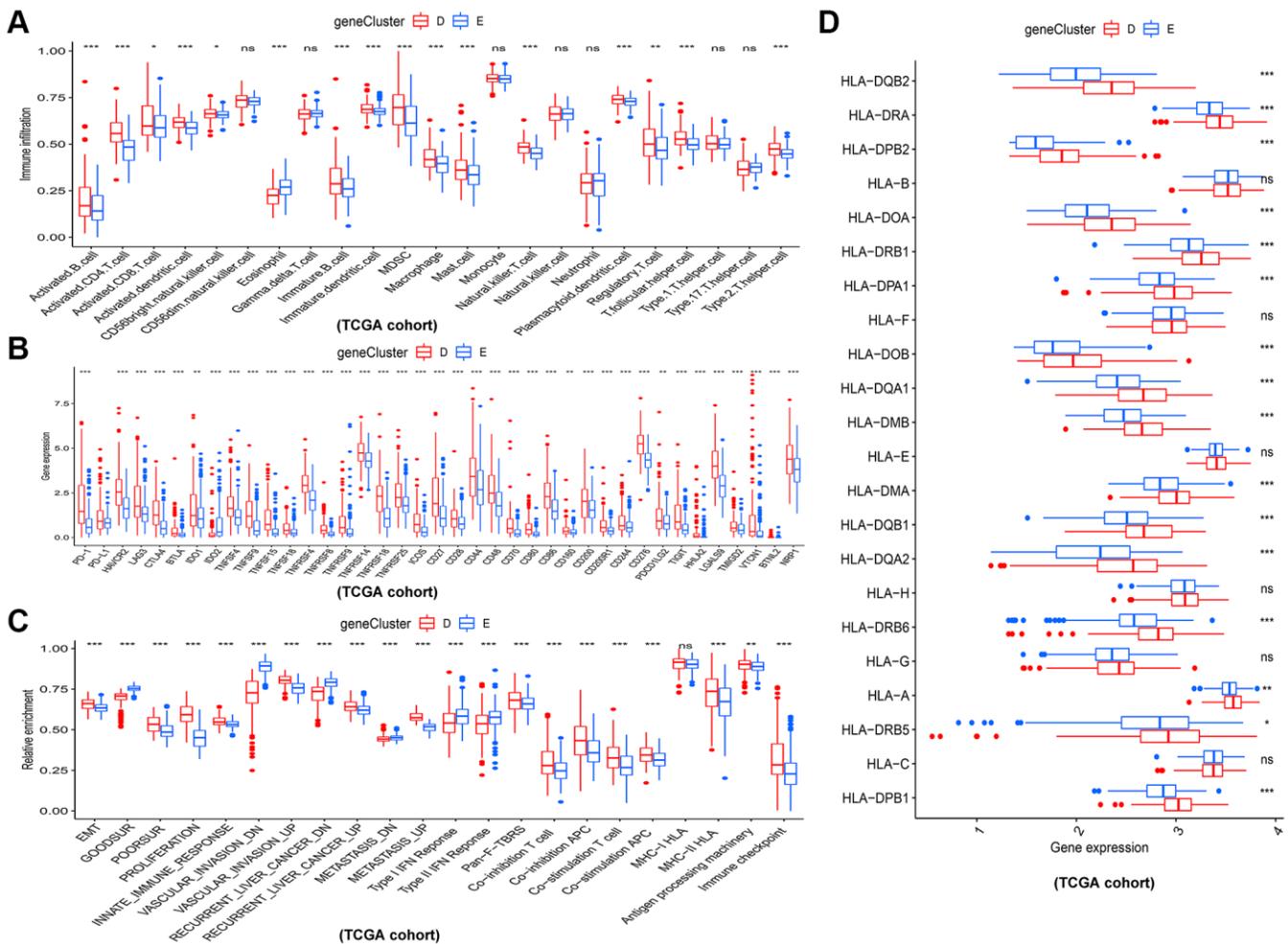

Figure 5. Different immune and cancer-related characteristics in gene cluster D and E. (**A**) The 23 kinds of immune cells in the two gene clusters. (**B**) Significant differences in expression of immune checkpoint between the two gene clusters. (**C**) The enrichment levels of cancer-related signatures in the two gene clusters. (**D**) Expression levels of HLA genes between gene cluster D and E. $^*p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.
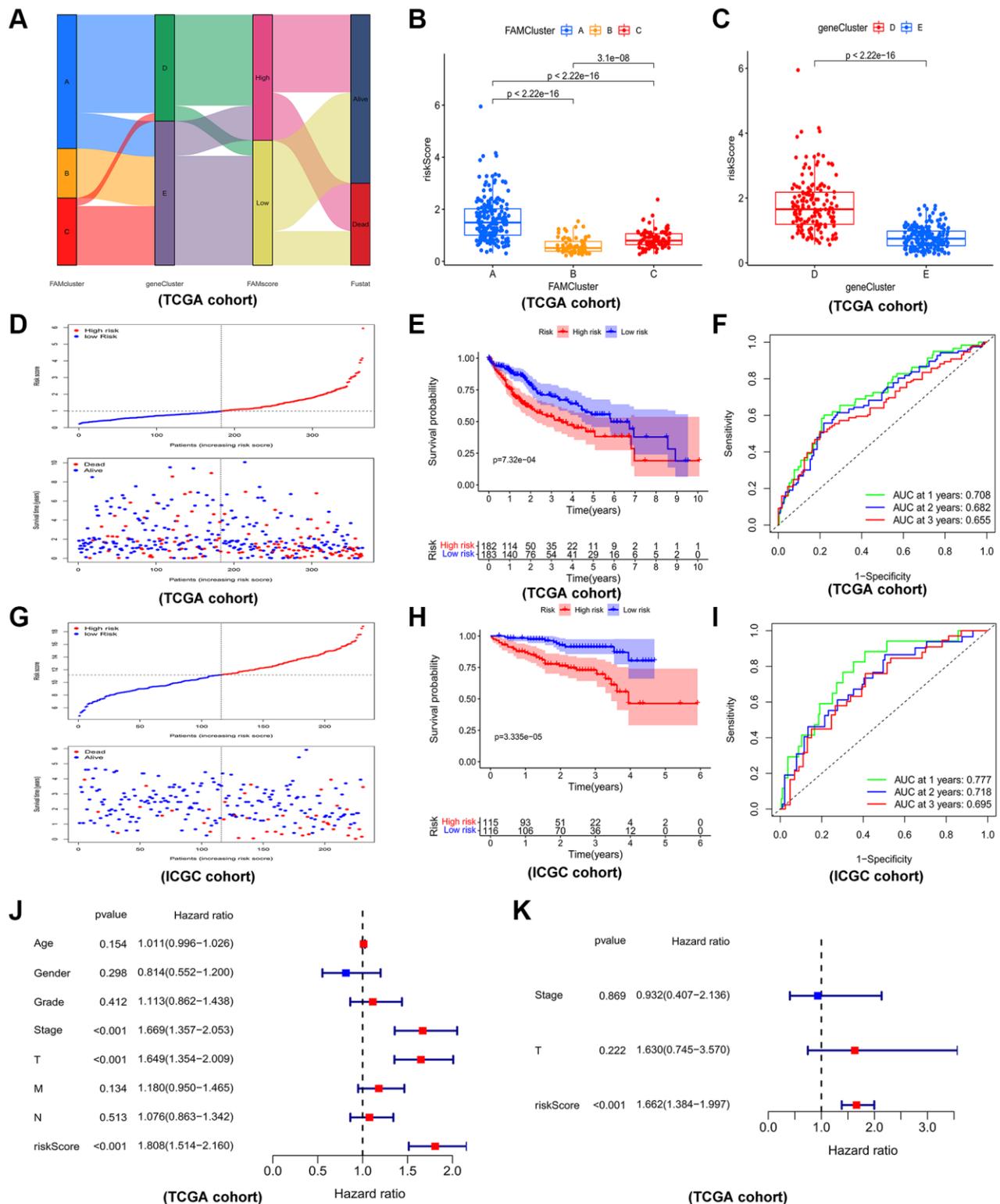
Figure 6. Construction and validation of prognostic risk model. (**A**) Alluvial diagram depicting the relationship of FAMcluster, genecluster, risk score (FAMscore) group and survival state. Boxplots of risk score in different FAMclusters (**B**) and geneclusters (**C**) Risk score distribution and scatter plots showing the risk score distribution and patient survival status in TCGA (**D**); Kaplan–Meier analysis of OS between the two groups in TCGA (**E**); ROC curves to predict the sensitivity and specificity of 1-, 3-, 5-year survival according the risk score in TCGA (**F**). Risk score distribution and scatter plots (**G**), Kaplan–Meier curves (**H**), ROC curves (**I**) of the risk model in ICGC cohort. The univariate (**J**) and multivariate (**K**) independent prognostic analysis of the model in TCGA cohort. Abbreviations: TCGA: the cancer genome atlas database; ICGC: International Cancer Genome Consortium; OS: overall survival; ROC: receiver operating characteristic.

Relationship of TME and the prognostic risk model in HCC

In the TCGA-LIHC cohort, we assessed the abundance of immune cells and cancer-related signatures by using the ssGSEA algorithm. Through the spearman method, the association among risk score and immune cells, cancer-related signatures levels were evaluated. As shown in the boxplots, the levels of immune cells (Supplementary Figure 9) and immune checkpoints (Figure 7A) in high-risk patients were higher than low-risk patients. Moreover, Figure 7B showed that high-risk patients also had higher abundance levels of bad prognostic signatures, such as EMT, poor survival, proliferation, vascular invasion, recurrent, metastasis signatures, and immune microenvironment signatures, such as innate immune response, pan-F-TBRS, co-inhibition APC, co-inhibition T cell, co-stimulation APC, co-stimulation T cell, MHC-I HLA, MHC-II HL, antigen processing machinery, and immune checkpoint compared to low-risk patients. Also, we conducted gene set enrichment analysis (GSEA) of HCC patients in different risk groups, and the result showed that high-risk group was enriched in Fc gamma R mediated phagocytosis, T cell receptor signaling pathway, Nod-like receptor signaling pathway, Fc epsilon Ri signaling pathway, while low risk group was enriched in PPAR signaling pathway and drug metabolism pathway (Figure 7C).

