Genes & Diseases (2022) 9, 116–127

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: http://ees.elsevier.com/gendis/default.asp

FULL LENGTH ARTICLE

Transcriptomic profiling of peroxisome-related genes reveals a novel prognostic signature in hepatocellular carcinoma

Liewang Qiu [a], Ke Zhan [b], Kija Malale [b], Xiaoling Wu [b,**], Zhechuan Mei [b,*]

[a] *Department of Gastroenterology, Yongchuan Hospital of Chongqing Medical University, Chongqing 402160, PR China*
[b] *Department of Gastroenterology, The Second Affiliated Hospital of Chongqing Medical University, Chongqing 400010, PR China*

Received 17 February 2020; received in revised form 25 March 2020; accepted 13 April 2020
Available online 21 April 2020

KEYWORDS
Gene signature;
Hepatocellular carcinoma;
Nomogram;
Peroxisome;
Prognosis

Abstract Emerging evidence suggests that peroxisomes play a role in the regulation of tumorigenesis and cancer progression. However, the prognostic value of peroxisome-related genes has been rarely investigated. This study aimed to establish a peroxisome-related gene signature for overall survival (OS) prediction in patients with hepatocellular carcinoma (HCC). First, univariate Cox regression analysis was employed to identify prognostic peroxisome-related genes in The Cancer Genome Atlas liver cancer cohort, and least absolute shrinkage and selection operator Cox regression analysis was used to construct a 10-gene signature. The risk score based on the signature was positively correlated with poor prognosis (HR = 4.501, 95% CI = 3.021–6.705, P = 1.39e−13). Second, multivariate Cox regression incorporating additional characteristics revealed that the signature was an independent predictor. Time-dependent ROC curves demonstrated good performance of the signature in predicting the OS of HCC patients. The prognostic performance was validated using International Cancer Genome Consortium HCC cohort data. Gene set enrichment analysis revealed that the signature-related alterations in biological processes mainly involved peroxisomal functions. Finally, we developed a nomogram model based on the gene signature and TNM stage, which showed a superior prognostic power (C-index = 0.702). Thus, our study revealed a novel peroxisome-related gene signature that may help improve personalized OS prediction in HCC patients.

* Corresponding author. Department of Gastroenterology, The Second Affiliated Hospital of Chongqing Medical University, No.76 Linjiang Road, Chongqing 400010, PR China.
** Corresponding author.
E-mail addresses: wuxiaoling@cqmu.edu.cn (X. Wu), meizhechuan@cqmu.edu.cn (Z. Mei).
Peer review under responsibility of Chongqing Medical University.

https://doi.org/10.1016/j.gendis.2020.04.010

Introduction

Hepatocellular carcinoma (HCC) is the most frequent pathological type of primary liver cancer, representing the sixth most common neoplasm and fourth leading cause of cancer-related deaths worldwide.[1] Although advancements have been made in the clinical management of HCC in the last decade, the prognosis of HCC patients is dismal, and the 5-year survival rate is 18%.[1] Prognostic evaluation is crucial for disease surveillance and selecting treatment strategies for HCC management. Several staging systems considering clinical and pathological patient characteristics are recommended in HCC management guidelines.[2–4] However, these assessment approaches have various limitations in patient stratification and require improvement.[1,5] In recent years, next-generation sequencing data have considerably enhanced our understanding of cancer biology; several studies have demonstrated the prognostic abilities of gene signatures established from different aspects in HCC patients.[6–8] However, further efforts are required to identify novel and robust prognosis predictors through bioinformatic analyses of genome data combined with the clinical features of HCC patients.

Aberrant metabolism is a hallmark of cancer.[9] Metabolic reprogramming involving glucose, lipid, and amino acid metabolism, redox balance, and the tumor microenvironment facilitates the survival and malignancy of hepatoma cells.[10] Several inhibitors targeting critical enzymes involved in metabolic reprogramming have entered preclinical and clinical trials for cancer therapy.[11] Peroxisomes, predominantly found in the liver, are single-membrane-enclosed organelles that contain enzymes that catalyze the metabolism of very-long-chain and branched-chain fatty acids, ether phospholipids, reactive oxygen species, and bile acids.[12–14] Recent studies revealed that peroxisomes play an essential role in the development and progression of HCC.[15] Specifically, multiple peroxisomal enzymes and related metabolic activities are altered in HCC and other cancers.[16–20] Additionally, modulation of peroxisomal enzyme inhibitors and gene expression has been demonstrated to either inhibit or promote tumor growth.[16,17,20]

Studies on the role of peroxisomes in HCC have mainly focused on specific peroxisomal enzymes, such as AGPS, HAO2, and PRDX1[17,20,21], whereas peroxisomal function at the subcellular level has been rarely investigated. In this study, we identified a novel prognostic signature from peroxisome-related genes in HCC cohorts via univariate Cox and least absolute shrinkage and selection operator (LASSO) Cox regression analyses. Further, we confirmed the signature to be a robust, independent predictor for risk stratification in HCC patients, and we established a superior nomogram combining the gene signature and a clinical feature (TNM stage). We hypothesize that this novel signature may contribute to strategizing personalized cancer management and that the relevant peroxisome-related genes may become prospective therapeutic targets for HCC.

Materials and methods

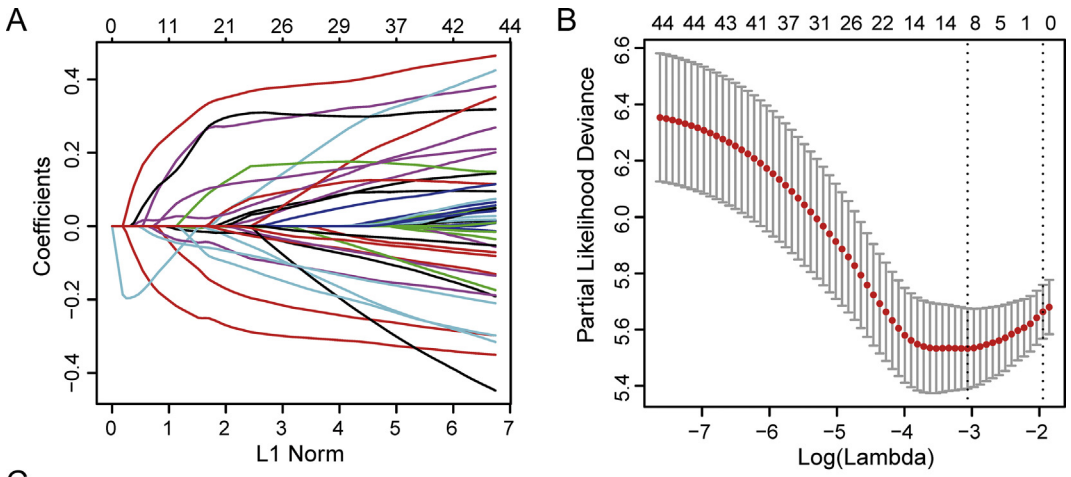
Collection of peroxisome-related gene sets

To generate a representative peroxisome-related gene list, we collected genetic data including a human peroxisomal gene set (100 genes) from the PeroxisomeDB 2.0 database,[22] a database that gathers and integrates curated information related to peroxisomes; a peroxisomal gene set (83 genes) from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database,[23] a well-known database for genes and pathways; and protein-coding genes from the proteome of human liver peroxisomes (60 genes).[24] In total, 113 non-redundant peroxisome-related genes were included in the analysis (Table S1).

