A Novel Nomogram Based on Lipid Metabolism-related Risk Gene Expression Can Better Predict Overall Survival for Hepatocellular Carcinoma

Qiliang Lu (drluqiliang@foxmail.com)
Medical College of Qingdao University, Qingdao, 266071, China

Wen Fu
Medical College of Qingdao University, Qingdao, 266071, China

Guangxiong Ouyang
Medical College of Qingdao University, Qingdao, 266071, China

zunqiang xiao
Department of Hepaticobiliarypancreatic and Minimally Invasive Surgery, Zhejiang Provincial People's Hospital, People's Hospital of Hangzhou Medical College, Hangzhou 310014 ,China

Jinhui Guo
Medical College of Qingdao University, Qingdao, 266071, China

Wen Sun
The Second Clinical Medical College of Zhejiang Chinese Medical University, Hangzhou, 310014, China

Xiangxiang Lei
Hangzhou Medical College, Hangzhou, Zhejiang 310000, China

Yu Shi
College of Materials Science and Enginnering, Zhejiang University of Technology ,Hangzhou 310014 ,China

Qiancheng Ma
Zhejiang University of Technology, Hangzhou 310014, China

ning zhu
Hangzhou Medical College, Hangzhou, Zhejiang 310000, China

Jie Zhao
Zhejiang University of Technology, Hangzhou 310014, China

Dongsheng Huang
The Key Laboratory of Tumor Molecular Diagnosis and Individualized Medicine of Zhejiang Province, Zhejiang Provincial People's Hospital Affiliated People's Hospital, Hangzhou Medical College, Hangz

Research Article

Keywords: hepatocellular carcinoma, lipid metabolism, LASSO, prognostic model, nomogram, biomarker

Posted Date: January 20th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-142865/v2

Abstract

Metabolic reprogramming has been proven to be a hallmark of cancer. The pathogenic factors involved in Hepatocellular carcinoma (HCC) lead to an abnormal lipid metabolism that facilitates the malignant transformation of liver cells. However, the association between lipid metabolism and the prognosis of HCC has not been systematically delineated. In this study, the training set comprised 221 patients from The Cancer Genome Atlas (TCGA) based on the gene expression details, whereas 230 patients within the International Cancer Genome Consortium (ICGC) comprised the validation set. Ten lipid metabolism-related risk genes were screened; they were found to be significantly related to the prognosis of HCC. The risk score was calculated based on ten screened lipid metabolism-related risk genes and was confirmed to be an independent prognostic factor for HCC even when excluding clinical features. Therefore, a novel nomogram integrating the risk score and other proven clinical attributes was constructed. The results of the area under the receiver operating characteristics curve (AUC), C index, and calibration plot supported the better predictive capacity of the nomogram over others. Treatment with metformin significantly positively affected the expression of four out of ten genes; this was beneficial to longer overall survival. The results provide a new insight into accurate prognostic prediction, as well as understanding the carcinogenesis and process of HCC.

1. Introduction

Liver cancers have the sixth-highest incidence of all cancers and are the fourth leading cause of cancer-related deaths. There are approximately 841,000 new cases and 782,000 deaths each year [1]. Its morbidity and mortality are increasing year by year. The American Cancer Society estimates that there would be more than 42000 new cases and 30000 deaths by the end of 2020 in the USA alone [2]. Hepatocellular carcinoma (HCC) accounts for 85 to 90 percent of primary liver cancers and had been a hot spot in cancer research. The efficacy of modern diagnostic and therapeutic options against HCC is unsatisfactory [3]. HCC with a 5-year survival of 18%, is only less malignant than pancreatic cancer [4]. Many studies have endeavored to develop an ideal tool for HCC prognosis prediction. However, the optimal models have not been established yet.

The initiation of HCC is closely related to the underlying liver disease, such as hepatitis B or C virus (HBV or HCV) infection, Aflatoxin B1 (AFB1) infection, or alcohol abuse. They, singly or synergistically, cause liver cell fat degeneration and lipid deposition, which leads to an imbalance in liver lipid metabolism and facilitate malignant transformation of liver cells [5]. Studies have reported that the plasma levels of triglycerides, cholesterol, free fatty acids, high- and low-density lipoproteins, and apolipoproteins were significantly reduced in most liver cancer patients [6,7]. In western countries, nonalcoholic fatty liver disease (NAFLD) may soon become the dominant causative factor in HCC [8]. Metabolic reprogramming can be a hallmark of cancer [9]. In early 1953, Medes *et al*. described the increased de novo lipid synthesis metabolic alteration in cancer. They concluded that essential lipids for cancer cell growth are obtained from the host [10]. Multiple studies have focused on lipid metabolism and the lipogenic phenotype in cancer cells [11]. Some potential drugs targeting lipid metabolic reprogramming have undergone clinical trials [12]. Metformin is commonly used for blood sugar control in diabetic patients, especially those with excessive body mass index (BMI). It inhibits hepatic gluconeogenesis and reduces hepatic glycogenolysis. Tseng CH found that metformin reduces the risk of HCC in a specific dose-response pattern [13]. Another study showed that type 2 diabetes promotes HCC through insulin resistance [14]. Metformin does not increase insulin secretion by stimulating islet B cells. It directly acts on the metabolic process of sugar, promotes the anaerobic glycolysis of sugar, and increases glucose uptake and utilization by peripheral tissues such as muscles and fat. This unique mechanism of action may help reduce insulin resistance and further benefit HCC patients.

Our research found a lipid metabolism-related HCC gene set related to the prognosis of HCC, and calculated an HCC prognostic risk score depending on screened lipid metabolism-related genes through the LASSO regression analysis. We established a nomogram for HCC prognostic prediction by combining the risk score with clinical factors.

2. Materials And Methods

2.1. Consistency Clustering Analysis

The lipid metabolism-related gene set were downloaded from the Gene Set Enrichment Analysis (GSEA) (https://www.gsea-msigdb.org/ gene set: GO_GLYCEROLIPID_METABOLIC_PROCESS). Gene expression data and clinical details were obtained from the cancer genome atlas (TCGA) (https://portal.gdc.cancer.gov/) and International Cancer Genome Consortium (ICGC) (https://icgc.org/) databases. The patients with unclear pathological diagnosis or follow-up times < 30 days were excluded. A total of 451 HCC patients were enrolled in the study. Among them, 221 cases from the TCGA were considered as the training set and 230 from the ICGC group as the validation set (Table supplement 1). A consistent clustering of metabolic genes set was conducted in the TCGA database using the "Consensus Cluster Plus" package of R (https://www.r-project.org). The cumulative distribution function (CDF) and consensus matrices were carried out to estimate the best numeral of clusters. The discrepancy of gene expression between the clusters was evaluated using principal component analysis (PCA) through R package named “princomp”.

2.2. Identification, Selection and Evaluation of DEGs

The R and “limma” Bioconductor packages were used to identify different expression genes (DEGs) with |logFC|>1 and FDR<0.05. A PPI network was constructed using String (Version10.5, http://string-db.org) with confidence >0.9 as a cut-off criterion for up- or down-regulated genes in both lipid metabolism and HCC. The analysis was executed to find out which pathways the screened DEGs enriched in, using DAVID (Database for Annotation, Visualization and Integrated Discovery, version 6.8, https://david-d.ncifcrf.gov), Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO).

2.3. Confirmation of Hub Risk Genes and Patients Grouping According to the Level of Risk Score

All DEGs were further screened using univariate analysis and LASSO regression. The filtered DEGs were then selected to build a risk score formula as follows: Risk score = (coefficient * expression of gene 1) + (coefficient * expression of gene 2) + ... + (coefficient * expression of gene X). The patients were separated into low- and high-risk groups based on the cut-off value defined by the median value of risk score in both the training and validation sets.