Figure 7. Connection among prognostic risk model and immune or cancer-related characteristics of HCC patients. (**A**) Significant differences in expression of immune checkpoints between the two groups. (**B**) The enrichment level of cancer-related signatures in the two groups. (**C**) Immune-related pathways enriched in the high-risk group. The correlation between genes in prognostic risk model and the infiltration level of 23 immune cells (**D**) and immune-related pathway (**E**). Red for positive associations and green for negative associations. Abbreviation: HCC: hepatocellular carcinoma. $^*p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Furthermore, the relationship between five genes in the model and immune cells was analyzed (Figure 7D). We observed that three high-risk genes (CCDC112, TRNP1 and CFL1) were significantly positively correlated with most immune cells, whereas significant negative correlation was observed between two low-risk genes (CYB5D2 and SLC22A1) and infiltration of immune cells. Consistently, Figure 7E displayed the result of correlation between five genes, risk score and immune related pathways.

Construction of nomogram in HCC

Comparison of genes mutations between the two risk groups revealed that high-risk patients had significantly higher mutation rates of TP53, MUC4, FLG, CSMD3, ARID1A, FAT3 than low-risk patients (Figure 8A, 8B). Moreover, high-risk HCC patients were remarkably associated with worse outcome, more advanced tumor stage and worse pathological grade (Figure 8C, 8D). To identify the reliability of this risk model in HCC patients, the prognostic nomogram plot containing the risk score and stage was constructed in TCGA-LIHC cohort (Figure 8E). Furthermore, calibration plot indicated excellent agreement between prediction and actual risk (Figure 8F). Overall, the risk model showed good prognostic value in HCC samples.

DISCUSSION

HCC seriously threatens human health with high mortality rate. While HCC can be managed with multiple treatments, patients with the disease have extremely low 5-year survival rates due to the fact that it is commonly diagnosed in advanced stages [28].

Currently, a number of immune checkpoint inhibitors (ICIs) have been approved by the FDA (Food and Drug Administration) to treat advanced HCC, including nivolumab [29] and pembrolizumab [30]. However, there are numerous disadvantages of ICI treatment, including low response rates and side effects. Therefore, new therapeutic targets and novel prognostic models are essential for HCC patients.

Metabolism reprogramming is critical for tumor initiation and progression, especially during HCC development [31]. Synthesis of fatty acids has been involved in energy metabolism and membrane production of tumor cells. Deregulated fatty acid metabolism has been regarded as a vital metabolic regulator in supporting cancer cell proliferation [32]. A remolded microenvironment caused by abnormally fatty acid metabolism could promote HCC progression. In this study, our objective was to assess the association of FAMs and the risk of HCC.

First, we explored the mutation and correlation state of 49 FAMs obtained from the Genecard database. The top three frequently mutated genes were ALB, APOB, and FASN. Missense mutation and C>T of FAMs were the most common mutations in HCC. Due to the high expression of ALB (20%) [33] and APOB's ability to facilitate VLDL secretion [34] (which consumes large amounts of energy), mutation of ALB or APOB may be inactivated to divert energy into cancer-relevant metabolic pathways [35]. According to the expression profiles of 49 FAMs, we determined 3 FAM clusters. Among 3 FAM clusters, cluster C had highest level of immune infiltration. Subsequently, differential analyses among 3 FAM clusters were employed. We screened 190 DEGs and showed them in a Venn plot. Based on the expression of 79 prognostic genes identified from 190 DEGs, HCC patients were grouped into 2 different gene clusters. Gene cluster D had worse survival rate, higher expression level of FAMs, and higher infiltration level of immune cells. Interestingly, gene cluster D also had higher enrichment of poor prognostic signatures, such as poor survival, liver cancer recurrent related signatures, cancer progression related signatures such as EMT, proliferation, vascular invasion, metastasis signatures, and several immune signatures, such as innate immune response, pan-F-TBRS, co-inhibition APC, co-inhibition T cell, co-stimulation APC, co-stimulation T cell, MHC-I HLA, MHC-II HLA, antigen processing machinery and immune checkpoint related signatures. These results indicated that FAMs appear to affect TME of HCC.

Moreover, based on 79 prognostic FAMs, a FAM-related model containing 5 genes (TRNP1, CCDC112, CFL1, CYB5D2, SLC22A1) was constructed by LASSO and multivariate Cox regression analysis in TCGA-LICH cohort. And we successfully confirmed the model using ICGC-JP cohort. HCC patients were categorized into two groups, high risk and low risk group. In both the TCGA and ICGC cohorts, the K-M curves showed that patients in the low group had better outcomes than those in the high group. The 1 year AUC of the model was 0.708, 0.777 in TCGA and ICGC cohort, respectively, which demonstrated that the accuracy of the risk model was excellent. The relation of our model and immune infiltration was also assessed. And the infiltration levels of immune cells were evaluated by ssGSEA. The analysis of relationship revealed that risk score was significantly positively correlated with infiltration of immune cell in HCC patients, especially CD56 bright natural killer cell, activated CD4 T cell and activated dendritic cell. The results of correlation analysis between each gene in model and HCC immunity were consistent with the properties of genes. For example, TRNP1, CCDC112, CFL1 are risk factors, then they were positively

correlated with the infiltration levels of most of immune cells, whereas the results of CYB5D2, SLC22A1 were contrary to this. Patients in high risk group had higher enrichment level of poor prognostic signatures, such as poor survival, liver cancer recurrent related signatures,

cancer progression related signatures such as EMT, proliferation, vascular invasion, metastasis signatures, and several immune signatures, such as innate immune response, pan-F-TBRS, co-inhibition APC, co-inhibition T cell, co-stimulation APC, co-stimulation T cell,
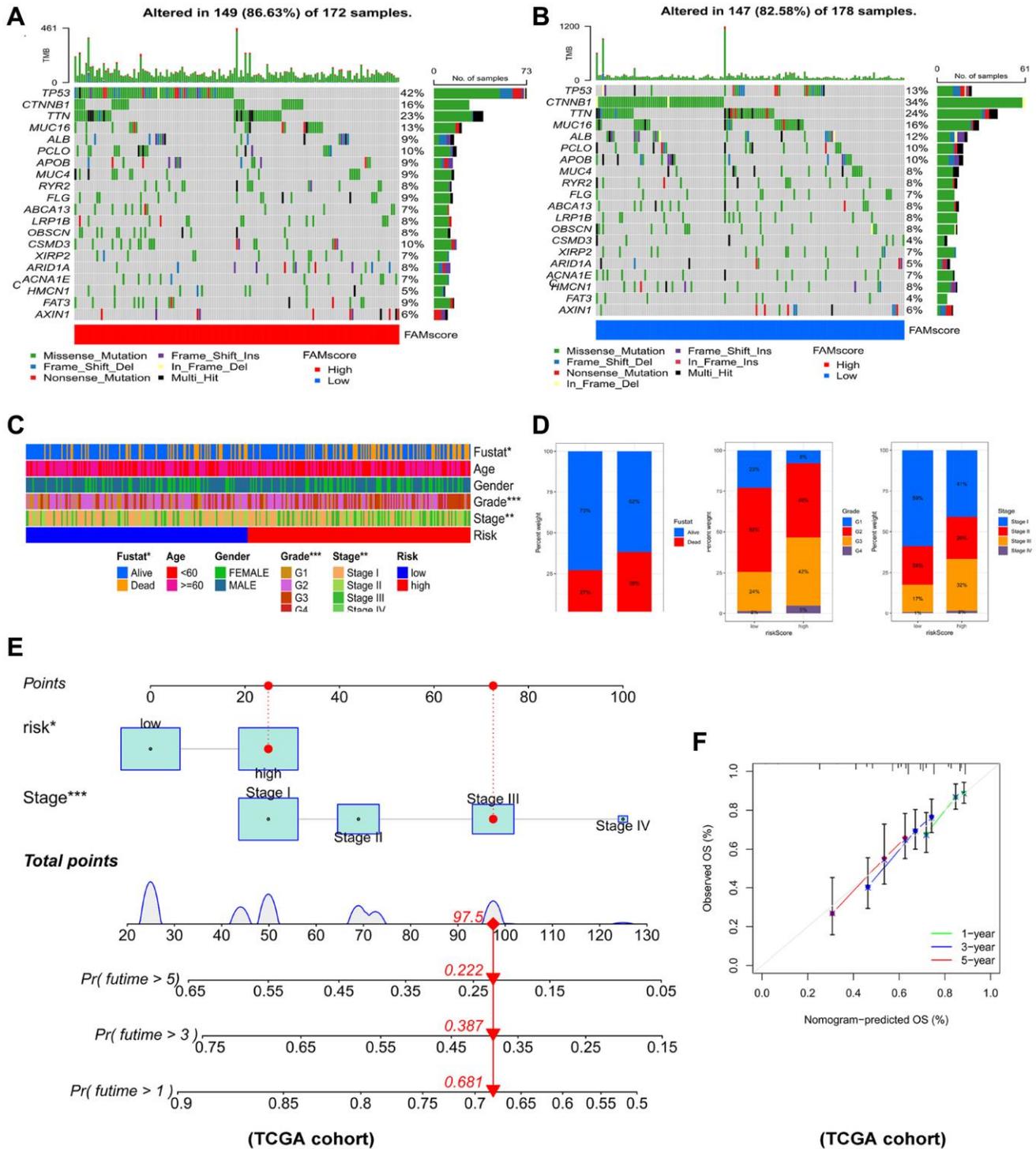
Figure 8. Connections between prognostic risk model and clinical characteristics of HCC patients. The waterfall plot of tumor somatic mutation established in (**A**) high risk group and (**B**) low risk group. (**C**) The heatmap of the model and clinical characteristics in TCGA-LIHC cohort. (**D**) Stacked bar plot of HCC survival state, pathological grade and tumor stage. (**E**) Nomogram for predicting the 1-,3- and 5- year OS of HCC patients. (**F**) Calibration curve of the program for predicting of 1-,3- and 5-year OS of HCC patients.