Data acquisition of HCC cohorts

All HCC datasets from The Cancer Genome Atlas (TCGA), International Cancer Genome Consortium (ICGC), and Gene Expression Omnibus (GEO) databases were collected (up to November 20, 2019). The TCGA Liver Hepatocellular Carcinoma (TCGA-LIHC) and ICGC Liver Cancer - RIKEN, Japan (LIRI-JP) cohorts were selected as study cohorts because they contain complete survival information and all the peroxisome-related genes could be mapped in their gene expression profiles.

For the TCGA-LIHC cohort, gene expression profiles produced using the Illumina HiSeq RNA-Seq platform were downloaded from the TCGA database (https://portal.gdc.cancer.gov/repository/). Protein-coding gene symbols corresponding to Ensemble IDs were transformed according to the Homo_sapiens.GRCh38.98.chr.gtf file (http://asia.ensembl.org/info/data/ftp/index.html). HTSeq-count gene expression profiles were normalized using the variance stabilizing transformation (VST) function in DESeq2 of the R package (R Core Team, https://www.R-project.org).[25] Clinical information and somatic mutation data were extracted from the TCGA database and cBioPortal (http://cbioportal.org/),[26] respectively. We collected data from 374 liver cancer samples from 371 patients. Twenty-three samples, including three with recurrent tumor, seven with incomplete survival information, seven with intrahepatic cholangiocarcinoma (ICC) or combined hepatocellular cholangiocarcinoma (cHCC/CC), and six with inferior quality according to the TCGA sample quality annotations file (https://gdc.cancer.gov/about-data/publications/pancanatlas), were excluded from further



C

Gene symbol	Type	Coefficient	HR	95% CI	*P*-value
MPV17	Risky	0.0167	1.816	1.327–2.484	0.0002
AGPS	Risky	0.1011	1.718	1.202–2.455	0.003
LDHA	Risky	0.2231	1.601	1.229–2.087	0.0005
TRIM37	Risky	0.0115	1.454	1.063–1.989	0.0191
PRDX1	Risky	0.1103	1.367	1.079–1.731	0.0097
ACSL6	Protective	−0.0299	0.842	0.75–0.946	0.0038
PECR	Protective	−0.0154	0.737	0.629–0.863	0.0002
ACAT1	Protective	−0.1801	0.664	0.558–0.789	<0.0001
MTARC2	Protective	−0.1248	0.657	0.566–0.763	<0.0001
ATAD1	Protective	−0.0058	0.622	0.427–0.907	0.0136

Figure 1 Development of the 10-gene signature based on data in the training cohort. (**A**) Ten-time cross-validation for tuning parameter screening in the LASSO Cox penalized regression model. (**B**) LASSO coefficient profiles of 47 peroxisome-related genes. (**C**) Forest plot of 10 signature genes selected by LASSO Cox regression. HR, hazard ratio; CI, confidence interval.

analysis. The remaining 351 HCC samples, from 351 patients, were designated as the training cohort.

For the ICGC LIRI-JP cohort, gene expression profiles, gene somatic mutation data, and clinical information were downloaded from the ICGC database (https://dcc.icgc.org/releases/release_28/).[27]

Raw RNA-Seq count data with annotated gene symbols were also normalized using VST. We collected data from 243 liver cancer samples from 232 patients. Forty-seven samples, including three with metastatic tumor, 30 with ICC or cHCC/CC, and 14 duplicated samples with low tumor cell percentage, were excluded from further analysis. The remaining 196 HCC samples, from 196 patients, were designated as the validation cohort.

Generation of a gene signature

To find prognostic peroxisome-related genes, we performed an univariate Cox proportional hazards regression analysis using the R package survival (https://cran.r-project.org/web/packages/survival/index.html) and we identified the genes significantly associated with overall survival (OS) in the training cohort. $P < 0.05$ was considered statistically significant in the Wald test. Hazard ratios (HRs) and 95% confident intervals (95% CIs) were calculated. A covariate with HR > 1 indicated a positive association with event hazards and a negative association with survival time. Least absolute shrinkage and selection operator (LASSO) penalized Cox regression was utilized to construct an optimal risk signature from the survival-associated genes using the R package glmnet (https://cran.r-project.org/web/packages/glmnet/index.html). The risk score for each patient in both cohorts was calculated by taking the sum of the LASSO regression coefficient for each signature gene multiplied with its corresponding expression value. Patients were subsequently median-dichotomized into high- and low-risk groups based on the risk scores in each cohort.

Table 1 Functions of genes in the prognostic gene signature.

No.	Gene Symbol	Full name	Function	Risk coefficient
1	*MPV17*	Mitochondrial Inner Membrane Protein MPV17	Regulation of reactive oxygen species metabolism	0.0167
2	*AGPS*	Alkylglycerone Phosphate Synthase	Participates in the pathway of ether lipid biosynthesis	0.1011
3	*LDHA*	Lactate Dehydrogenase A	Catalyzes conversion of pyruvate to lactate under anaerobic conditions	0.2231
4	*TRIM37*	Tripartite Motif Containing 37	E3 ubiquitin-protein ligase	0.0115
5	*PRDX1*	Peroxiredoxin 1	Thiol-specific peroxidase	0.1103
6	*ACSL6*	Acyl-CoA Synthetase Long Chain Family Member 6	Catalyzes the formation of acyl-CoA from fatty acids, ATP, and CoA	−0.0299
7	*PECR*	Peroxisomal Trans-2-Enoyl-CoA Reductase	Participates in chain elongation of fatty acids	−0.0154
8	*ACAT1*	Acetyl-CoA Acetyltransferase 1	Enzyme that catalyzes the reversible formation of acetoacetyl-CoA	−0.1801
9	*MTARC2*	Mitochondrial Amidoxime Reducing Component 2	Catalyzes the reduction of N-oxygenated molecules	−0.1248
10	*ATAD1*	ATPase Family AAA Domain Containing 1	AAA^{+}-protein involved in peroxisome biogenesis and function	−0.0058

Evaluation and validation of the gene signature

Schoenfeld residuals were calculated to validate the proportional hazards (PH) assumption of the multigene signature Cox model. PH assumption was satisfied if both individual and global *P*-values were >0.05. Kaplan–Meier analysis and the log-rank test were employed to estimate and visualize OS distributions, using R packages survival and survminer (https://cran.r-project.org/web/packages/survminer/index.html). $P < 0.05$ was considered statistically significant in the log-rank test. Multivariate Cox regression analysis was performed to evaluate the independence of the signature from additional clinical features and gene somatic mutation status. To assess the predictive performance of the signature, we conducted time-dependent receiver-operating characteristic (ROC) analysis and calculated the area under the curve (AUC) values using the R package survcomp (http://www.bioconductor.org/packages/release/bioc/html/survcomp.html).