2.4. Analysis of Screened Risk Genes

Survival curves of screened DEGs were drawn according to the gene expression in the training and validation sets using Kaplan-Meier analysis. Later, the proteins encoded by screened risk genes were analyzed using The Human Protein Atlas (https://www.proteinatlas.org) to figure out if there was a distribution difference between HCC and adjacent tissues. We used GSEA (http://software.broadinstitute.org/gsea) to find potential functional annotations about these genes.

2.5. The Correlation between the Risk Score and Clinical Features

We compared the prognostic process and clinical characteristics of the high and low-risk groups using Kaplan-Meier analysis in both the training and validation sets. The relationship between the risk score and clinical characteristic was assessed using univariate and multivariate Cox analysis.

2.6. Construction and Evaluation of the Nomogram

Sex, age, TNM stage, and risk score were selected as prognostic factors to establish a nomogram. The area under the receiver operating characteristic curve (AUC), C index, and calibration plot were performed to assess the precision of the nomogram in both the training and validation sets.

2.7. Effect of Metformin Treatment on Risk Genes

The expression changes of all screened risk genes in GSE69850 was investigated to probe the correlation between risk genes and metformin intake in HCC. This included 9 samples of HepG2 cells handled by metformin and another 39 control samples handled by dimethyl sulfoxide (DMSO), through unpaired t-tests in GraphPad Prism 8.0.

3. Results

3.1. Subtype of HCC Owing to the Lipid Metabolism-related Gene Set

The patients were divided into 2 clusters (K=2) by “consensus” to unscramble the correlation of lipid metabolism-related genes expression with outcome of HCC, (Figure 1A-C). PCA revealed two clusters that presented significant differences (Figure 1D). A chi-square test showed the difference in sex, age, tumor grade, and TNM stage between two clusters (Table supplement 1).

3.2. Selection and Evaluation of The Prognostic Lipid Metabolism-related Genes

Lipid-metabolism DEGs (214) were identified between HCC tissues and adjacent non-tumor tissues (Figure 1E). A PPI network was built to inspect the interaction among the 214 DEGs, by utilizing the STING tool (Figure 1F), and GO and KEGG analysis were performed. The GO analysis showed three main DEGs pathways including glycerolipid metabolic process, plasma lipoprotein particle, and phosphoric ester hydrolase activity (Figure 2A). The KEGG analysis showed three main DEGs pathways containing glycerolipid, phospholipid, and glycerophospholipid metabolic process (Figure 2B).

3.3. Ten Screened Lipid Metabolism-related Genes

Ten risk DEGs (*ACSL3*, *LCLAT1*, *LPCAT1*, *PIGU*, *PLA2G7*, *PLEKHA8*, *PON1*, *PTPMT1*, *SOCS2*, *TBL1XR1*) were chosen after univariate and LASSO regression analysis (Figure 1G and 1H), to originate a risk score formula. Patients were classified into high and low-risk groups in terms of the median risk score (0.71) (Figure supplement 1). The Kaplan-Meier plot showed that the cohorts at the high-risk group had a shorter survival than those at low risk, both were investigated in two sets (Figure 1I and 1J). GSEA revealed that all ten risk genes were involved in ten pathways. The high-risk score group mainly enriched in vascular smooth muscle contraction, hypertrophic cardiomyopathy (hcm), neuroactive ligand-receptor interaction, calcium signal pathway, dilated cardiomyopathy pathways, whereas the low-risk score group mainly enriched in homologous recombination, cell cycle, pyrimidine metabolism, RNA degradation, and spliceosome pathways (Figure 2C-L).

3.4. Eight of Ten Lipid Metabolism-related Genes Were Related to HCC Prognosis

High expression of *ACSL3*, *LCLAT1*, *LPCAT1*, *PIGU* (Figure 3A-D), *PTPMT1* (Figure 3H) and *TBL1XR1* (Figure 3J) and low expression of *PON1* (Figure 3G) and *SOCS2* (Figure 3I) were negatively correlated with a good outcome in HCC patients in the TCGA database. *PLA2G7* and *PLEKHA8* showed no significant correlation (Figure 3E-F). However, in the ICGC database, high expression of *ACSL3*, *LCLAT1*, *LPCAT1*, *PIGU* (Figure supplement 2A-D), and *TBL1XR1* (Figure supplement 2J) and low expression of *PON1* (Figure supplement 2G) were negatively correlated with a good outcome in HCC patients. We also searched The Human Protein Atlas to investigate all ten proteins encoded by ten key DEGs; Eight proteins (*ACSL3*, *LCLAT1*, *LPCAT1*, *PIGU*, *PTPMT1*, *TBL1XR1*, *PON1*, and *SOCS2*). Immunohistochemistry showed significant differences in the distribution of all eight proteins between cancers and adjacent tissues (Figure 3K-Z). The dependence of all eight related genes and survival rates in our study were statistically significant.

3.5. The Risk Score Was An Independent Prognostic Factor for HCC

We investigated the correlation of the risk score and clinicopathological factors of HCC patients in two sets. There were significant differences in grade, T stage, and TNM stage but no significant differences in sex and age between the high and low-risk groups (Table 1). Moreover, poor differentiation, higher T stage, and worse TNM stage are positively correlated with the risk score (Figure 4A-F). Layered comparison displayed that the high-risk group suffered a shorter survival in both the training and validation sets, in the same conditions of sex, age, tumor grade, T stage, and TNM stage (Figure 4G-N and Figure supplement 3A-D).

Table 1
Correlation between the risk score and clinicopathological factors of HCC patients in two sets

Features	TCGA (N=221)				ICGC(N=230)			
	Low risk (N=111)	High risk (N=110)	χ^2	*P* value	Low risk (N=124)	High risk (N=106)	χ^2	*P* value
Sex			1.827	0.176			2.144	0.143
female	30	39			28	33		
male	81	71			96	73		
Age			2.049	0.152			0.349	0.53
≤60	57	67			25	25		
>60	54	43			99	81		
TNM stage			10.938	0.008			21.683	0
	66	42			22	14		
	18	28			69	37		
	25	39			31	40		
	2	1			2	15		
Grade			15.006	0.002				
	16	11						
	60	36						
	32	56						
	3	7						
T stage			9.816	0.019				
T1	66	43						
T1	18	29						
T3	24	31						
T4	3	7						
M stage			0.329	1				
M0	109	109						
M1	2	1						
N stage			3.069	0.1222				
N0	111	107						
N1	0	3						

Univariate and multivariate COX regression analyses were carried out to contrast the risk score with typical clinical factors including sex, age, tumor grade, and TNM stage. Univariate COX regression indicated that the risk score was an

important prognostic factor (training set: HR=3.213, 95%CI [1.917-5.384], *P*<0.001, Table 2; validation set: HR=3.586, 95%CI [1.82-7.066], *P*<0.001, Table 2). Additionally, the results of multivariate COX regression showed that the risk score was an independent prognostic factor (training set: HR=3.129, 95%CI [1.8269-5.360], *P*<0.001, Table 2; validation set: HR=2.7158, 95%CI [1.3289-5.5503], *P*<0.001, Table 2).