MHC-I HLA, MHC-II HLA, antigen processing machinery and immune checkpoint related signatures.

TRNA1, CCDC112, and CFL1 expression were substantially higher in HCC tissues than in normal tissues, whereas CYB5D2 and SLC22A1 expression were significantly lower. These results were found in the TCGA, ICGC, GEO, HPA, and CPTAC databases. Liu et al. reported TRNP1 as a risk factor of four-gene model for predicting OS in HCC patients [36]. TRNP1 is essential for neural development and cell self-renewal [37]. As a hypoxia-responsive gene, CFL1 contributes to hypoxia-induced HCC progression by activating PLD1/AKT signals [38]. In a mechanism study, knockdown of CFL1 increased F-actin levels and disrupted the balance between F-actin and G-actin, which resulted in aggressiveness inhibition of HCC cells [39]. Researches have reported that decreased level of CYB5D2 is associated with breast cancer progression [40]. SLC22A1 downregulation correlates with worse patient outcomes and tumor progression [41]. It is thought that the development of HCC is accompanied by aberrant SLC22A1 variants, which may greatly affect the sorafenib levels in the affected intracellular concentrations in HCC [42]. However, there is still a lack of knowledge about how TRNP1 and CCDC112 affect the development and prognosis of HCC.

In recognition of the clinical utility of the model in predicting over survival in HCC patients, using the risk score and stage together, a nomogram was created to predict the 1-, 3-, and 5-year survival rates for HCC in TCGA cohort. The calibration plot verified the accuracy of nomograms.

Nevertheless, our study has several shortcomings. First, molecular mechanisms of these genes need to be uncovered by additional functional experiments. Second, an additional experiment is needed for further verification of model genes. Finally, since the study was analyzed on data from public database, the risk model needs to be validated by our own clinical cohort.

Taken together, we identified 3 FAM clusters, 2 gene clusters and established a novel 5-gene prognostic model for HCC patients. Fatty acid metabolism-related genes exhibited synergy with immune activation. We hope the prognostic model may help improve immunotherapy for HCC in the future.

CONCLUSION

Our findings investigated molecular cluster and prognostic model about fatty acid metabolism in HCC, and highlighted a potential strategy for targeting the immunometabolism of HCC.

Abbreviations

HCC: Hepatocellular Carcinoma; FAM: fatty acid metabolism; TCGA: The cancer genome atlas database; ICGC: International Cancer Genome Consortium; DEGs: differentially expressed genes; LASSO: Least absolute shrinkage and selection operator; TKI: Tyrosine kinase inhibitors; TME: tumor micro-environment; FAMs: fatty acid metabolism-related genes; LIHC: liver hepatocellular carcinoma; FPKM: fragments per kilobase million; TPM: transcripts per kilobase million; MSigDB: molecular signatures database; GSVA: gene set variation analysis; ssGSEA: single-sample gene set enrichment analysis; K-M curve: Kaplan-Meier curve; ROC curve: receiver operating characteristic curve; CNVs: copy number alterations; ssGSEA: single-sample gene-set enrichment analysis; PCA: principal component analysis; GO: Gene ontology; KEGG: The Kyoto Encyclopedia of Genes and Genomes; OS: overall survival; EMT: epithelial-mesenchymal transition; APC: antigen presenting cell; MHC: major histocompatibility complex; HLA: human leukocyte antigen; AUC: Area under curve; ICGC-JP: ICGC-Japan; ICIs: immune checkpoint inhibitors; FDA: Food and Drug Administration.

AUTHOR CONTRIBUTIONS

XL and AL designed and monitored the research. ZR initially screened and analyzed the data. DG drafted the manuscript. YL, ZS, GW performed statistical analyses. NQ revised the manuscript. All authors approved the submitted version.

ACKNOWLEDGMENTS

CONFLICTS OF INTEREST

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

FUNDING

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.
https://doi.org/10.1038/s41572-020-00240-3
PMID:33479224

2. Vogel A, Martinelli E, and ESMO Guidelines Committee. Updated treatment recommendations for hepatocellular carcinoma (HCC) from the ESMO Clinical Practice Guidelines. Ann Oncol. 2021; 32:801–5.
https://doi.org/10.1016/j.annonc.2021.02.014
PMID:33716105

3. 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

4. Zhao Y, Zhang YN, Wang KT, Chen L. Lenvatinib for hepatocellular carcinoma: From preclinical mechanisms to anti-cancer therapy. Biochim Biophys Acta Rev Cancer. 2020; 1874:188391.
https://doi.org/10.1016/j.bbcan.2020.188391
PMID:32659252

5. 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

6. Hall Z, Chiarugi D, Charidemou E, Leslie J, Scott E, Pellegrinet L, Allison M, Mocciaro G, Anstee QM, Evan GI, Hoare M, Vidal-Puig A, Oakley F, et al. Lipid Remodeling in Hepatocyte Proliferation and Hepatocellular Carcinoma. Hepatology. 2021; 73:1028–44.
https://doi.org/10.1002/hep.31391
PMID:32460431

7. Currie E, Schulze A, Zechner R, Walther TC, Farese RV Jr. Cellular fatty acid metabolism and cancer. Cell Metab. 2013; 18:153–61.
https://doi.org/10.1016/j.cmet.2013.05.017
PMID:23791484

8. Hoy AJ, Nagarajan SR, Butler LM. Tumour fatty acid metabolism in the context of therapy resistance and obesity. Nat Rev Cancer. 2021; 21:753–66.
https://doi.org/10.1038/s41568-021-00388-4
PMID:34417571

9. Nagarajan SR, Butler LM, Hoy AJ. The diversity and breadth of cancer cell fatty acid metabolism. Cancer Metab. 2021; 9:2.
https://doi.org/10.1186/s40170-020-00237-2
PMID:33413672

10. Ringel AE, Drijvers JM, Baker GJ, Catozzi A, García-Cañaveras JC, Gassaway BM, Miller BC, Juneja VR, Nguyen TH, Joshi S, Yao CH, Yoon H, Sage PT, et al. Obesity Shapes Metabolism in the Tumor Microenvironment to Suppress Anti-Tumor Immunity. Cell. 2020; 183:1848–66.e26.
https://doi.org/10.1016/j.cell.2020.11.009
PMID:33301708

11. Marien E, Meister M, Muley T, Gomez Del Pulgar T, Derua R, Spraggins JM, Van de Plas R, Vanderhoydonc F, Machiels J, Binda MM, Dehairs J, Willette-Brown J, Hu Y, et al. Phospholipid profiling identifies acyl chain elongation as a ubiquitous trait and potential target for the treatment of lung squamous cell carcinoma. Oncotarget. 2016; 7:12582–97.
https://doi.org/10.18632/oncotarget.7179
PMID:26862848

12. Madak-Erdogan Z, Band S, Zhao YC, Smith BP, Kulkoyluoglu-Cotul E, Zuo Q, Santaliz Casiano A, Wrobel K, Rossi G, Smith RL, Kim SH, Katzenellenbogen JA, Johnson ML, et al. Free Fatty Acids Rewire Cancer Metabolism in Obesity-Associated Breast Cancer via Estrogen Receptor and mTOR Signaling. Cancer Res. 2019; 79:2494–510.
https://doi.org/10.1158/0008-5472.CAN-18-2849
PMID:30862719

13. Wang YN, Zeng ZL, Lu J, Wang Y, Liu ZX, He MM, Zhao Q, Wang ZX, Li T, Lu YX, Wu QN, Yu K, Wang F, et al. CPT1A-mediated fatty acid oxidation promotes colorectal cancer cell metastasis by inhibiting anoikis. Oncogene. 2018; 37:6025–40.
https://doi.org/10.1038/s41388-018-0384-z
PMID:29995871

14. Wang L, Li C, Song Y, Yan Z. Inhibition of carnitine palmitoyl transferase 1A-induced fatty acid oxidation suppresses cell progression in gastric cancer. Arch Biochem Biophys. 2020; 696:108664.
https://doi.org/10.1016/j.abb.2020.108664
PMID:33157102

15. Han S, Wei R, Zhang X, Jiang N, Fan M, Huang JH, Xie B, Zhang L, Miao W, Butler AC, Coleman MA, Vaughan AT, Wang Y, et al. CPT1A/2-Mediated FAO Enhancement-A Metabolic Target in Radioresistant Breast Cancer. Front Oncol. 2019; 9:1201.
https://doi.org/10.3389/fonc.2019.01201
PMID:31803610

16. Wu Y, Fabritius M, Ip C. Chemotherapeutic sensitization by endoplasmic reticulum stress: increasing the efficacy of taxane against prostate cancer. Cancer Biol Ther. 2009; 8:146–52.
https://doi.org/10.4161/cbt.8.2.7087
PMID:19182512

17. Veglia F, Tyurin VA, Blasi M, De Leo A, Kossenkov AV, Donthireddy L, To TKJ, Schug Z, Basu S, Wang F, Ricciotti E, DiRusso C, Murphy ME, et al. Fatty acid transport protein 2 reprograms neutrophils in cancer. Nature. 2019; 569:73–8.
https://doi.org/10.1038/s41586-019-1118-2
PMID:30996346

18. Fujimoto A, Furuta M, Totoki Y, Tsunoda T, Kato M, Shiraishi Y, Tanaka H, Taniguchi H, Kawakami Y, Ueno M, Gotoh K, Ariizumi S, Wardell CP, et al. Whole-genome mutational landscape and characterization of noncoding and structural mutations in liver cancer. Nat Genet. 2016; 48:500–9.
https://doi.org/10.1038/ng.3547
PMID:27064257

19. Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, Szcześniak MW, Gaffney DJ, Elo LL, Zhang X, Mortazavi A. Erratum to: A survey of best practices for RNA-seq data analysis. Genome Biol. 2016; 17:181.
https://doi.org/10.1186/s13059-016-1047-4
PMID:27565134

20. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res. 2018; 28:1747–56.
https://doi.org/10.1101/gr.239244.118
PMID:30341162

21. 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

22. Long Q, Huang C, Meng Q, Peng J, Yao F, Du D, Wang X, Zhu W, Shi D, Xu X, Qi X, Deng W, Chen M, et al. TNF Patterns and Tumor Microenvironment Characterization in Head and Neck Squamous Cell Carcinoma. Front Immunol. 2021; 12:754818.
https://doi.org/10.3389/fimmu.2021.754818
PMID:34691075

23. 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

24. Iasonos A, Schrag D, Raj GV, Panageas KS. How to build and interpret a nomogram for cancer prognosis. J Clin Oncol. 2008; 26:1364–70.
https://doi.org/10.1200/JCO.2007.12.9791
PMID:18323559

25. Sebestyén E, Singh B, Miñana B, Pagès A, Mateo F, Pujana MA, Valcárcel J, Eyras E. Large-scale analysis of genome and transcriptome alterations in multiple tumors unveils novel cancer-relevant splicing networks. Genome Res. 2016; 26:732–44.
https://doi.org/10.1101/gr.199935.115
PMID:27197215

26. Koch A, Joosten SC, Feng Z, de Ruijter TC, Draht MX, Melotte V, Smits KM, Veeck J, Herman JG, Van Neste L, Van Criekinge W, De Meyer T, van Engeland M. Analysis of DNA methylation in cancer: location revisited. Nat Rev Clin Oncol. 2018; 15:459–66.
https://doi.org/10.1038/s41571-018-0004-4
PMID:29666440

27. Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, Chen X, Taipale J, Hughes TR, Weirauch MT. The Human Transcription Factors. Cell. 2018; 172:650–65.
https://doi.org/10.1016/j.cell.2018.01.029
PMID:29425488

28. Zheng J, Kuk D, Gönen M, Balachandran VP, Kingham TP, Allen PJ, D'Angelica MI, Jarnagin WR, DeMatteo RP. Actual 10-Year Survivors After Resection of Hepatocellular Carcinoma. Ann Surg Oncol. 2017; 24:1358–66.
https://doi.org/10.1245/s10434-016-5713-2
PMID:27921192

29. El-Khoueiry AB, Sangro B, Yau T, Crocenzi TS, Kudo M, Hsu C, Kim TY, Choo SP, Trojan J, Welling TH Rd, Meyer T, Kang YK, Yeo W, 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; 389:2492–502.
https://doi.org/10.1016/S0140-6736(17)31046-2
PMID:28434648

30. Verset G, Borbath I, Karwal M, Verslype C, Van Vlierberghe H, Kardosh A, Zagonel V, Stal P, Sarker D, Palmer DH, Vogel A, Edeline J, Cattan S, et al. Pembrolizumab Monotherapy for Previously Untreated Advanced Hepatocellular Carcinoma: Data from the Open-Label, Phase II KEYNOTE-224 Trial. Clin Cancer Res. 2022; 28:2547–54.
https://doi.org/10.1158/1078-0432.CCR-21-3807
PMID:35421228

31. Rui L. Energy metabolism in the liver. Compr Physiol. 2014; 4:177–97.
https://doi.org/10.1002/cphy.c130024
PMID:24692138

32. Li Z, Zhang H. Reprogramming of glucose, fatty acid and amino acid metabolism for cancer progression. Cell Mol Life Sci. 2016; 73:377–92.
https://doi.org/10.1007/s00018-015-2070-4
PMID:26499846

33. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Å, Kampf C, Sjöstedt E, Asplund A, Olsson I, Edlund K, Lundberg E, et al. Proteomics. Tissue-based map of the human proteome. Science. 2015; 347:1260419.
https://doi.org/10.1126/science.1260419
PMID:25613900

34. Egusa G, Beltz WF, Grundy SM, Howard BV. Influence of obesity on the metabolism of apolipoprotein B in humans. J Clin Invest. 1985; 76:596–603.
https://doi.org/10.1172/JCI112011
PMID:4031064

35. Fernandez-Banet J, Lee NP, Chan KT, Gao H, Liu X, Sung WK, Tan W, Fan ST, Poon RT, Li S, Ching K, Rejto PA, Mao M, Kan Z. Decoding complex patterns of genomic rearrangement in hepatocellular carcinoma. Genomics. 2014; 103:189–203.
https://doi.org/10.1016/j.ygeno.2014.01.003
PMID:24462510

36. Liu J, Zhang SQ, Chen J, Li ZB, Chen JX, Lu QQ, Han YS, Dai W, Xie C, Li JC. Identifying Prognostic Significance of RCL1 and Four-Gene Signature as Novel Potential Biomarkers in HCC Patients. J Oncol. 2021; 2021:5574150.
https://doi.org/10.1155/2021/5574150
PMID:34257652

37. Esgleas M, Falk S, Forné I, Thiry M, Najas S, Zhang S, Mas-Sanchez A, Geerlof A, Niessing D, Wang Z, Imhof A, Götz M. Trnp1 organizes diverse nuclear membrane-less compartments in neural stem cells. EMBO J. 2020; 39:e103373.
https://doi.org/10.15252/embj.2019103373
PMID:32627867

38. Yao B, Li Y, Chen T, Niu Y, Wang Y, Yang Y, Wei X, Liu Q, Tu K. Hypoxia-induced cofilin 1 promotes hepatocellular carcinoma progression by regulating the PLD1/AKT pathway. Clin Transl Med. 2021; 11:e366.
https://doi.org/10.1002/ctm2.366
PMID:33784016

39. Zhang L, Chai Z, Kong S, Feng J, Wu M, Tan J, Yuan M, Chen G, Li Z, Zhou H, Cheng S, Xu H. Nujiangexanthone A Inhibits Hepatocellular Carcinoma Metastasis via Down Regulation of Cofilin 1. Front Cell Dev Biol. 2021; 9:644716.
https://doi.org/10.3389/fcell.2021.644716
PMID:33791303

40. Ojo D, Rodriguez D, Wei F, Bane A, Tang D. Downregulation of CYB5D2 is associated with breast cancer progression. Sci Rep. 2019; 9:6624.
https://doi.org/10.1038/s41598-019-43006-y
PMID:31036830

41. Heise M, Lautem A, Knapstein J, Schattenberg JM, Hoppe-Lotichius M, Foltys D, Weiler N, Zimmermann A, Schad A, Gründemann D, Otto G, Galle PR, Schuchmann M, Zimmermann T. Downregulation of organic cation transporters OCT1 (SLC22A1) and OCT3 (SLC22A3) in human hepatocellular carcinoma and their prognostic significance. BMC Cancer. 2012; 12:109.
https://doi.org/10.1186/1471-2407-12-109
PMID:22439694

42. Herraez E, Lozano E, Macias RI, Vaquero J, Bujanda L, Banales JM, Marin JJ, Briz O. Expression of SLC22A1 variants may affect the response of hepatocellular carcinoma and cholangiocarcinoma to sorafenib. Hepatology. 2013; 58:1065–73.
https://doi.org/10.1002/hep.26425
PMID:23532667