Gene set enrichment analysis (GSEA)

To investigate alterations in the molecular signaling pathways the signature genes are involved in, we conducted GSEA using the officially recommended software (GSEA v4.0.3)[28] to analyze gene enrichment in the datasets. Risk scores based on the gene signature were used to classify the patients into high-risk and low-risk groups. The gene sets used in our study (c2.cp.kegg.v7.0.symbols.gmt) were downloaded from the Molecular Signatures Database (MSigDB, http://software.broadinstitute.org/gsea/msigdb/index.jsp). Nominal $P < 0.05$ and false discovery rate (FDR) < 0.25 were set as thresholds to determine the statistical significance of the normalized enrichment score (NES).

Construction and evaluation of the nomogram

The risk scores and characteristics significant in the univariate Cox analysis were selected to construct a nomogram using the R packages survival and rms (https://cran.r-project.org/web/packages/rms/index.html). The concordance index (C-index) was calculated to assess the performance of the prediction model. Calibration plots were drawn to evaluate the concordance between actual and predicted survival. All statistical analyses performed were two-tailed, and $P < 0.05$ was considered statistically significant.

Results

Baseline characteristics of HCC patients

Data from 547 HCC patients from the TCGA-LIHC cohort ($n = 351$) and the ICGC LIRI-JP cohort ($n = 196$) were examined in this study. Detailed baseline characteristics of the patients in both cohorts are listed in Table S2. The component ratios of clinical features and gene mutation status differed significantly ($P < 0.05$) between the two independent cohorts.

Development of a peroxisome-related gene signature from the training cohort

We identified 47 genes significantly associated with the OS of HCC patients (Fig. S1) based on univariate Cox regression analysis of 113 peroxisome-related genes in the TCGA-LIHC (training) cohort. Forty genes with HRs <1 were considered as protective genes, whereas 7 genes with HRs >1 were considered as risky genes.

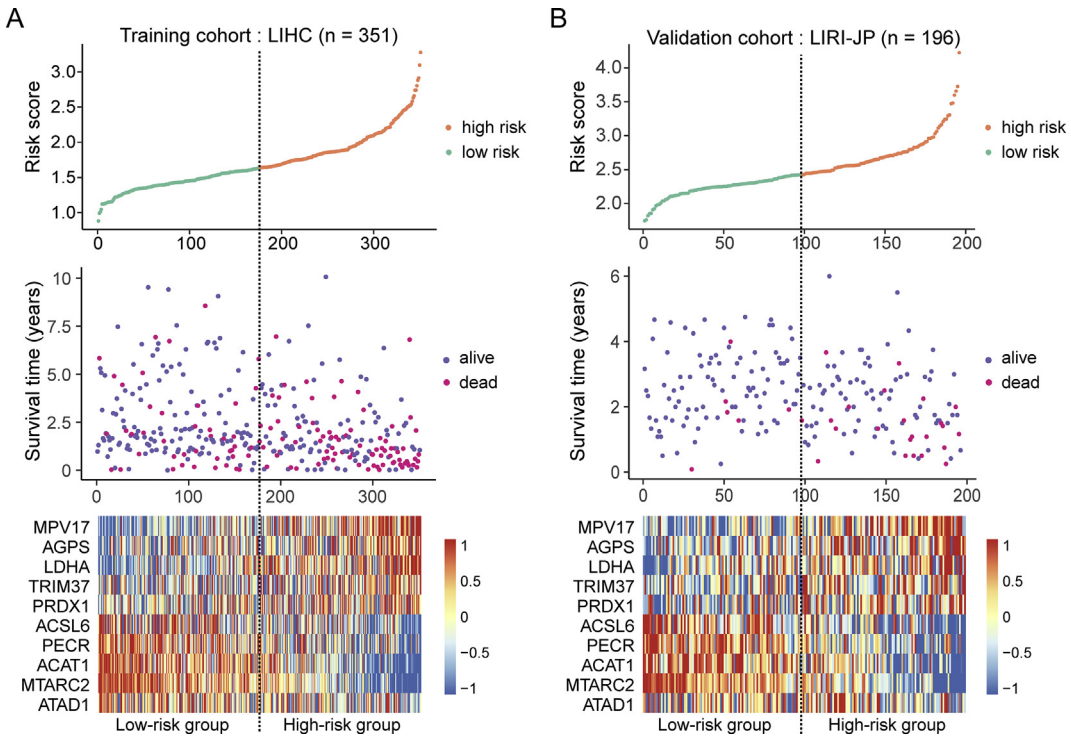


Figure 2 Distributions of risk scores, OS, and survival status, and heatmaps of gene expression profiles of signature genes in the training (A) and validation (B) cohorts. The dotted line indicates patients were median-dichotomized into the low-risk group and high-risk group.

To select optimal predictive genes, we applied LASSO Cox regression to 47 prognostic genes and thus identified 10 genes, i.e., *ACAT1*, *AGPS*, *ATAD1*, *LDHA*, *MTARC2*, *PECR*, *ASCL6*, *MPV17*, *PRDX1*, and *TRIM37*, with nonzero LASSO Cox regression coefficients (Fig. 1). All individual and global Schoenfeld test *P*-values were >0.05 (Fig. S2), indicating no violation of the PH assumption; therefore, the gene signature was established using these 10 genes. The functions and coefficients of these 10 genes were shown in Table 1, which mainly involved with a variety of metabolic enzymes. The risk scores for each patient were computed using the following formula: risk score = (−0.1801 × *ACAT1* expression) + (0.1011 × *AGPS* expression) + (−0.0058 × *ATAD1* expression) + (0.2231 × *LDHA* expression) + (−0.1248 × *MTARC2* expression) + (−0.0154 × *PECR* expression) + (−0.0299 × *ACSL6* expression) + (0.0167 × *MPV17* expression) + (0.1103 × *PRDX1* expression) + (0.0115 × *TRIM37* expression). High- and low-risk groups were classified according to the median cutoff in each cohort.

The distributions of the signature-based risk scores, OS status, survival time, and gene expression profiles for the training and validation cohorts are plotted in Fig. 2. The heatmaps demonstrate that the risky genes *MPV17*, *AGPS*, *LDHA*, *TRIM37*, and *PRDX1* exhibit higher expression in the high-risk group, whereas the protective genes *ASCL6*, *PECR*, *ACAT1*, *MTARC2*, and *ATAD1* exhibit higher expression in the low-risk group.

Evaluation and validation of the 10-gene signature in HCC cohorts

We evaluated the OS predictive ability of the 10-gene signature in the training and validation cohorts. The OS rates of patients in the high-risk group were significantly lower in the training ($P < 0.0001$) and validation cohorts ($P = 0.00038$) (Fig. 3A, C).

To evaluate the predictive ability of the signature further, we conducted time-dependent ROC curve analysis for OS at different time points. The AUC values for 1-, 3-, and 5-year OS were 0.715, 0.704, and 0.691, respectively, for the training cohort (Fig. 3B) and 0.759, 0.738, and 0.724, respectively, for the validation cohort (Fig. 3D).