Table 2
Univariate and multivariate Cox regression analysis of the clinical features and risk score for OS in TCGA and ICGC database

	TCGA univariate analysis		**TCGA multivariate analysis**		**ICGC univariate analysis**		**ICGC multivariate analysis**	
Variables	HR (95% CI)	*P* value	HR (95% CI)	*P* value	HR (95% CI)	*P* value	HR (95% CI)	*P* value
Age								
≤60	Reference		Reference		Reference		Reference	
>60	1.069(0.664-1.721)	0.784	1.051(0.639-1.730)	0.845	0.769(0.385-1.536)	0.457	0.637(0.312-1.300)	0.214
Sex								
Female	Reference		Reference		Reference		Reference	
Male	0.770(0.471-1.257)	0.269	1.149(0.678-1.948)	0.606	0.476(0.253-0.896)	0.002	0.418(0.212-0.823)	0.117
TNM stage								
	Reference		Reference		Reference		Reference	
	1.691(0.818-3.498)	0.156	1.235(0.582-2.618)	0.583	6.247(0.8337-46.81)	0.075	7.284(0.970-54.695)	0.054
	4.225(2.406-7.418)	<0.001	3.822(2.152-6.785)	<0.001	7.598(0.992-58.19)	0.051	9.357(1.207-72.524)	0.032
	8.331(2.448-28.353)	<0.001	7.431(2.035-227.129)	0.002	26.834 (3.393-212.24)	0.002	25.031(3.131-201.919)	0.003
Risk								
Low risk	Reference		Reference		Reference		Reference	
High risk	3.213(1.917-5.384)	<0.001	3.129(1.827-5.360)	<0.001	3.586(1.82-7.07)	<0.001	2.716(1.829-5.550)	0.006

3.6. The Predictive Efficacy of the Nomogram for HCC Prognosis Was Better Than both TNM Stage and Risk Score Alone

A nomogram amalgamating the risk score and clinicopathological features was constructed to predict the overall survival (Figure 5A). In the training set, the AUCs of nomograms for predicting 1-, 3- and 5-year over survival were correspondingly 0.811, 0.799 and 0.799 (Figure 5B-D), the C index was 0.749 and the calibration plots for survival probabilities were displayed in Figure 5E-F. A validation set comprising 230 HCC patients was picked to identify the precision of this nomogram. The AUCs were respectively 0.856, 0.759, and 0.670 (Figure supplement 4A-C), C index was 0.753 and the calibration plots were exhibited (Figure supplement 4D-E).

3.7. Metformin Was Conducive to HCC Prognosis

Unpaired *t*-test results revealed that metformin intake was associated with the expression changes of four genes (*LCLAT1*, *PIGU*, *PON1, PTPMT1*) in GSE69850 ($p<0.05$) (Figure 3AA-AD). DMSO was used as a control.

4. Discussion

Fast-multiplying cancer cells mainly draw energy from increasing aerobic glycolysis, which is known as the “Warburg Effect” [15]. However, the metabolic alterations of cancer cells primarily involve lipid metabolism [16]. Several studies have identified core gene expression signatures for predicting the malignancy of HCC and the outcomes of the patients [17]. However, it is difficult to accurately understand the role of lipid metabolism-related gene sets in liver cancer. In this study, for the first time, a nomogram was built, based on lipid metabolism-related genes and clinical features, to predict the prognosis of HCC. The lipid metabolism-related gene set was downloaded using GSEA. Gene expression profiles and clinical characteristics were determined. The 221 patients in the TCGA database were regarded as the training set, whereas another 230 patients in the ICGC database served as the validation set.

First, a series of screenings on the lipid metabolism-related gene sets were conducted and 10 DEGs associated with the prognosis of HCC patients were picked: *ACSL3*, *LCLAT1*, *LPCAT1*, *PIGU*, *PLA2G7*, *PLEKHA8*, *PON1*, *PTPMT1*, *SOCS2*, *TBL1XR1*. *ACSL3* is ubiquitously expressed in the prostate, brain, and other tissues including the liver. The protein belongs to long-chain fatty-acid-coenzyme A ligase family which is essential in lipid biosynthesis and fatty acid degradation. Chang *et al*. illustrated that endoplasmic reticulum (ER) stress-induced the expression of *ACSL3*. Meanwhile, *ACSL3* shRNA inhibited the induction of lipid accumulation [18]. This phenomenon recommended that *ACSL3* may be a novel therapeutic target towards lipid dysregulation. Tushiro Migita *et al*. showed that *ACSL3* contributes to the growth of castration-resistant prostate cancer (CRPC) through intratumoral steroidogenesis [19]. Haarith Ndiaye *et al*, utilizing immunohistochemical analyses of HCC tissues and subcellular fractionation of cultured HepG2 cells, discovered the increasing expression of *ACSL3* in HCC in contrast to normal liver [20]. In our study, HCC patients with ACSL3 high expression encountered a worse survival rate than those with low expression in both TCGA and ICGC databases. Lysocardiolipin acyltransferase 1 (*LCLAT1*), a cardiolipin-remodeling enzyme of mammalian mitochondrial cardiolipin, modulates mitochondrial membrane potential, cardiolipin remodeling, reactive oxygen species generation, and apoptosis of alveolar epithelial cells [21]. One study demonstrated that *LCLAT1* causes insulin resistance [22]. Another study demonstrated that insulin resistance promotes HCC process. There were no reports about *LCLAT1* on tumors. In our study, metformin intake was related to decreased *LCLAT1* expression. For this study, high expression of *LCLAT1* predicted a poor prognosis in both the training and validation sets. Lysophosphatidylcholine acyltransferase 1 (*LPCAT1*) participates in phospholipid metabolism, particularly in the process of converting lysophosphatidylcholine into phosphatidylcholine when acyl-CoA exists. Bi *et al*. identified *LPCAT1* as a hub among signaling, tumor growth, and the expression of genetically driven growth factor receptors [23]. Mounting evidence suggests that alteration in *LPCAT* activities is involved in the pathological processes, such as NAFLD, viral infections, and cancer [24]. Several studies found that *LPCAT1* is upregulated or overexpressed in human cancers, including colorectal, renal, prostate, lung, and breast cancer [25,26]. Moreover, *LPCAT1* upregulation leads to poor prognosis by promoting progression and recurrence of breast and prostate cancer [27,28]. *LPCAT1* also stimulates brain metastasis of lung adenocarcinoma [29]. For HCC cells cultured in favorable conditions, *LPCAT1* modulated phospholipid composition and distribution. Moreover, *LPCAT1* overexpression promoted HCC cell proliferation, invasion, and migration. *LPCAT1* knockdown produced the opposite effect.[30] In our study, the high expression of *LPCAT1* resulted in poor prognosis in the training set but not obtained in the validation set. Phosphatidylinositol glycan anchor biosynthesis class U (*PIGU*), encoding GPI transamidase fifth subunit, was confirmed as an oncogene for bladder cancer [31]. For HCC, *PIGU* was a significant stage-specific DEG [32].