A

B

Supplementary Figure 1. The results of correlation analysis in HCC. (**A**) Correlation analysis among 49 FAMs in HCC patients. Red represents positive correlation; blue represents negative correlation. (**B**) The correlations of immune cells in HCC patients. Red dots denoting activated pathways and blue dots denoting inhibited pathways.
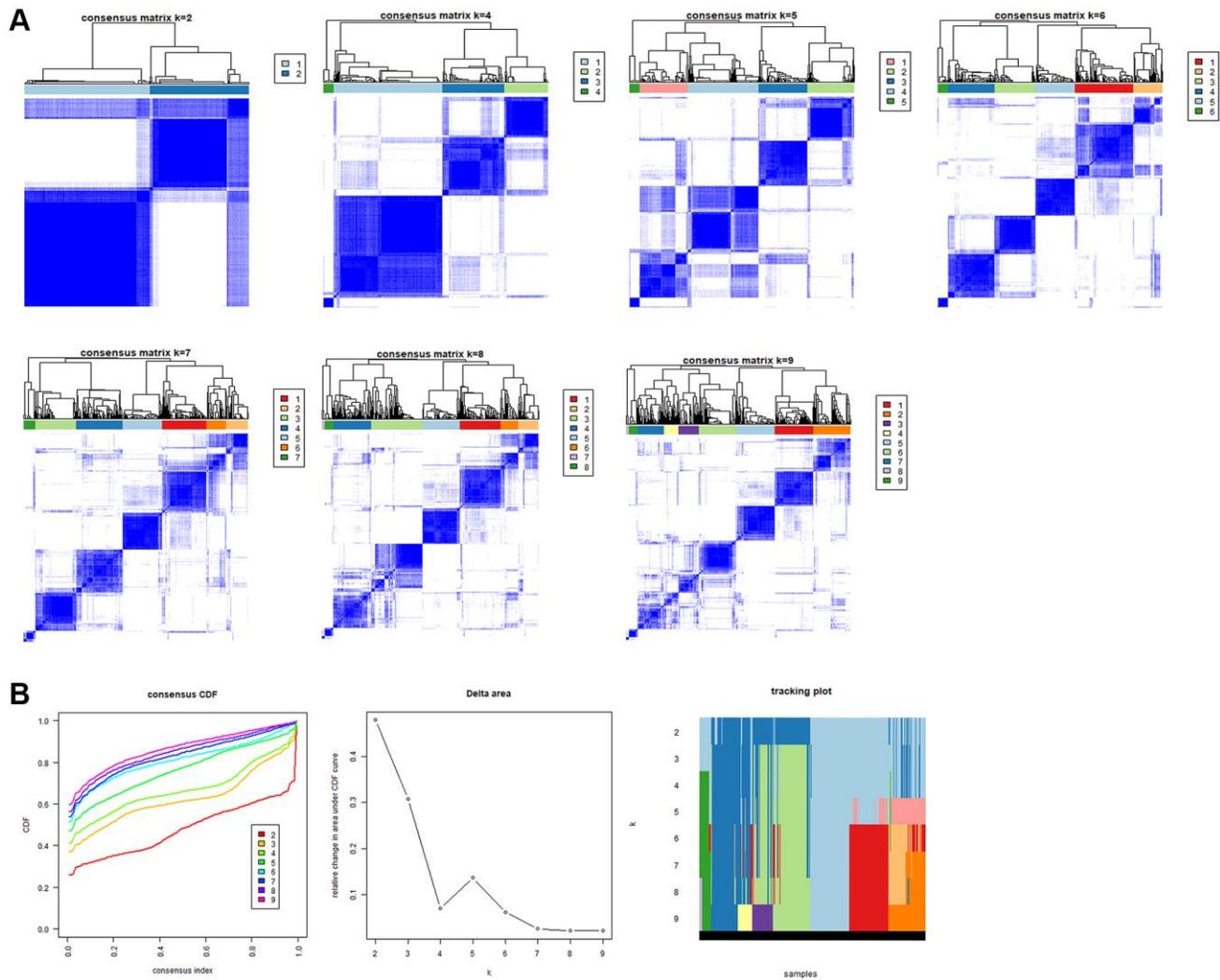
Supplementary Figure 2. The supplementary results of unsupervised clustering analysis based on 49 FAMs. (**A**) Consensus matrices of 49 FAMs in HCC. (**B**) The cumulative distribution function (CDF), relative change in area under the CDF curve, and tracking plot.
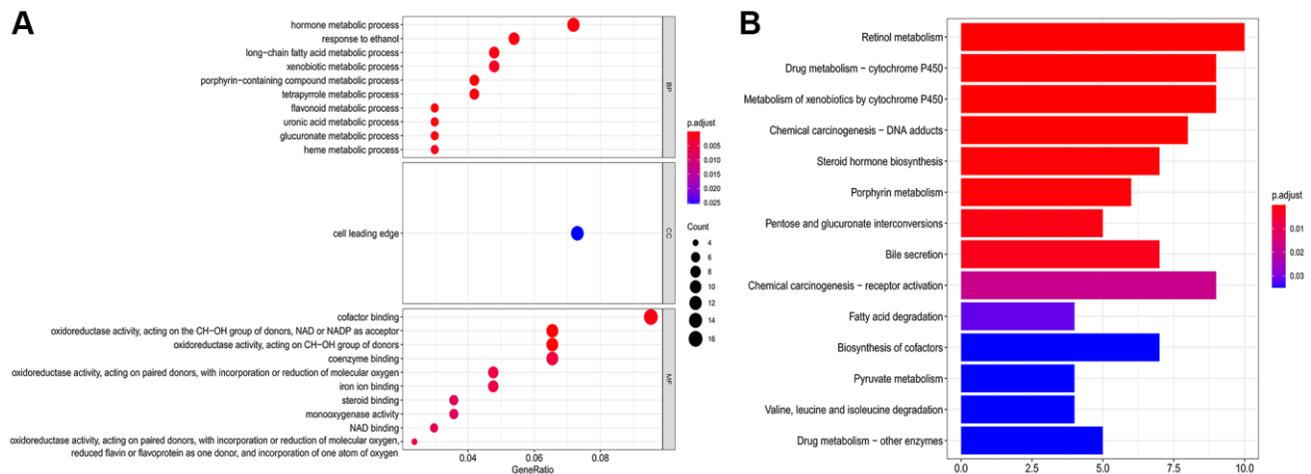
Supplementary Figure 3. GO (**A**) and KEGG (**B**) analysis of 190 DEGs among three FAM clusters.
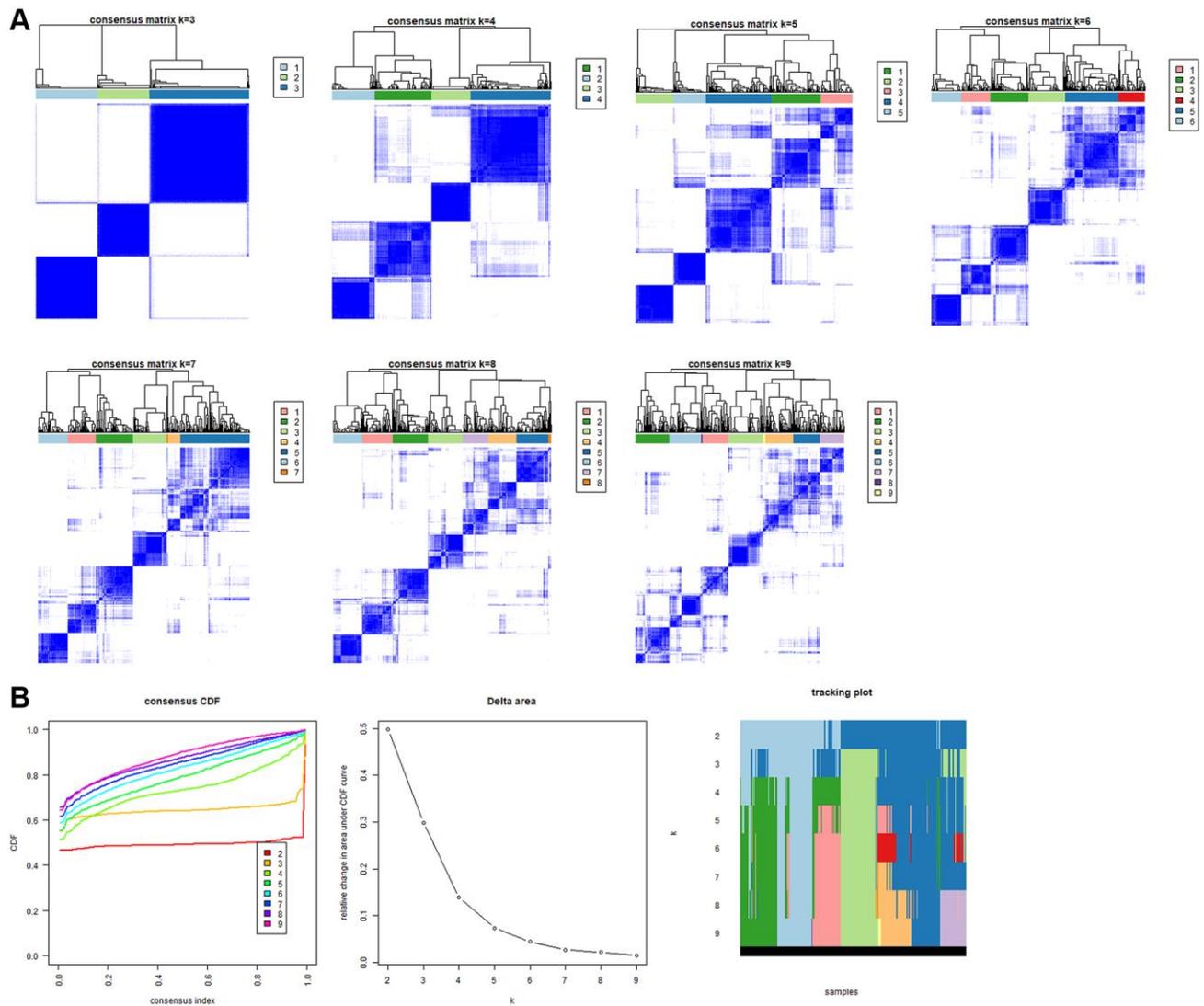