Univariate Cox regression analysis confirmed that the risk score was positively associated with poor prognosis of patients in the training (HR = 4.501, 95% CI = 3.021—6.705, $P = 1.39e{-}13$) and validation cohorts (HR = 6.572, 95% CI = 3.026—14.274, $P = 1.95e{-}06$) (Fig. 4). In addition, the TNM stage was significantly prognostic in both cohorts, whereas tumor size, hepatic vein invasion, and TP53 mutation status were predictive only in the validation cohort.

Next, we performed multivariate Cox regression analysis adjusted for additional characteristics significant in the univariate Cox regression analysis to examine the independence of the 10-gene signature (Fig. 4). The signature retained highly prognostic value after adjustment in both HCC cohorts (TCGA-LIHC: HR = 3.808, 95%

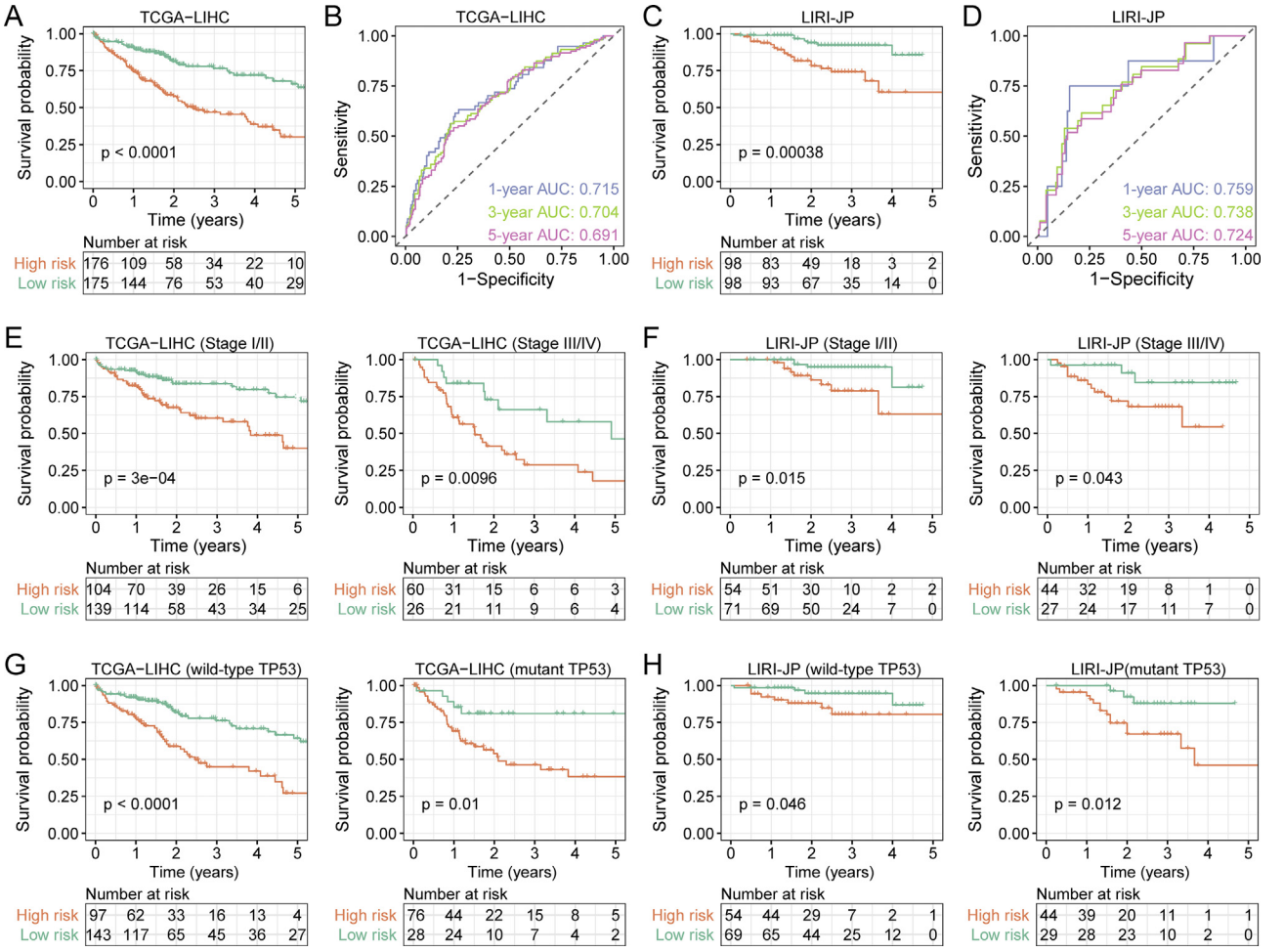


Figure 3 Prognostic performance of the 10-gene signature in HCC cohorts. (A) Kaplan—Meier plot of OS in the training cohort. (B) Time-dependent ROC curves with calculated AUCs at 1-, 3-, and 5-year OS based on the gene signature in the training cohort. (C) Kaplan—Meier plot of OS in the validation cohorts. (D) Time-dependent ROC curves with calculated AUCs at 1-, 3-, and 5-year OS based on the gene signature in the validation cohorts. (E, F) Kaplan—Meier plots of OS in subgroups with different tumor stages in the training (E) and validation (F) cohorts. (G, H) Kaplan—Meier plots of OS in subgroups with different TP53 mutation status in the training (G) and validation (H) cohorts.

CI = 2.461—5.892, P = 1.92e—09; ICGC LIRI-JP: HR = 4.567, 95% CI = 1.822—11.449, P = 0.001). Kaplan—Meier analysis of OS according to TNM stage and TP53 mutation status revealed that the signature successfully discriminated OS rates in the different subgroups in both the training and validation cohorts (all log-rank test, $P < 0.05$) (Fig. 3E—H).

The 10-gene signature was also compared with six recently reported multigene biomarkers for predicting the OS of HCC patients.[29—34] Risk scores were calculated in accordance with the reports. Time-dependent ROC curve analysis revealed that our signature performed well in predicting the prognosis across distinct datasets (Fig. 5), indicating its robustness.

Altered KEGG pathways in high- and low-risk subgroups

We conducted GSEA comparing the high- and low-risk groups in each HCC cohort to determine the biological pathway alterations underlying the signature. In total, 23 and 16 KEGG pathways were found to be enriched in the TCGA-LIHC and ICGC LIRI-JP cohorts, respectively (Table S3). Commonly enriched KEGG pathways in both cohorts are shown in Fig. 6. No significantly enriched pathway was identified in the high-risk group for both cohorts. Enriched KEGG pathways in the low-risk group were mainly related to peroxisome, peroxisome proliferator-activated receptor (PPAR) signaling, and other metabolic pathways, including primary bile acid biosynthesis, fatty acid metabolism, drug metabolism cytochrome p450, and nine amino acid metabolic pathways.