Additionally, consistent with our study, *PIGU* overexpression was reported as an independent predictive factor for poor prognosis in HCC and the incorporation of *PIGU* expression with a typical TNM stage was thought to elevate prognostic stratification [33]. Phospholipase A2 group VII (*PLA2G7*) catalyzes the activation of the platelet-activating factor. *PLA2G7* defects lead to platelet-activating factor acetylhydrolase deficiency. Moreover, knocking out *PLA2G7* leads to the absence of the activity of soluble lipoprotein-associated phospholipase A2 [34]. Most studies involving *PLA2G7* focus on the process of inflammatory interaction or lipid metabolism in Coronary heart disease, stroke, diabetes, and obesity, but few on tumors [35–38]. Pleckstrin homology domain-containing A8 (*PLEKHA8*), also known as *FAPP2*, participates in vesicle maturation and promotes cytoplasmic lipid transfer. Chen *et al*. demonstrated that *PLEKHA8* overexpression promotes human colon cancer cell growth via an active Wnt signaling [39]. Paraoxonase 1 (*PON1*) is a restricted expression toward the liver, which exhibits lactonase and ester hydrolase activity. The enzyme is synthesized in the liver and kidney and binds to high-density lipoprotein (HDL) particles after being secreted into circulation, and hydrolyzes thiolactones and xenobiotics. Sun *et al*. reported that serum *PON1* level could be used to distinguish early hepatocellular carcinoma from liver cirrhosis with a sensitivity of 71.4% and 95.2% and a specificity of 94.7% and 78.9%, respectively [40]. Ding *et al*. found that the serum level of *PON1* was better than AFP for microvascular invasion prediction and did not fluctuate significantly with the change of tumor size in HCC patients [41]. In our study, *PON1* showed a significant predictive capability for survival rate. *PON1* low expression indicated a better prognosis. Protein tyrosine phosphatase mitochondrial 1 (*PTPMT1*) was a crucial intermediate in cardiolipin biosynthesis and hematopoietic stem cell differentiation [42,43]. In pancreatic beta cells, the downregulation of *PTPMT1* led to an elevation of insulin production and cellular ATP levels [44]. A study reported that *PTPMT1* downregulation promoted cancer cell death [45]. Another reported modulating *PTPMT1* alternative splicing would ameliorate cancer cell radioresistance [46]. Suppressor of cytokine signaling 2 (*SOCS2*) encodes SOCS2 family proteins, which are negative regulators of cytokine receptor signaling via JAK/SATA pathway. *SOCS2* is a well-known cancer suppressor. It inhibits the progression and metastasis in colon, breast, and lung cancer [47–49]. An experiment in mice indicated that *SOCS2* is a modulator of obesity via regulating the metabolic pathways depending on adipocytes' size. Moreover, *SOCS2* also serves as an inflammatory regulator through controlling cell differentiation or recruitment into adipose tissue and cytokines release during the progression of obesity [50]. Another study proved that *SOCS2* plays a protective role in acute liver injury through balancing immune response and oxidative stress [51]. Chen *et a*l. elucidated a mechanism of epigenetic alteration in HCC; *SOCS2* expression was suppressed by methyltransferase-like 3 (*METTL3*) via an m6A-YTHDF2-dependent process [52]. Ren *et al*. concluded that high expression of *SOCS2* inhibits HCC progression via the JAK/STAT pathway related to downregulating miR-196a or miR-196b.[53] In our results, the high expression of *SOCS2* also displayed a protective effect for HCC patients. Transducing (beta)-like 1X-linked receptor 1 (*TBL1XR1*), belonging to WD40 repeat-containing family, presents the sequence identity of *TBL1X* and is required for transcriptional activation. Mutations or recurrent translocations in this gene have been frequently observed in intellectual disability and infrequently in some tumors. Several studies showed that the upregulation of *TBL1XR1* not only promotes cancer cells (including lung, cervical, ovarian, breast and gastric cancer) proliferation, migration, invasion, and metastasis,[54–57] but also suppresses chemotherapy sensitivity in nasopharyngeal carcinoma [58]; therefore causing a bad outcome to cancer patients. Guo *et al*. demonstrated that *TBLR1* was a pivotal oncogene of HCC. Synchronous exhibition about cell proliferation, antiapoptosis, and angiogenesis were observed in both HCC cell lines *in vitro* and samples *in vivo* when the *TBLR1* gene is silenced [59]. For our study, high expression of *TBL1XR1* was statistically related to poor survival. The predictive capacity of all ten DEGs in HCC prognosis was demonstrated regardless of clinical features. Therefore, a risk score based on ten DEGs was calculated.

A nomogram integrating the risk score and clinicopathological features was later constructed. Then, ROC curves were carried out to compare the prognostic values among the nomogram, risk score, age, sex, and TNM stage. The result

suggested that the nomogram could better predict HCC prognostic process. Furthermore, we found that Metformin intake was associated with decreased *LCLAT1*, *PIGU*, and *PTPMT1* expression and increased *PON1* expression. These trends matched the calculated prognosis trends. Therefore, this study speculates that the four genes may offer an effective therapeutic target of HCC with abnormal lipid metabolism. Of note, the lipid metabolism-related risk genes remained an independent prognostic factor even with the exclusion of clinical features. So, combining the risk score and other proven features could produce a better prediction of HCC.

5. Conclusion

5.1 A Novel Nomogram Based on the Lipid Metabolism-related Risk Gene Was Identified for Better Prognostic Prediction of HCC

In conclusion, the lipid metabolism-related gene set was identified as a predictor of the prognosis of HCC. Then, a risk score based on selected ten lipid metabolism-related genes was carried out. A nomogram, combining the risk score and clinical characteristics, was established; the nomogram strongly correlated the survival rate of HCC patients. Furthermore, the selected lipid metabolism-related genes provide new insight into accurate prognostication, and understanding the carcinogenesis and process of HCC. Overall, a novel nomogram based on the lipid metabolism-related risk gene was identified for better prognostic prediction of HCC.

Declarations

Availability of data and materials

All data supporting the findings in our research are available in

TCGA at (https://portal.gdc.cancer.gov) andICGC at (https://icgc.org/). Both two databases are open access for the public.

Acknowlegements

Thank for open access of TCGA and ICGC database.

Fundings

This study was supported in part by the National Natural Science Foundation of China (81672474,81874049), the Co-construction of Provincial and Department Project (WKJ-ZJ-1919), the National Science and Technology Major Project for New Drug (No. 2017ZX09302003004).

Contributions

QL, WF, QM and GO conceived the project and wrote the draft. JG, ZX, WS, XL, NZ, YS accomplished data analysis. ZX, QL, JZ andNZ attended discussion and modified the draft. DH attended discussion and reviewed the manuscript.

Ethics declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

All authors declare no competing interests.

References

1. Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A. & Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin***68**, 394-424 (2018).
2. Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2020. *CA Cancer J Clin***70**, 7-30 (2020).
3. Villanueva, A. Hepatocellular Carcinoma. *N Engl J Med***380**, 1450-1462 (2019).
4. Jemal, A. *et al.* Annual Report to the Nation on the Status of Cancer, 1975-2014, Featuring Survival. *J Natl Cancer Inst***109**, (2017).
5. Jeannot, E. *et al.* Increased incidence of aflatoxin B1-induced liver tumors in hepatitis virus C transgenic mice. *Int J Cancer***130**, 1347-1356 (2012).
6. Motta, M., Giugno, I., Ruello, P., Pistone, G., Di Fazio, I. & Malaguarnera, M. Lipoprotein (a) behaviour in patients with hepatocellular carcinoma. *Minerva Med***92**, 301-305 (2001).
7. Basili, S., Andreozzi, P., Vieri, M., Maurelli, M., Cara, D., Cordova, C. & Alessandri, C. Lipoprotein (a) serum levels in patients with hepatocarcinoma. *Clin Chim Acta***262**, 53-60 (1997).
8. Younossi, Z. *et al.* Nonalcoholic Steatohepatitis Is the Fastest Growing Cause of Hepatocellular Carcinoma in Liver Transplant Candidates. *Clin Gastroenterol Hepatol***17**, 748-755.e3 (2019).
9. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell***144**, 646-674 (2011).
10. MEDES, G., THOMAS, A. & WEINHOUSE, S. Metabolism of neoplastic tissue. IV. A study of lipid synthesis in neoplastic tissue slices in vitro. *Cancer Res***13**, 27-29 (1953).
11. Menendez, J. A. & Lupu, R. Fatty acid synthase and the lipogenic phenotype in cancer pathogenesis. *Nat Rev Cancer***7**, 763-777 (2007).
12. Piccinin, E., Villani, G. & Moschetta, A. Metabolic aspects in NAFLD, NASH and hepatocellular carcinoma: the role of PGC1 coactivators. *Nat Rev Gastroenterol Hepatol***16**, 160-174 (2019).
13. Tseng, C. H. Metformin and risk of hepatocellular carcinoma in patients with type 2 diabetes. *Liver Int***38**, 2018-2027 (2018).
14. Chettouh, H., Lequoy, M., Fartoux, L., Vigouroux, C. & Desbois-Mouthon, C. Hyperinsulinaemia and insulin signalling in the pathogenesis and the clinical course of hepatocellular carcinoma. *Liver Int***35**, 2203-2217 (2015).
15. Pavlova, N. N. & Thompson, C. B. The Emerging Hallmarks of Cancer Metabolism. *Cell Metab***23**, 27-47 (2016).
16. Brault, C. & Schulze, A. The Role of Glucose and Lipid Metabolism in Growth and Survival of Cancer Cells. *Recent Results Cancer Res***207**, 1-22 (2016).
17. Li, N., Li, L. & Chen, Y. The Identification of Core Gene Expression Signature in Hepatocellular Carcinoma. *Oxid Med Cell Longev***2018**, 3478305 (2018).
18. Chang, Y. S. *et al.* ACSL3 and GSK-3β are essential for lipid upregulation induced by endoplasmic reticulum stress in liver cells. *J Cell Biochem***112**, 881-893 (2011).
19. Migita, T. *et al.* ACSL3 promotes intratumoral steroidogenesis in prostate cancer cells. *Cancer Sci***108**, 2011-2021 (2017).