Supplementary Figure 4. The supplementary results of unsupervised clustering analysis based on 79 prognostic DEGs. (**A**) Consensus matrices of 79 prognostic DEGs in HCC for k = 3~9. (**B**) The cumulative distribution function (CDF), relative change in area under the CDF curve, and tracking plot.
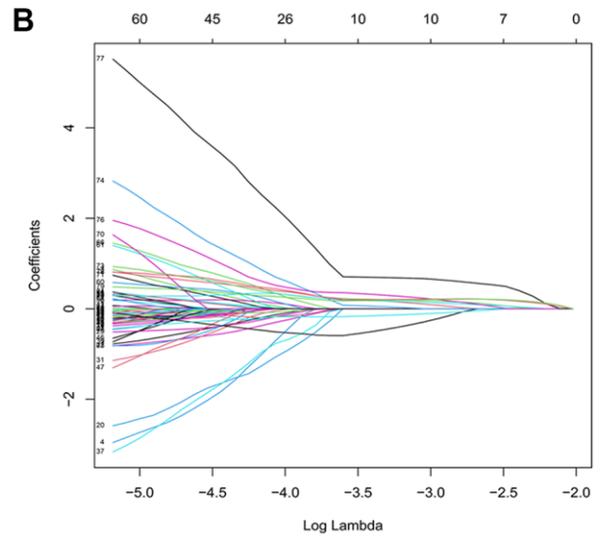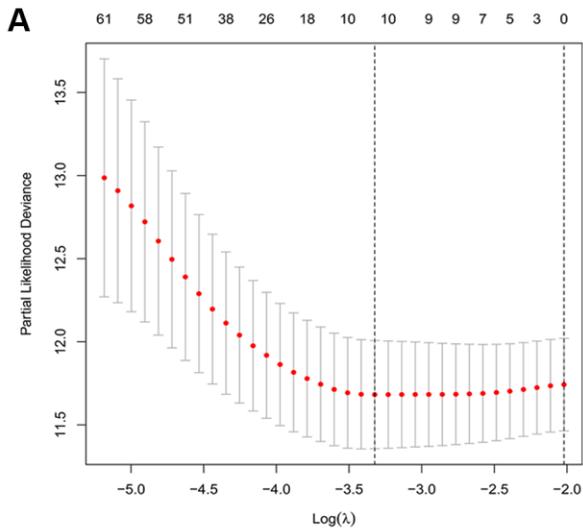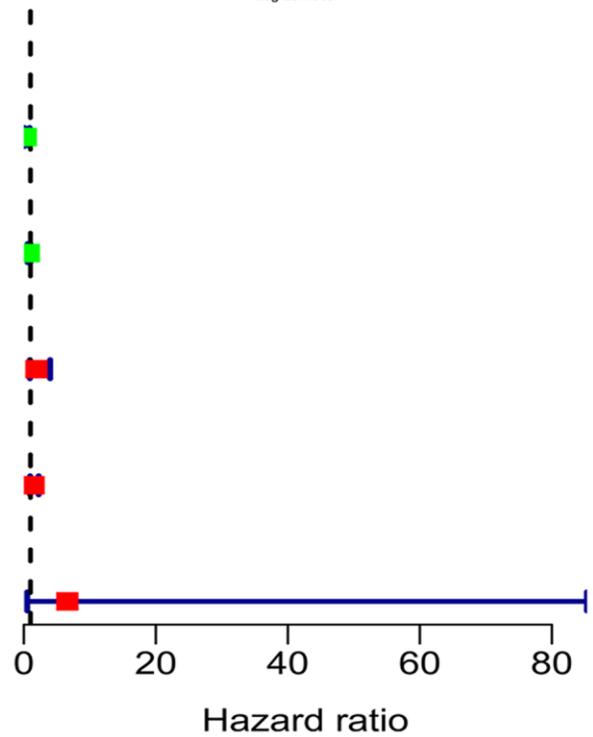
Supplementary Figure 5. Construction of risk prognostic model. (**A**) Screening of optimal parameter (lambda) at which the vertical was drawn. (**B**) LASSO coefficient profiles of the 10 DEGs with non-zero coefficients determined by the optimal lambda. (**C**) Multivariate Cox regression of five candidate genes. Green represents low-risk genes and red represents high-risk genes.
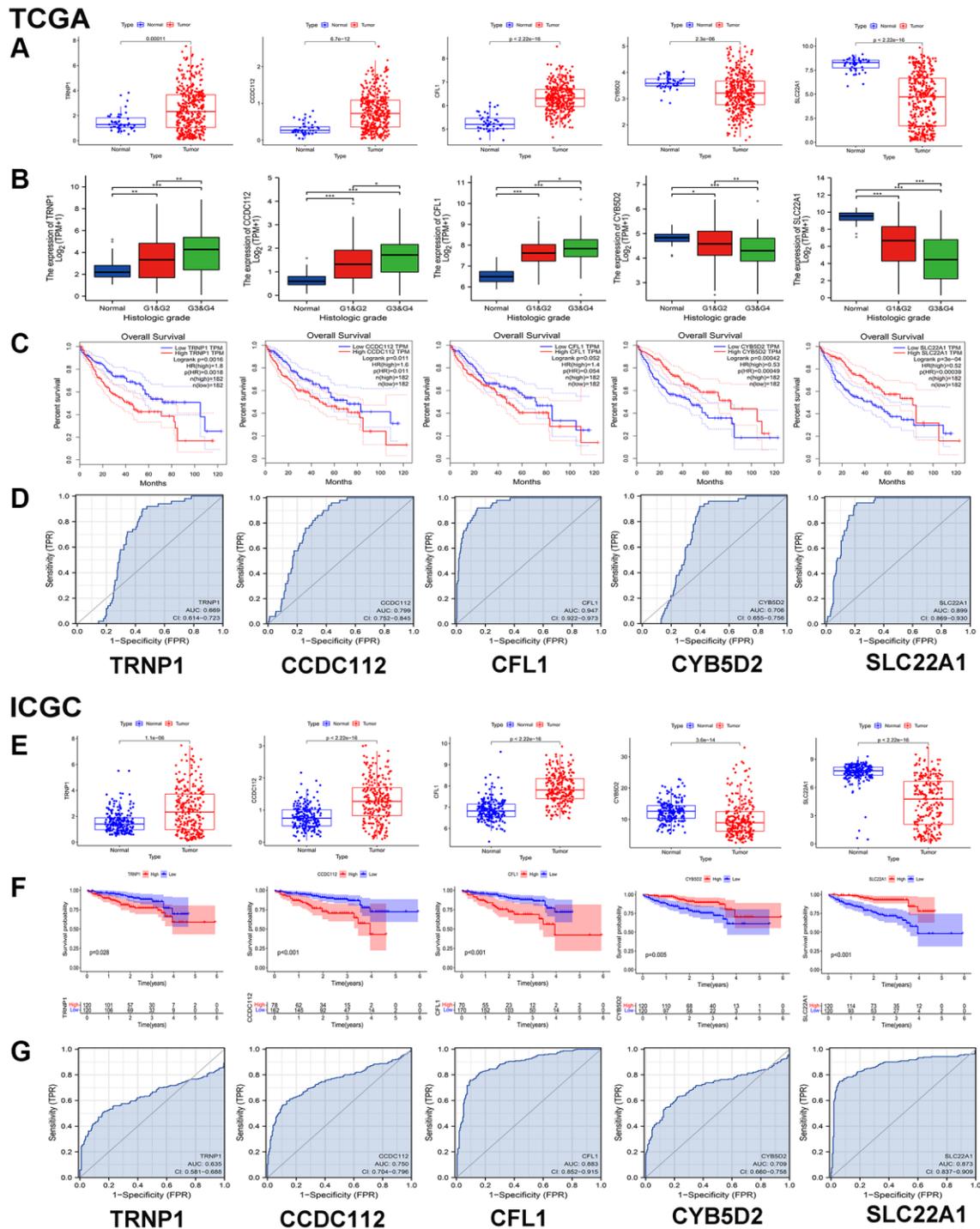
Supplementary Figure 6. The expression of five model genes between tumor and normal tissues in TCGA database (**A**). The expression of five model genes among different pathological