Establishment of a nomogram based on the peroxisome-related gene signature

To quantitatively predict the prognosis of HCC patients, we developed a nomogram integrating the risk score and the independent clinical risk factor (TNM stage) using TCGA-LIHC cohort data (Fig. 7A). Based on multivariate Cox

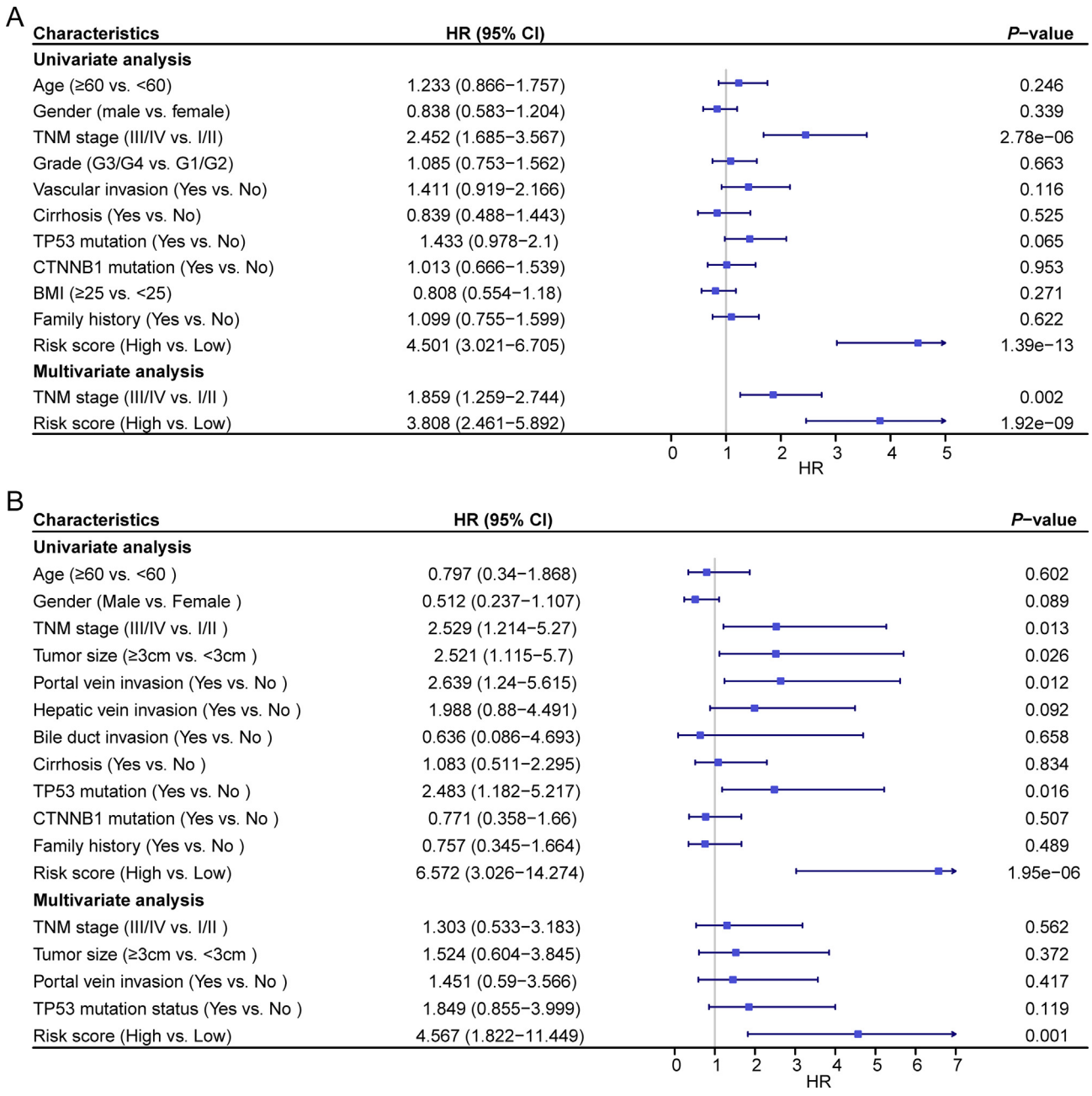


Figure 4 Univariate and multivariate Cox regression analyses of the relation between the signature and clinical characteristics in the TCGA-LIHC (A) and the LIRI-JP (B) cohorts.

analysis, each variate was assigned scaled points in the nomogram. We drew two horizontal straight lines to dispense the points for risk score and TNM stage, respectively. The total points for each patient were then computed by taking the sum of all variate points. The predicted survival probabilities at 1, 2, and 3 years were obtained by drawing a vertical line between the total point line and each prognostic line. The calibration curves indicated that the predicted and actual survival had a good consistency (Fig. 7B). The nomogram indicated that the risk score had a higher weight than the TNM stage. The C-index was 0.611, 0.693, and 0.702 for the TNM stage, risk score, and nomogram model, respectively. The nomogram was validated in the ICGC LIRI-JP cohort, and the 1-, 2-, and 3-year calibration curves are presented in Fig. 7C. Taken together, these results suggest that the developed nomogram is an optimal model for predicting the prognosis of HCC patients in comparison with individual risk factors.

Discussion

Advances in multi-omics technologies have greatly improved our understanding of cancer development and