20. Ndiaye, H., Liu, J. Y., Hall, A., Minogue, S., Morgan, M. Y. & Waugh, M. G. Immunohistochemical staining reveals differential expression of ACSL3 and ACSL4 in hepatocellular carcinoma and hepatic gastrointestinal metastases. *Biosci Rep***40**, (2020).
21. Huang, L. S. *et al.* The mitochondrial cardiolipin remodeling enzyme lysocardiolipin acyltransferase is a novel target in pulmonary fibrosis. *Am J Respir Crit Care Med***189**, 1402-1415 (2014).
22. Xia, Y., Hong, H., Ye, L., Wang, Y., Chen, H. & Liu, J. Label-free quantitative proteomic analysis of right ventricular remodeling in infant Tetralogy of Fallot patients. *J Proteomics***84**, 78-91 (2013).
23. Bi, J. *et al.* Oncogene Amplification in Growth Factor Signaling Pathways Renders Cancers Dependent on Membrane Lipid Remodeling. *Cell Metab***30**, 525-538.e8 (2019).
24. Wang, B. & Tontonoz, P. Phospholipid Remodeling in Physiology and Disease. *Annu Rev Physiol***81**, 165-188 (2019).
25. Mansilla, F. *et al.* Lysophosphatidylcholine acyltransferase 1 (LPCAT1) overexpression in human colorectal cancer. *J Mol Med (Berl)***87**, 85-97 (2009).
26. Du, Y. *et al.* Lysophosphatidylcholine acyltransferase 1 upregulation and concomitant phospholipid alterations in clear cell renal cell carcinoma. *J Exp Clin Cancer Res***36**, 66 (2017).
27. Lebok, P. *et al.* Up-regulation of lysophosphatidylcholine acyltransferase 1 (LPCAT1) is linked to poor prognosis in breast cancer. *Aging (Albany NY)***11**, 7796-7804 (2019).
28. Grupp, K. *et al.* High lysophosphatidylcholine acyltransferase 1 expression independently predicts high risk for biochemical recurrence in prostate cancers. *Mol Oncol***7**, 1001-1011 (2013).
29. Wei, C. *et al.* LPCAT1 promotes brain metastasis of lung adenocarcinoma by up-regulating PI3K/AKT/MYC pathway. *J Exp Clin Cancer Res***38**, 95 (2019).
30. Morita, Y. *et al.* Lysophosphatidylcholine acyltransferase 1 altered phospholipid composition and regulated hepatoma progression. *J Hepatol***59**, 292-299 (2013).
31. Guo, Z. *et al.* CDC91L1 (PIG-U) is a newly discovered oncogene in human bladder cancer. *Nat Med***10**, 374-381 (2004).
32. Sarathi, A. & Palaniappan, A. Novel significant stage-specific differentially expressed genes in hepatocellular carcinoma. *BMC Cancer***19**, 663 (2019).
33. Cao, J., Wang, P., Chen, J. & He, X. PIGU overexpression adds value to TNM staging in the prognostic stratification of patients with hepatocellular carcinoma. *Hum Pathol***83**, 90-99 (2019).
34. Saleheen, D. *et al.* Human knockouts and phenotypic analysis in a cohort with a high rate of consanguinity. *Nature***544**, 235-239 (2017).
35. Gregson, J. M. *et al.* Genetic invalidation of Lp-PLA2 as a therapeutic target: Large-scale study of five functional Lp-PLA2-lowering alleles. *Eur J Prev Cardiol***24**, 492-504 (2017).
36. Esenwa, C. C. & Elkind, M. S. Inflammatory risk factors, biomarkers and associated therapy in ischaemic stroke. *Nat Rev Neurol***12**, 594-604 (2016).
37. Musunuru, K. & Kathiresan, S. Surprises From Genetic Analyses of Lipid Risk Factors for Atherosclerosis. *Circ Res***118**, 579-585 (2016).
38. Casas, J. P. *et al.* PLA2G7 genotype, lipoprotein-associated phospholipase A2 activity, and coronary heart disease risk in 10 494 cases and 15 624 controls of European Ancestry. *Circulation***121**, 2284-2293 (2010).
39. Chen, J., Li, L., Zhou, Z., Yu, S., Li, Y. & Gao, Y. FAPP2 promotes tumor cell growth in human colon cancer through activation of Wnt signaling. *Exp Cell Res***374**, 12-18 (2019).
40. Sun, C., Chen, P., Chen, Q., Sun, L., Kang, X., Qin, X. & Liu, Y. Serum paraoxonase 1 heteroplasmon, a fucosylated, and sialylated glycoprotein in distinguishing early hepatocellular carcinoma from liver cirrhosis patients. *Acta*

*Biochim Biophys Sin (Shanghai)***44**, 765-773 (2012).

41. Ding, G. Y. *et al.* Serum PON1 as a biomarker for the estimation of microvascular invasion in hepatocellular carcinoma. *Ann Transl Med***8**, 204 (2020).
42. Yu, W. M. *et al.* Metabolic regulation by the mitochondrial phosphatase PTPMT1 is required for hematopoietic stem cell differentiation. *Cell Stem Cell***12**, 62-74 (2013).
43. Zhang, J. *et al.* Mitochondrial phosphatase PTPMT1 is essential for cardiolipin biosynthesis. *Cell Metab***13**, 690-700 (2011).
44. Pagliarini, D. J. *et al.* Involvement of a mitochondrial phosphatase in the regulation of ATP production and insulin secretion in pancreatic beta cells. *Mol Cell***19**, 197-207 (2005).
45. Niemi, N. M., Lanning, N. J., Westrate, L. M. & MacKeigan, J. P. Downregulation of the mitochondrial phosphatase PTPMT1 is sufficient to promote cancer cell death. *PLoS One***8**, e53803 (2013).
46. Sheng, J. *et al.* SRSF1 modulates PTPMT1 alternative splicing to regulate lung cancer cell radioresistance. *EBioMedicine***38**, 113-126 (2018).
47. Zheng, Z., Li, X., You, H., Zheng, X. & Ruan, X. LncRNA SOCS2-AS1 inhibits progression and metastasis of colorectal cancer through stabilizing SOCS2 and sponging miR-1264. *Aging (Albany NY)***12**, (2020).
48. Liu, Y., Yang, Y., Du, J., Lin, D. & Li, F. MiR-3613-3p from carcinoma-associated fibroblasts exosomes promoted breast cancer cell proliferation and metastasis by regulating SOCS2 expression. *IUBMB Life* (2020).
49. Kim, J. H. *et al.* Alterations in the p53-SOCS2 axis contribute to tumor growth in colon cancer. *Exp Mol Med***50**, 3 (2018).
50. Val, C. H. *et al.* SOCS2 modulates adipose tissue inflammation and expansion in mice. *J Nutr Biochem***76**, 108304 (2020).
51. Monti-Rocha, R. *et al.* SOCS2 Is Critical for the Balancing of Immune Response and Oxidate Stress Protecting Against Acetaminophen-Induced Acute Liver Injury. *Front Immunol***9**, 3134 (2018).
52. Chen, M. *et al.* RNA N6-methyladenosine methyltransferase-like 3 promotes liver cancer progression through YTHDF2-dependent posttranscriptional silencing of SOCS2. *Hepatology***67**, 2254-2270 (2018).
53. Ren, W., Wu, S., Wu, Y., Liu, T., Zhao, X. & Li, Y. MicroRNA-196a/-196b regulate the progression of hepatocellular carcinoma through modulating the JAK/STAT pathway via targeting SOCS2. *Cell Death Dis***10**, 333 (2019).
54. Zhang, T. *et al.* TBL1XR1 is involved in c-Met-mediated tumorigenesis of human nonsmall cell lung cancer. *Cancer Gene Ther***27**, 136-146 (2020).
55. Liao, L. M., Zhang, F. H., Yao, G. J., Ai, S. F., Zheng, M. & Huang, L. Role of Long Noncoding RNA 799 in the Metastasis of Cervical Cancer through Upregulation of TBL1XR1 Expression. *Mol Ther Nucleic Acids***13**, 580-589 (2018).
56. Wu, X. *et al.* Nuclear TBLR1 as an ER corepressor promotes cell proliferation, migration and invasion in breast and ovarian cancer. *Am J Cancer Res***6**, 2351-2360 (2016).
57. Zhou, Q. *et al.* Transducin (β)-like 1 X-linked receptor 1 promotes gastric cancer progression via the ERK1/2 pathway. *Oncogene***36**, 1873-1886 (2017).
58. Chen, S. P. *et al.* Transducin β-like 1 X-linked receptor 1 suppresses cisplatin sensitivity in nasopharyngeal carcinoma via activation of NF-κB pathway. *Mol Cancer***13**, 195 (2014).
59. Guo, Y. *et al.* Theranostical nanosystem-mediated identification of an oncogene and highly effective therapy in hepatocellular carcinoma. *Hepatology***63**, 1240-1255 (2016).