grade (**B**) in TCGA database. The K-M curves of five genes in TCGA database (**C**). ROC curves of five genes in TCGA database (**D**). The expression of five model genes between tumor and normal tissues (**E**), the K-M curves five model genes (**F**), the ROC curves of five genes (**G**) in ICGC database.
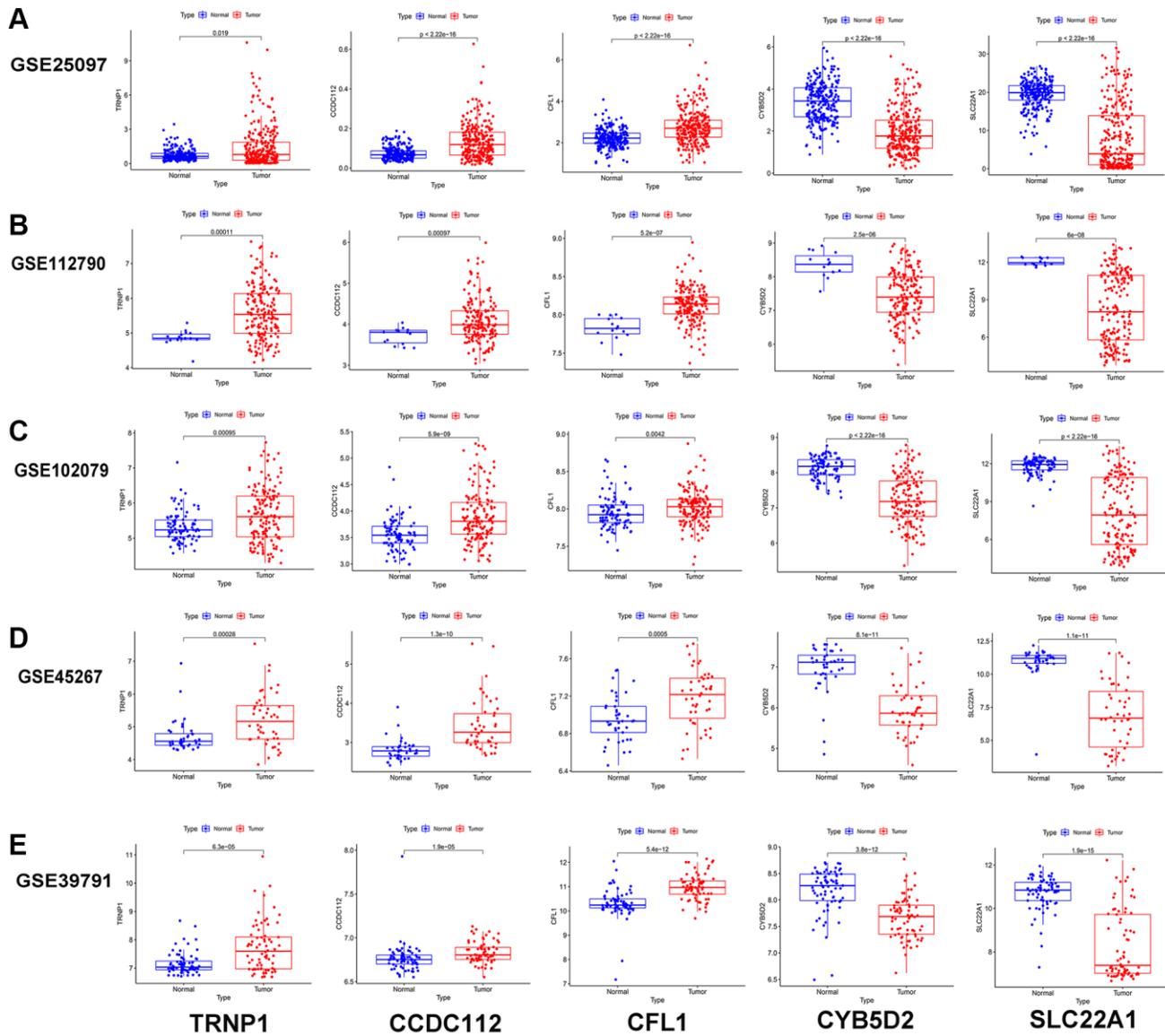
Supplementary Figure 7. Five genes expression in GSE25097 (**A**), GSE112790 (**B**), GSE102079 (**C**), GSE45267 (**D**), GSE39791 (**E**) dataset.
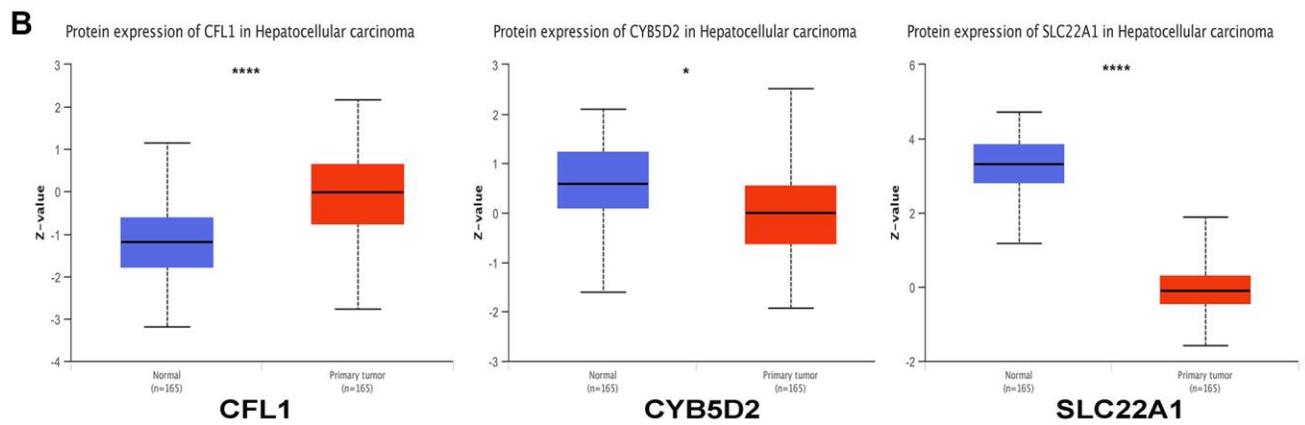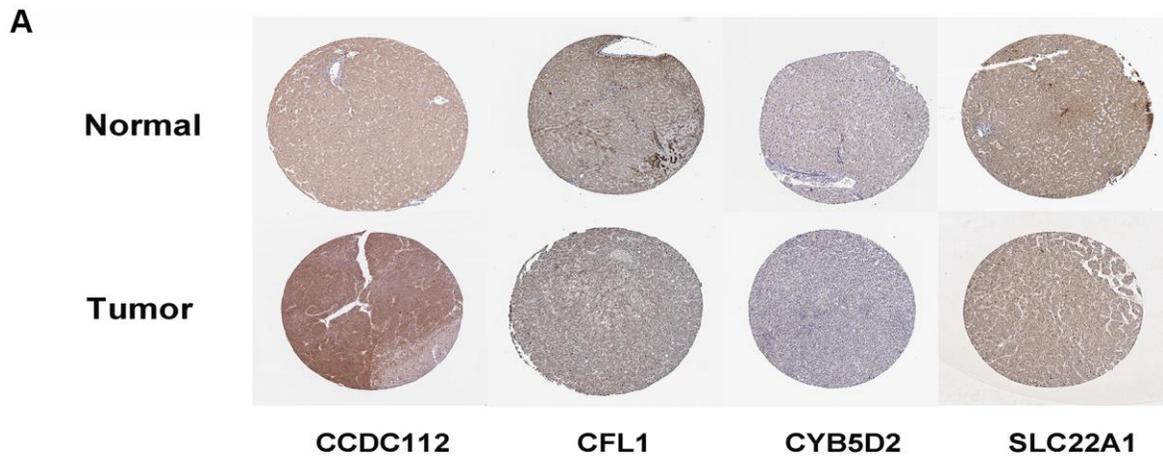
Supplementary Figure 8. Immunohistochemistry analysis of CCDC112, CFL1, CYB5D2, SLC22A1 expression in HPA database (**A**); Protein levels of CFL1, CYB5D2, SLC22A1 in CPTAC database (**B**).
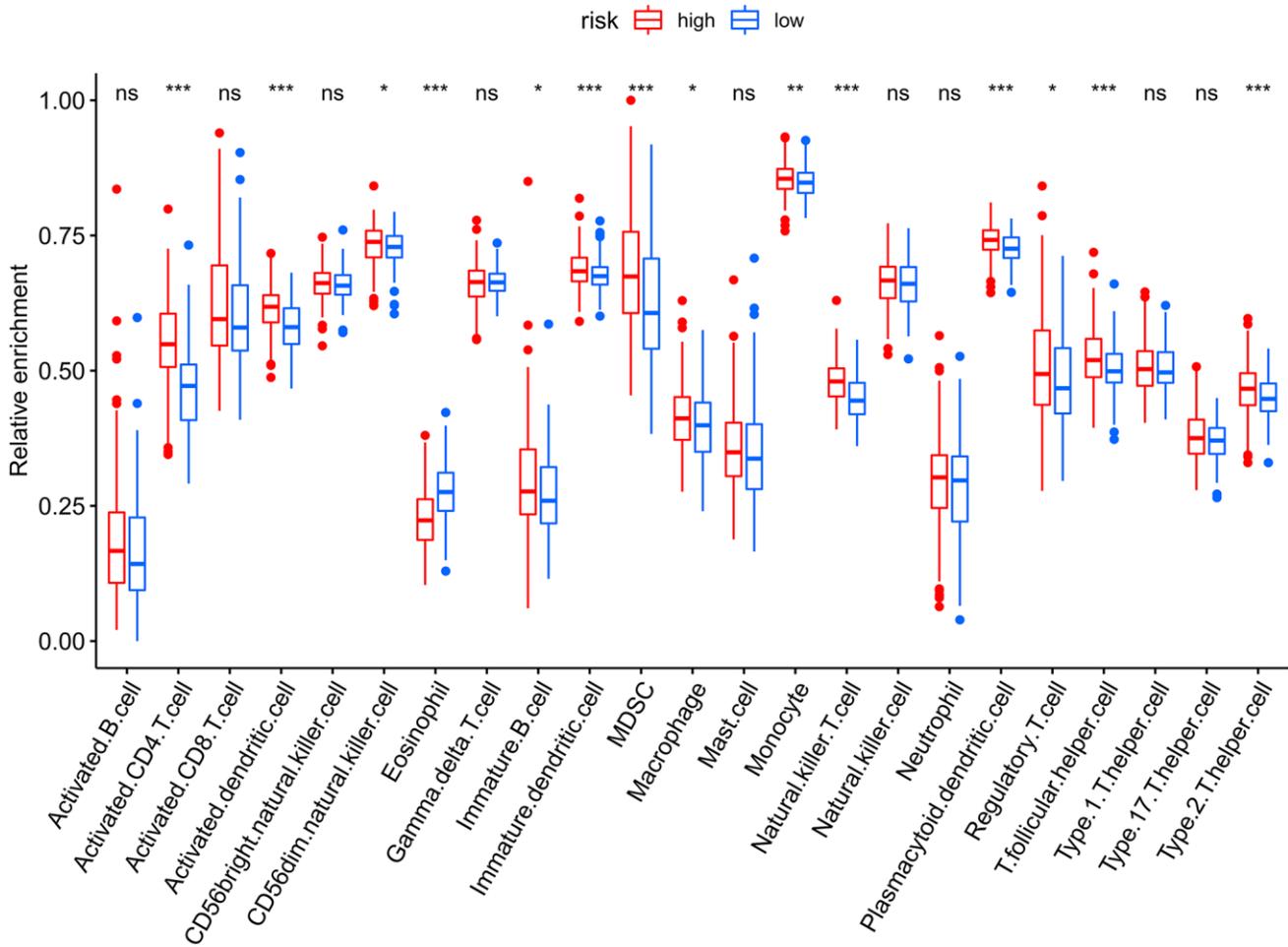
Supplementary Figure 9. The infiltration levels of 23 kinds of immune cells of TME in the two risk groups.