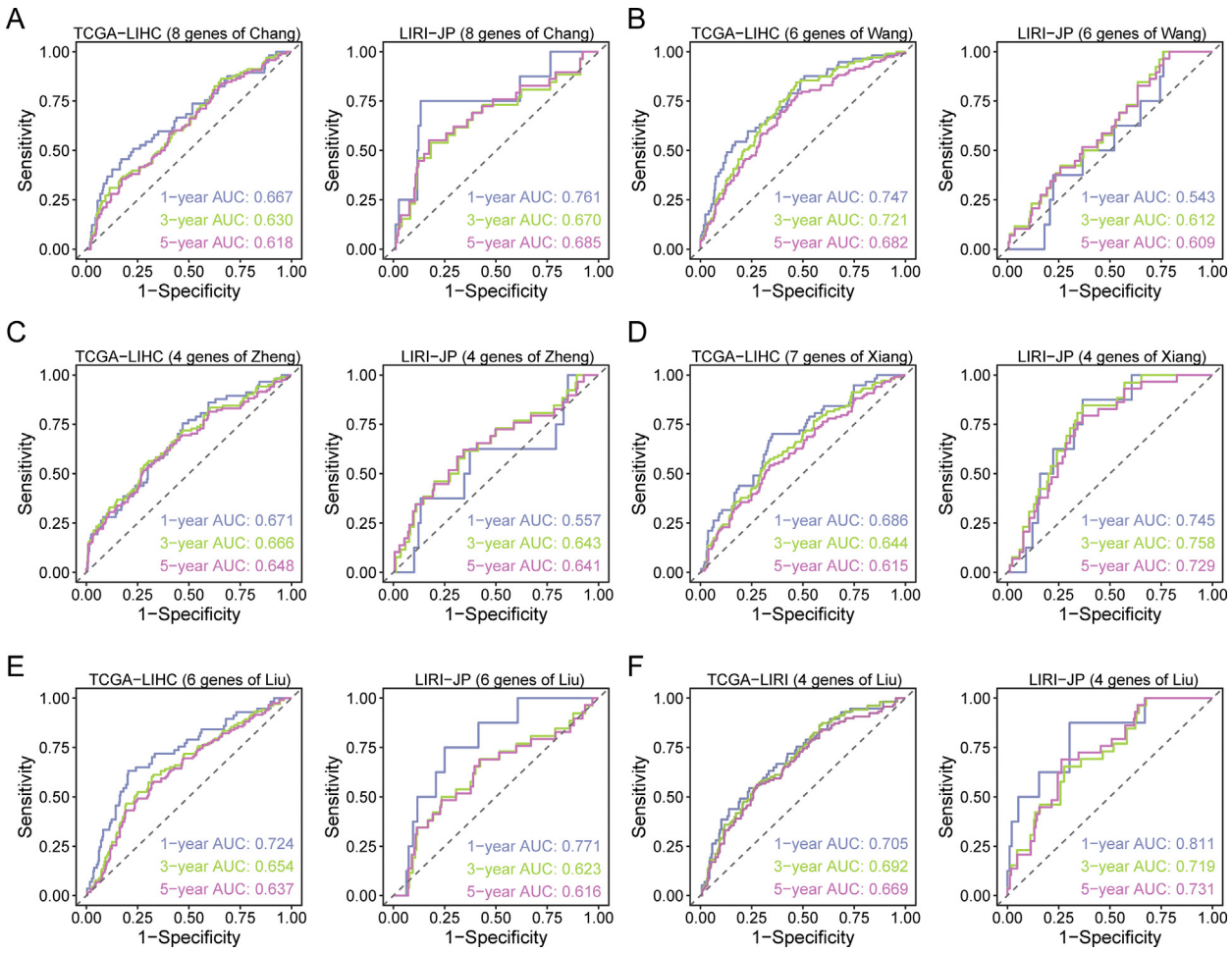


Figure 5 Time-dependent ROC curves with calculated AUCs at 1-, 3-, and 5-year OS of six other gene signatures reported in previous studies in the TCGA-LIHC and LIRI-JP cohorts.

progression and provided new promising approaches in cancer prevention, diagnosis, and therapy.[35,36] In this study, we identified a novel peroxisome-related gene signature that independently predicts OS in HCC cohorts. It improved the predictive ability of the TNM stage and successfully differentiated the OS of patients in similar stages. In addition, we developed a nomogram based on the 10-gene signature that demonstrated enhanced predictive ability in prognosis prediction.

The 10 genes included in our signature encode enzymes that are either predominantly or partially localized in the peroxisomes.[24] *MPV17*, *AGPS*, *LDHA*, *TRIM37*, and *PRDX1* were positively associated with poor survival, whereas *ASCL6*, *PECR*, *ACAT1*, *MTARC2*, and *ATAD1* appeared to be protective genes in the TCGA-LIHC cohort. *AGPS*, which encodes a critical peroxisomal enzyme involved in ether lipid synthesis, has been reported as an oncogene in multiple cancers.[17] AGPS inhibitor 1a lowered ether lipid levels and inhibit the survival and migration in several types of cancer cells.[37] As a new peroxisomal constituent in the human liver identified by proteomics,[24] LDHA catalyzes the interconversion of pyruvate to lactate and promotes various malignant features, including proliferation, metastasis, and immune escape.[21,38,39] *TRIM37*, encoding an E3 ubiquitin ligase that is partially located in peroxisomes, also promotes malignancy in multiple cancers.[40–42] *PRDX1* is an antioxidant enzyme of the peroxiredoxin family and reportedly functions as an oncoprotein in various types of solid tumors.[18,43,44] ASCL6 belongs to long-chain acyl-CoA synthetases participating in lipid metabolism. High ASCL6 levels predicted a better prognosis in acute myeloid leukemia, suggesting that *ASCL6* is a potential protective gene.[45] Based on their respective performances in the HCC cohorts, *AGPS*, *LDHA*, *TRIM37*, *PRDX1*, and *ASCL6* showed similar risk-elevating or protective functions in previous studies. By contrast, *ACAT1* is considered an oncogene, as ACAT1 inhibitors impaired tumor progression.[46–48] Interestingly, we found that *ACAT1* expression was significantly reduced in the high-risk group and correlated with a better OS; however, further investigation is needed to validate this result. The remaining signature genes (*ATAD1*, *MTARC2*, *MPV17*, and *PECR*), which also may act as oncoproteins or tumor suppressors, have been seldom reported. Further evaluation is required to fully elucidate the potential roles and mechanisms of the identified 10 genes in HCC.