Figures

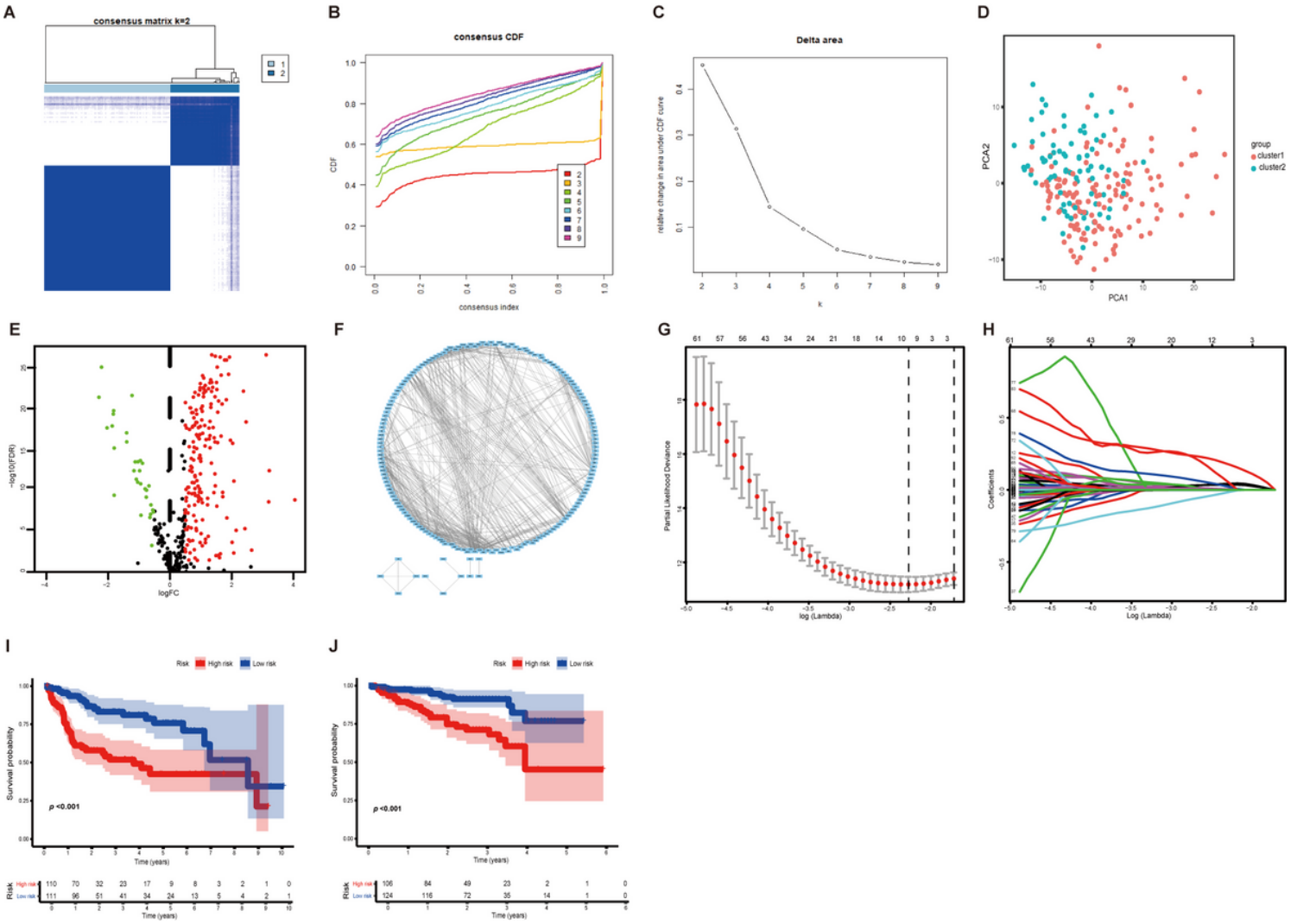

Figure 1

Subtype of HCC based on the lipid metabolism-related gene set

A, consensus clustering matrix of 221 HCC samples for K=2.

B, consensus clustering for K=2 to K=10;

C, Relative change in area under CDF curve according to various k values;

D, survival analysis of HCC patients in cluster1 and cluster2;

E, volcano plot of DEGs, red dots represent upregulated genes and green dots represent downregulated genes;

F, protein-protein interaction (PPI) network;

G, “leave-one-out-cross-validation” for parameter selection in LASSO Cox proportional hazards regression;

H, LASSO coefficient profiles of the ten robust prognostic genes; I-J, Kaplan-Meier survival curve for high and low risk cohort in the training set and validation set.