Supplementary Tables

Supplementary Table 1. Fatty acid metabolism-related genes.

Gene symbol	Description	Category	Relevance score
FASN	Fatty Acid Synthase	Protein Coding	88.04946899
FABP2	Table-S1:Fatty Acid Binding Protein 2	Protein Coding	80.34545898
LIPA	Lipase A, Lysosomal Acid Type	Protein Coding	79.07958984
INS	Insulin	Protein Coding	78.31687164
SLC17A5	Solute Carrier Family 17 Member 5	Protein Coding	75.35491943
FADS1	Fatty Acid Desaturase 1	Protein Coding	73.69827271
HADHA	Hydroxyacyl-CoA Dehydrogenase Trifunctional Multienzyme Complex Subunit Alpha	Protein Coding	73.44226837
FABP4	Fatty Acid Binding Protein 4	Protein Coding	72.57804871
PPARG	Peroxisome Proliferator Activated Receptor Gamma	Protein Coding	70.81949615
FABP1	Fatty Acid Binding Protein 1	Protein Coding	70.25164032
CPT2	Carnitine Palmitoyltransferase 2	Protein Coding	69.19077301
ACADM	Acyl-CoA Dehydrogenase Medium Chain	Protein Coding	66.55280304
MTR	5-Methyltetrahydrofolate-Homocysteine Methyltransferase	Protein Coding	65.85922241
ALB	Albumin	Protein Coding	65.84506226
FADS2	Fatty Acid Desaturase 2	Protein Coding	65.0801239
FABP3	Fatty Acid Binding Protein 3	Protein Coding	63.80797577
DDC	Dopa Decarboxylase	Protein Coding	62.60398865
FA2H	Fatty Acid 2-Hydroxylase	Protein Coding	62.45428848
PPARA	Peroxisome Proliferator Activated Receptor Alpha	Protein Coding	61.77651215
BAAT	Bile Acid-CoA:Amino Acid N-Acyltransferase	Protein Coding	61.54239273
CPT1A	Carnitine Palmitoyltransferase 1A	Protein Coding	61.52598572
ACADVL	Acyl-CoA Dehydrogenase Very Long Chain	Protein Coding	61.33808136
HADHB	Hydroxyacyl-CoA Dehydrogenase Trifunctional Multienzyme Complex Subunit Beta	Protein Coding	60.69838715
MMACHC	Metabolism Of Cobalamin Associated C	Protein Coding	60.51202393
ABCD1	ATP Binding Cassette Subfamily D Member 1	Protein Coding	60.29699326
LPL	Lipoprotein Lipase	Protein Coding	58.79434204
FABP5	Fatty Acid Binding Protein 5	Protein Coding	57.8329277
FAAH	Fatty Acid Amide Hydrolase	Protein Coding	57.01290512
PNPLA3	Patatin Like Phospholipase Domain Containing 3	Protein Coding	56.81323242
CYP2D6	Cytochrome P450 Family 2 Subfamily D Member 6	Protein Coding	55.99719238
AKR1D1	Aldo-Keto Reductase Family 1 Member D1	Protein Coding	54.9125061
CYP2C19	Cytochrome P450 Family 2 Subfamily C Member 19	Protein Coding	53.54149628
NR1H4	Nuclear Receptor Subfamily 1 Group H Member 4	Protein Coding	53.41639328
ACADS	Acyl-CoA Dehydrogenase Short Chain	Protein Coding	53.31412506
ACOX1	Acyl-CoA Oxidase 1	Protein Coding	52.61158371
ADIPOQ	Adiponectin, C1Q And Collagen Domain Containing	Protein Coding	52.42185211
SCD	Stearoyl-CoA Desaturase	Protein Coding	52.2116394
SLC27A1	Solute Carrier Family 27 Member 1	Protein Coding	52.18554688
FABP12	Fatty Acid Binding Protein 12	Protein Coding	52.18073273
FABP6	Fatty Acid Binding Protein 6	Protein Coding	52.11978912
CD36	CD36 Molecule	Protein Coding	52.04457474
LIPC	Lipase C, Hepatic Type	Protein Coding	51.4848938
APOB	Apolipoprotein B	Protein Coding	50.64818192

APOE	Apolipoprotein E	Protein Coding	50.5888176
PHYH	Phytanoyl-CoA 2-Hydroxylase	Protein Coding	50.57295227
HADH	Hydroxyacyl-CoA Dehydrogenase	Protein Coding	50.49472809
CYP3A4	Cytochrome P450 Family 3 Subfamily A Member 4	Protein Coding	50.4631424
AMACR	Alpha-Methylacyl-CoA Racemase	Protein Coding	50.12758636
HSD17B4	Hydroxysteroid 17-Beta Dehydrogenase 4	Protein Coding	50.03477859

Supplementary Table 2. 79 prognostic genes identified by unicox analysis.

ID	HR	HR.95L	HR.95H	*p* value
CYP3A4	0.778076355	0.631093988	0.95929105	0.018822658
RANBP3L	0.485929735	0.257884755	0.915632671	0.025573005
SORD	0.553252858	0.315242258	0.970963495	0.039146754
SEC14L2	0.556356443	0.35219986	0.878854671	0.011952572
AQP9	0.677623041	0.493691285	0.930081205	0.016015165
ALAS1	0.341885541	0.131021095	0.892113769	0.028289079
ADH1B	0.605729859	0.438281266	0.837153423	0.002392409
ECM2	0.490288961	0.28657424	0.838816724	0.009282406
PACRG	0.489240618	0.272397527	0.87870248	0.016722256
GRHPR	0.320862856	0.105618949	0.97475854	0.044956842
APOC4	0.564618825	0.325778952	0.978560512	0.041629443
ZNF385B	0.599719641	0.383262564	0.938426242	0.025210962
MYRIP	0.654531516	0.44584435	0.960899259	0.030497356
UNC119	2.406159309	1.076484474	5.378249999	0.032390758
COBLL1	0.401391037	0.187490206	0.859323633	0.018756828
CYP3A43	0.494188942	0.307679981	0.79375561	0.003553024
ST6GALNAC4	4.577439148	2.318914107	9.035672813	1.16456E-05
TMSB10	3.52740543	1.218053353	10.21514291	0.020149295
TRNP1	2.23135374	1.514507558	3.287497302	4.9203E-05
KCTD17	2.14699596	1.313742559	3.50874806	0.002297419
ADI1	0.209793466	0.067094199	0.655992608	0.007257318
CD24	1.433969982	1.041136815	1.975023724	0.027329796
GTF3C6	13.57443192	3.603410607	51.13633223	0.000116072
PAFAH1B3	2.512457785	1.407123335	4.486063136	0.001841278
PDK4	0.541413324	0.326464404	0.897887746	0.017440487
IKBKE	2.254088928	1.312611493	3.87084596	0.003219696
ANKRD13D	3.523342088	1.528676787	8.120709085	0.003115166
LRRC1	1.808283501	1.125661734	2.904859535	0.014308862
ARPC3	10.79177831	1.894809774	61.4639426	0.007361876
GLTP	17.37701093	3.613247439	83.57039306	0.000366508
SLC2A6	1.715164093	1.051338438	2.798135939	0.030738336
BTNL9	0.366002192	0.201388798	0.665169096	0.000975159
OAZ1	18.94482761	2.560843715	140.1516583	0.003965072
CIB2	1.766576726	1.09180951	2.858367965	0.020464379
SH3BGRL3	6.03881903	1.808128189	20.16855635	0.003471103

NFKBIE	2.896363352	1.295007694	6.477892534	0.009613449
GNAZ	1.915092976	1.255313579	2.92164537	0.002568985
GRAMD1A	2.411861311	1.140248758	5.101584143	0.021259918
PLP2	2.429269296	1.239697358	4.760314505	0.009710398
P2RX4	3.309419925	1.426343897	7.678555126	0.00532122
BAK1	3.838358962	1.703445451	8.648941187	0.001174366
DYNLT1	13.36161281	3.008589083	59.34100404	0.000654496
CAPG	2.309200988	1.383797324	3.853461134	0.001358784
CCDC149	3.732128772	1.652195378	8.430470968	0.001536791
TRAPPC4	14.09480342	4.075780094	48.74244409	2.92156E-05
C11orf80	2.560164894	1.288077049	5.088549859	0.007312492
SMPD2	5.253812801	1.795266114	15.37518518	0.002461619
RGS19	3.177490631	1.340891369	7.529652994	0.008630722
ARPC1B	4.21079628	1.520584487	11.66051966	0.005667732
S100A6	1.83503394	1.073343413	3.13725274	0.026511316
C12orf49	2.691628705	1.383349689	5.237189949	0.003551498
PHLDA2	2.157304562	1.382549384	3.366218252	0.000706858
EGLN3	2.568763817	1.705472772	3.869043035	6.34526E-06
MMP9	1.493898792	1.071488971	2.082833945	0.017924239
RNFT2	3.528666911	1.790133293	6.955621806	0.000270858
NRM	2.327320495	1.207843706	4.484372159	0.011594105
GABARAPL1	0.387683107	0.17874024	0.840874958	0.016453825
GLS	2.339010916	1.280372041	4.272954962	0.005712296
PDE4A	1.807937851	1.009799228	3.236919955	0.046285649
TMEM189	6.608824384	2.229300737	19.59204473	0.000659563
ANKRD24	0.30219409	0.136574227	0.668656669	0.003144648
TRPM2	2.771002254	1.427282838	5.37977007	0.002603816
ABHD12	13.09979991	2.834703327	60.5371137	0.000987256
CCDC112	3.57319329	1.807429626	7.064015164	0.000250133
ASRGL1	2.309893898	1.433019387	3.723334011	0.000588197
RBM38	3.92752773	1.482062082	10.408116	0.005937778
CFL1	128.1863798	14.51617911	1131.960955	1.25869E-05
DOK1	2.09453848	1.020148304	4.300444777	0.043976754
IFI27L2	1.761646639	1.025608679	3.025909339	0.040210618
TMED9	8.276218821	1.054130849	64.97845883	0.044420152
SAAL1	6.133981755	1.871604613	20.10346197	0.002745506
MMD	3.088608202	1.519324905	6.278775918	0.001836454
CYB5D2	0.179105725	0.05980064	0.536430053	0.002120822
RRM2	3.054808696	1.703875719	5.476840867	0.000177522
AKR1D1	0.6053703	0.414352169	0.884448612	0.009465756
C16orf45	0.50778156	0.289491837	0.890671447	0.018088686
SLC22A1	0.623023973	0.474713929	0.817669016	0.000646909
FETUB	0.702168912	0.515209768	0.956971726	0.025195346
CD14	0.373053627	0.139944271	0.994460202	0.048716011

Supplementary Table 3. The regression coefficient of five candidate genes.

ID	coef	HR	HR.95L	HR.95H	*p* value
TRNP1	0.382912477	1.466549668	0.954978512	2.25216369	0.080206219
CCDC112	0.650210449	1.915943995	0.921571858	3.98323946	0.081641119
CFL1	1.885657032	6.59068331	0.509098487	85.32161766	0.148951447
CYB5D2	−1.230989748	0.292003425	0.09700108	0.879021141	0.028575575
SLC22A1	−0.290320823	0.748023546	0.557853321	1.003022129	0.052403777