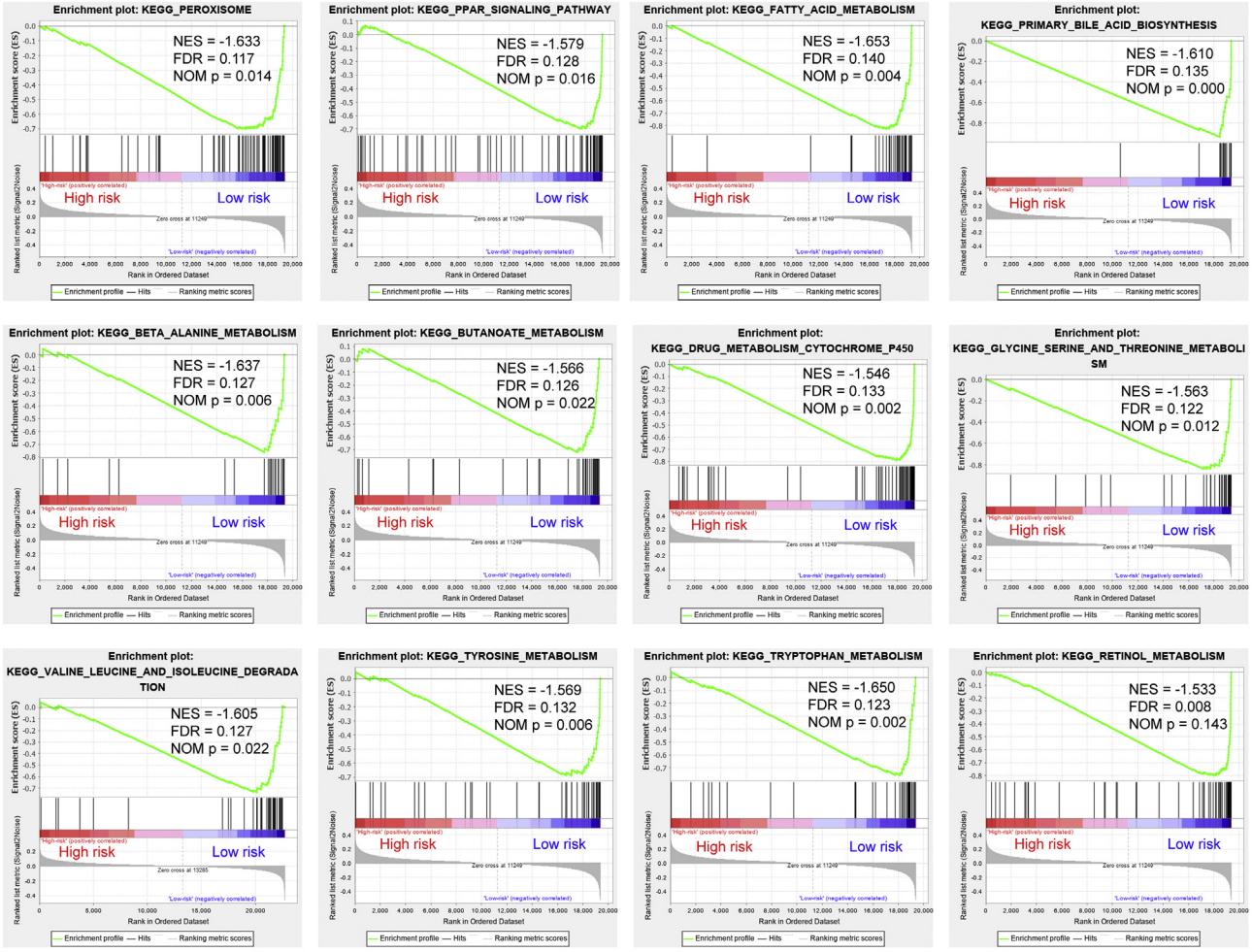


Figure 6 GSEA plots of commonly enriched KEGG pathways in the TCGA-LIHC cohort.

Evaluation and validation of the 10-gene signature in two independent HCC cohorts demonstrated that it effectively discriminated HCC patients with poor prognosis. Additionally, the results after incorporation of traditional clinical characteristics indicated that the signature can serve as an independent predictor. Our signature was found to be robust across distinct datasets when compared with six recently reported multi-gene biomarkers for liver cancer. Finally, a nomogram model composed of the risk score and TNM stage was developed for patient stratification.

Results of GSEA in both HCC cohorts revealed that the signature was significantly associated with the peroxisome, PPAR signaling, and other metabolic pathways, suggesting that it is tightly related to metabolic reprogramming in HCC. As multifunctional organelles, peroxisomes have been reported to be either directly or indirectly related to the enriched biological processes identified. Interestingly, all enriched pathways were related to the low-risk group, which indicates that peroxisome-related metabolism may benefit the survival of HCC patients. Identifying the mechanisms underlying the gene signature may enhance our understanding of the role of peroxisomes in tumorigenesis and progression.

The current risk signature has several advantages. First, it was established based on two independent and strictly screened HCC cohorts, in which samples with non-HCC pathology and inferior quality were excluded. It is a precise risk signature derived directly from HCC rather than liver cancer cohorts. Second, our signature demonstrated a more robust performance than common clinical features and other gene signatures reported previously. Third, it reflected a specific biological background and significantly correlated with several characteristics of metabolic reprogramming in HCC and may be useful in distinguishing patients with different metabolic features.

However, as our study was based on retrospective data, prospective validation of the signature is warranted in further research. In addition, the predictive performance is expected to improve if multi-omics data can be appropriately integrated into the analyses. Furthermore, the current gene signature is not yet suitable for immediate clinical application because its application requires presetting the risk-score thresholds and data normalization in a large pre-collected cohort.[49]

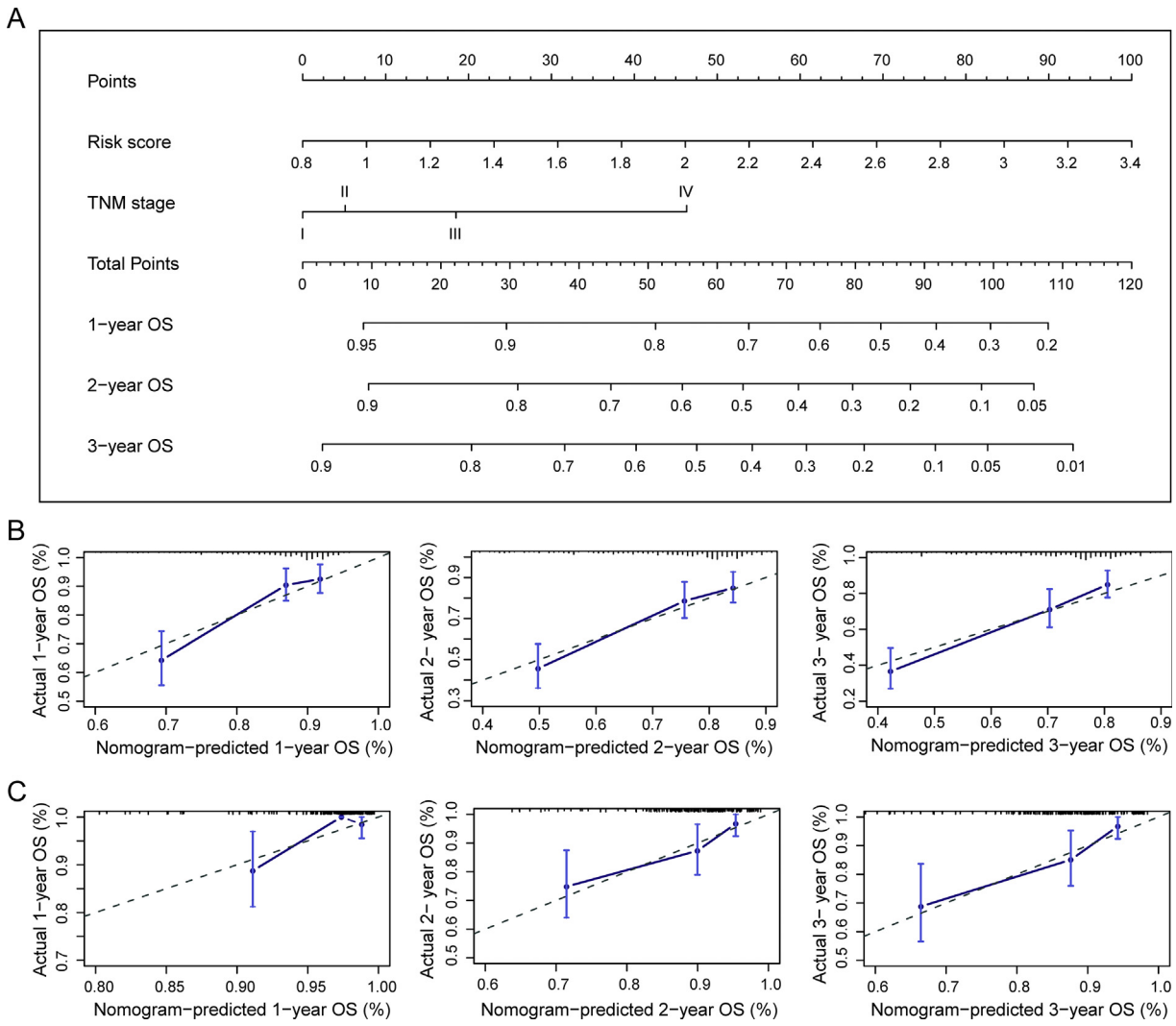


Figure 7 Construction of a nomogram for OS prediction. (**A**) Nomogram combining the gene signature with TNM stage. (**B**, **C**) Calibration curve of the nomogram for predicting the probability of OS at 1, 2, and 3 years in the TCGA-LIHC (B) and in the LIRI-JP (C) cohorts.