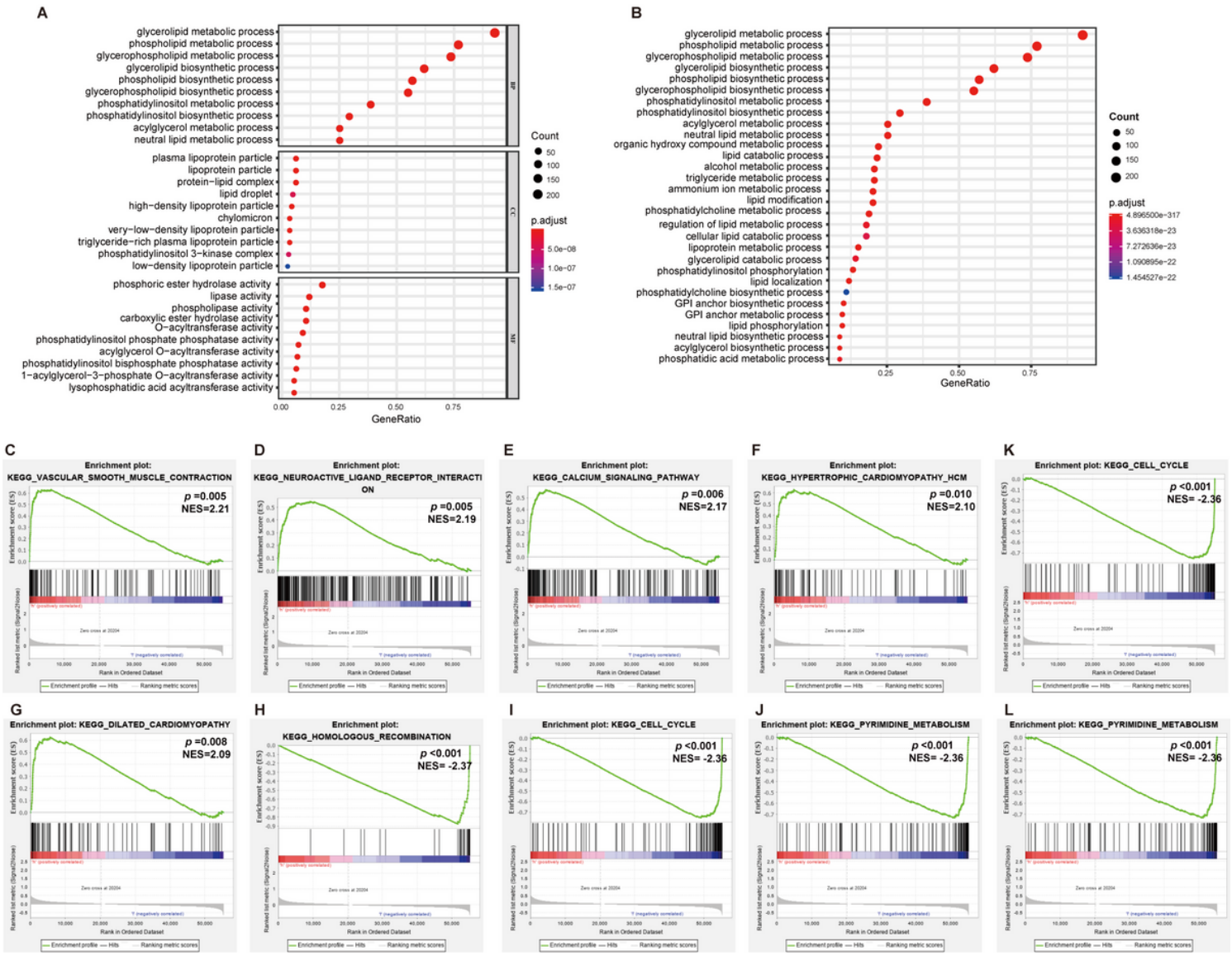

Figure 2

A-B, GO and KEGG analysis of DEGs;

C-L, functional enrichment analysis through GSEA.

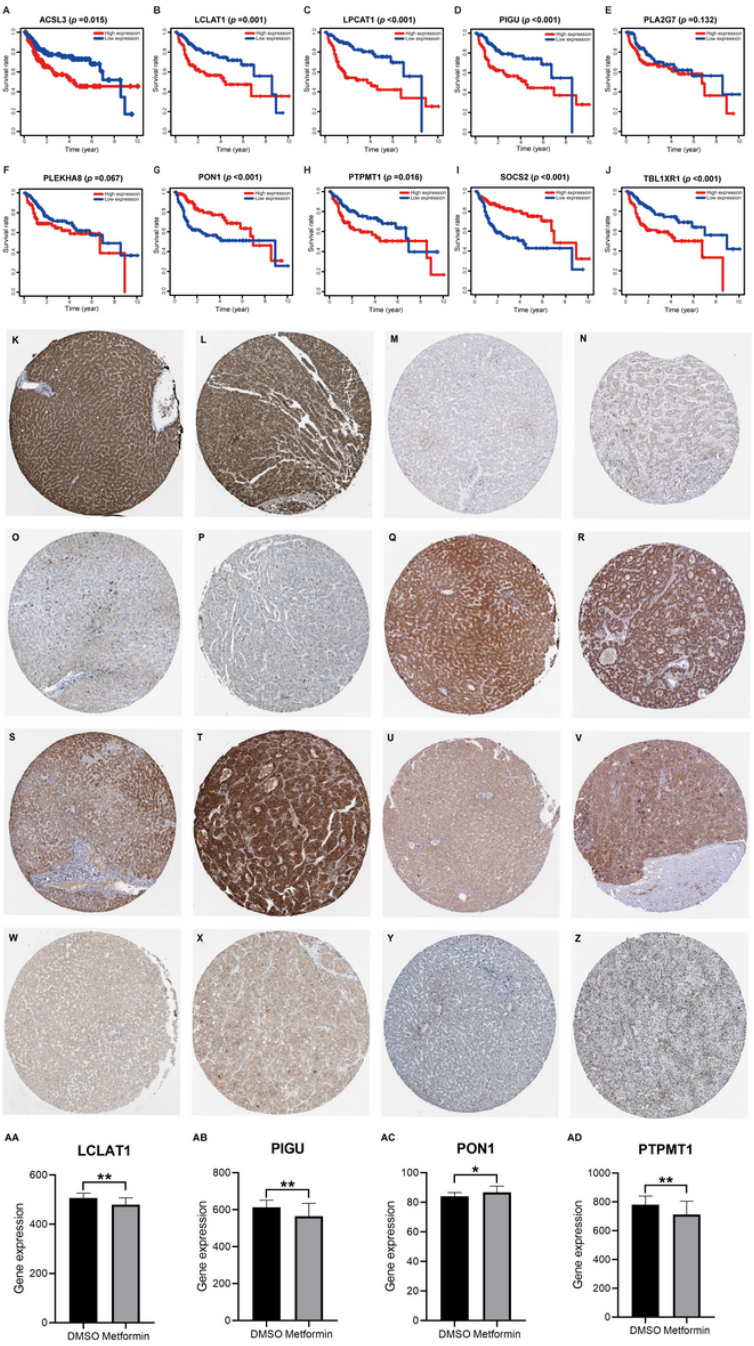

Figure 3

A-J, Kaplan-Meier analysis about the relevance between ten risk genes expression and survival rate of HCC patients in training set;

K-Z, Immunohistochemistry of eight proteins expression between cancer and adjacent tissues. K, ACSL3 normal; L, ACSL3 tumor; M, LCLAT1 normal; N, LCLAT1 tumor; O, LPCAT1 normal; P, LPCAT1 tumor; Q, PIGU normal; R, PIGU tumor; S, PON1 normal; T, PON1 tumor; U, PTPMT1 normal; V, PTPMT1 tumor; W, SOCS2 normal; X, SOCS2 tumor; W, TBL1XR1 normal; Z, TBL1XR1 tumor;

AA-AD, The difference between four genes expression and metformin treatment presented statistically significance. * P<0.05; **P<0.01.

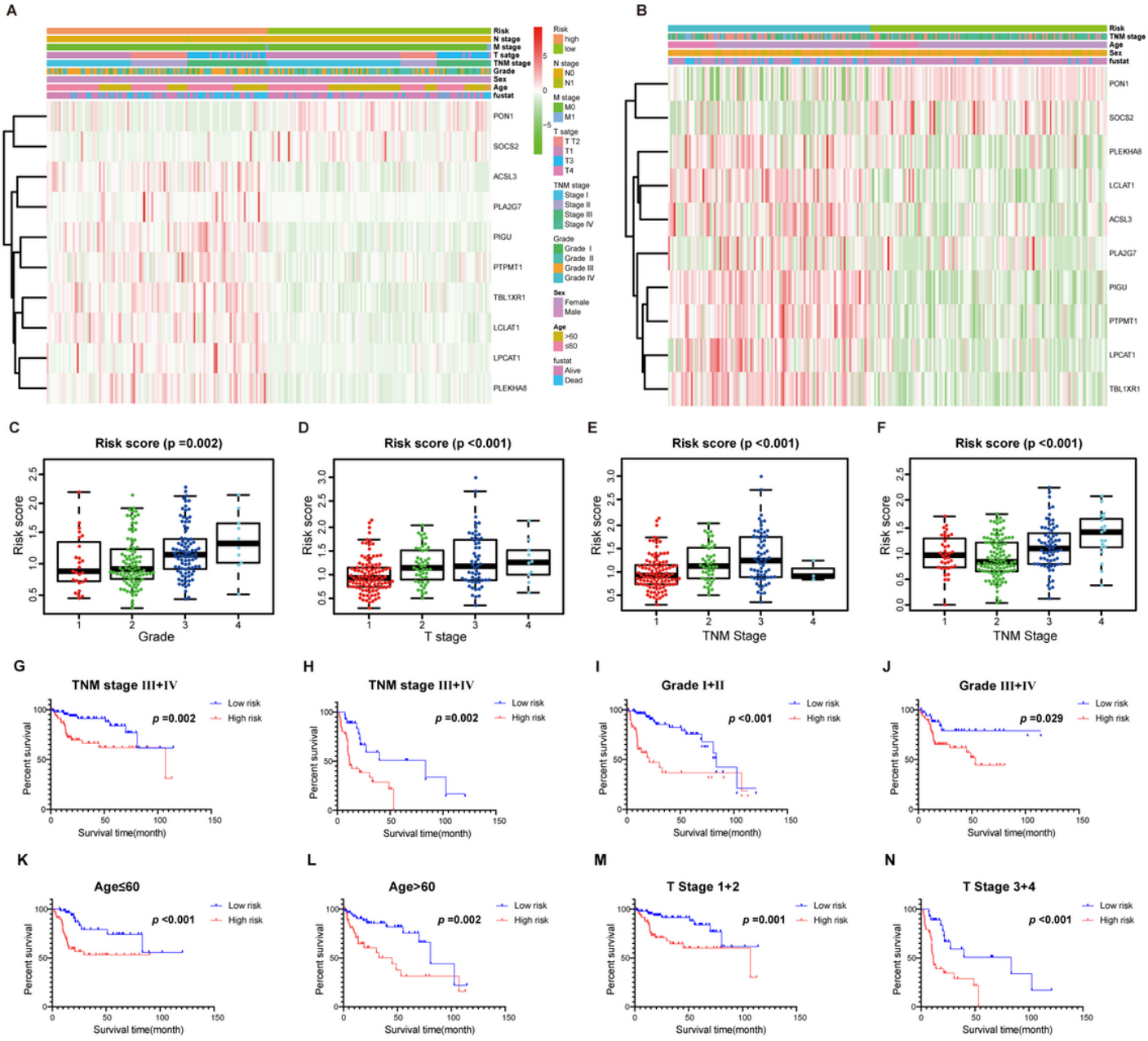

Figure 4

Correlation between risk score and clinical characteristics

A-B, heat map of the association of risk scores and clinicopathological features in training set and validation set;

C-E, box plot of the association of risk scores and tumor grade, T stage and TNM stage in training set;

F, box plot of the association of risk scores and TNM stage in validation set;

G-N, Layered comparison for the same conditions of age, tumor grade, T stage and TNM stage in training set.

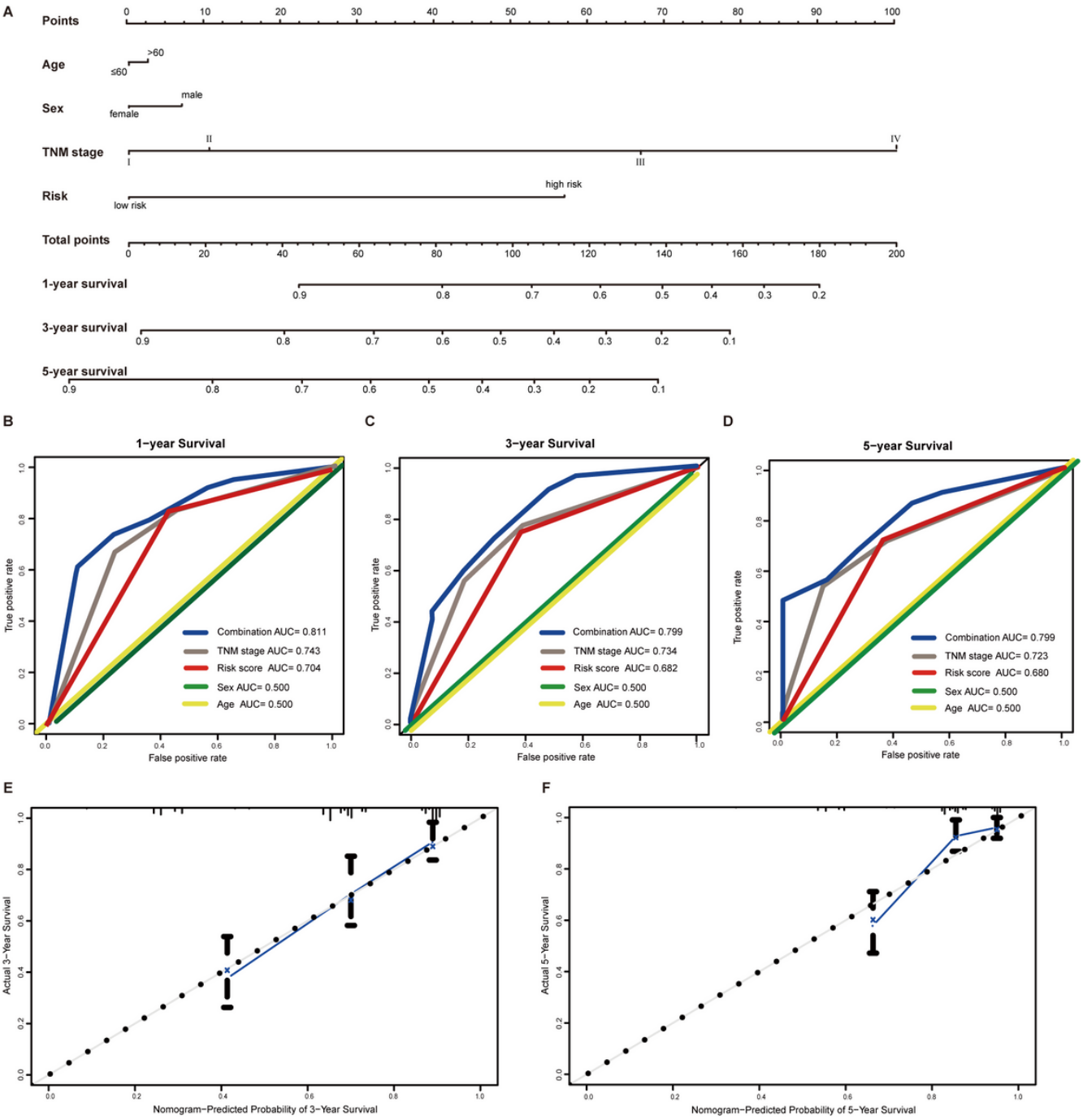

Figure 5

Construction and evaluation of ten lipid metabolism-related genes nomogram in training set

A, prognostic nomogram for the prediction of 1-, 3- and 5-years overall survival in HCC;

B-D, AUCs for the 1-, 3-, 5-year overall survival of HCC patients in training set, respectively;

E-F, Calibration curve of nomogram models for 3-, 5-year overall survival in training set.

Supplementary Files

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

- FigureSupplement1L.tif
- FigureSupplement2L.tif
- FigureSupplement3L.tif
- FigureSupplement4L.tif
- LegendsofFigureSupplement14.docx
- Tablesupplement1.docx