Conclusion

We identified a novel prognostic signature comprising 10 peroxisomal genes for HCC. The 10-gene signature may be a potentially valuable prognostic marker for HCC. Application of the 10-gene signature to personalized clinical treatment may have a positive impact on the prognosis of HCC patients.

Author contributions

ZM, LQ and XW conceived, designed, or planned the study. LQ downloaded and analyzed the data. LQ, KZ and KM wrote the first draft and prepared figures. ZM and XW helped to interpret the data. All authors revised and reviewed this work, and approved the final manuscript.

Funding

This work was supported by grant from Ministry of Science and Technology of the People's Republic of China (No. 2017ZX10203202-004-005).

Data availability statement

The datasets generated and analyzed during the current study are available in the TCGA database (https://portal.gdc.cancer.gov/repository/) and ICGC database (https://dcc.icgc.org/releases/release_28/).

Conflict of interests

The authors declare no conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.gendis.2020.04.010.

References

1. Villanueva A. Hepatocellular carcinoma. *N Engl J Med*. 2019; 380(15):1450—1462.

2. Omata M, Cheng AL, Kokudo N, et al. Asia-Pacific clinical practice guidelines on the management of hepatocellular carcinoma: a 2017 update. *Hepatol Int*. 2017;11(4):317—370.
3. Marrero JA, Kulik LM, Sirlin CB, et al. Diagnosis, staging, and management of hepatocellular carcinoma: 2018 practice guidance by the American association for the study of liver diseases. *Hepatology*. 2018;68(2):723—750.
4. European Association for the Study of the Liver. EASL Clinical Practice Guidelines. Management of hepatocellular carcinoma. *J Hepatol*. 2018;69(1):182—236.
5. Forner A, Reig M, Bruix J. Hepatocellular carcinoma. *Lancet*. 2018;391(10127):1301—1314.
6. Villanueva A, Hoshida Y, Battiston C, et al. Combining clinical, pathology, and gene expression data to predict recurrence of hepatocellular carcinoma. *Gastroenterology*. 2011;140(5): 1501—1512.
7. Li G, Xu W, Zhang L, et al. Development and validation of a CIMP-associated prognostic model for hepatocellular carcinoma. *EBioMedicine*. 2019;47:128—141.
8. Yan Y, Lu Y, Mao K, et al. Identification and validation of a prognostic four-genes signature for hepatocellular carcinoma: integrated ceRNA network analysis. *Hepatol Int*. 2019;13(5): 618—630.
9. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011;144(5):646—674.
10. De Matteis S, Ragusa A, Marisi G, et al. Aberrant metabolism in hepatocellular carcinoma provides diagnostic and therapeutic opportunities. *Oxid Med Cell Longev*. 2018;2018:7512159.
11. Nakagawa H, Hayata Y, Kawamura S, Yamada T, Fujiwara N, Koike K. Lipid metabolic reprogramming in hepatocellular carcinoma. *Cancers (Basel)*. 2018;10(11):447.
12. Islinger M, Voelkl A, Fahimi HD, Schrader M. The peroxisome: an update on mysteries 2.0. *Histochem Cell Biol*. 2018;150(5): 443—471.
13. Lodhi IJ, Semenkovich CF. Peroxisomes: a nexus for lipid metabolism and cellular signaling. *Cell Metab*. 2014;19(3): 380—392.
14. Walker CL, Pomatto LCD, Tripathi DN, Davies KJA. Redox regulation of homeostasis and proteostasis in peroxisomes. *Physiol Rev*. 2018;98(1):89—115.
15. Dahabieh MS, Di Pietro E, Jangal M, et al. Peroxisomes and cancer: the role of a metabolic specialist in a disease of aberrant metabolism. *Biochim Biophys Acta Rev Cancer*. 2018; 1870(1):103—121.
16. Lloyd MD, Yevglevskis M, Lee GL, Wood PJ, Threadgill MD, Woodman TJ. alpha-Methylacyl-CoA racemase (AMACR): metabolic enzyme, drug metabolizer and cancer marker P504S. *Prog Lipid Res*. 2013;52(2):220—230.
17. Benjamin DI, Cozzo A, Ji X, et al. Ether lipid generating enzyme AGPS alters the balance of structural and signaling lipids to fuel cancer pathogenicity. *Proc Natl Acad Sci U S A*. 2013; 110(37):14912—14917.
18. Turner-Ivey B, Manevich Y, Schulte J, et al. Role for Prdx1 as a specific sensor in redox-regulated senescence in breast cancer. *Oncogene*. 2013;32(45):5302—5314.
19. Cai M, Sun X, Wang W, et al. Disruption of peroxisome function leads to metabolic stress, mTOR inhibition, and lethality in liver cancer cells. *Cancer Lett*. 2018;421:82—93.
20. Mattu S, Fornari F, Quagliata L, et al. The metabolic gene HAO2 is downregulated in hepatocellular carcinoma and predicts metastasis and poor survival. *J Hepatol*. 2016;64(4):891—898.
21. Feng Y, Xiong Y, Qiao T, Li X, Jia L, Han Y. Lactate dehydrogenase A: a key player in carcinogenesis and potential target in cancer therapy. *Cancer Med*. 2018;7(12):6124—6136.
22. Schluter A, Real-Chicharro A, Gabaldon T, Sanchez-Jimenez F, Pujol A. PeroxisomeDB 2.0: an integrative view of the global peroxisomal metabolome. *Nucleic Acids Res*. 2010;38(Database issue):D800—D805.
23. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res*. 2017;45(D1):D353—D361.
24. Gronemeyer T, Wiese S, Ofman R, et al. The proteome of human liver peroxisomes: identification of five new peroxisomal constituents by a label-free quantitative proteomics survey. *PLoS One*. 2013;8(2):e57395.
25. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
26. Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal*. 2013;6(269):pl1.
27. Fujimoto A, Furuta M, Totoki Y, et al. Whole-genome mutational landscape and characterization of noncoding and structural mutations in liver cancer. *Nat Genet*. 2016;48(5): 500—509.
28. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43):15545—15550.
29. Zheng Y, Liu Y, Zhao S, et al. Large-scale analysis reveals a novel risk score to predict overall survival in hepatocellular carcinoma. *Cancer Manag Res*. 2018;10:6079—6096.
30. Xiang XH, Yang L, Zhang X, et al. Seven-senescence-associated gene signature predicts overall survival for Asian patients with hepatocellular carcinoma. *World J Gastroenterol*. 2019; 25(14):1715—1728.
31. Wang Z, Teng D, Li Y, Hu Z, Liu L, Zheng H. A six-gene-based prognostic signature for hepatocellular carcinoma overall survival prediction. *Life Sci*. 2018;203:83—91.
32. Chang WH, Forde D, Lai AG. A novel signature derived from immunoregulatory and hypoxia genes predicts prognosis in liver and five other cancers. *J Transl Med*. 2019;17(1):14.
33. Liu GM, Xie WX, Zhang CY, Xu JW. Identification of a four-gene metabolic signature predicting overall survival for hepatocellular carcinoma. *J Cell Physiol*. 2020;235(2):1624—1636.
34. Liu GM, Zeng HD, Zhang CY, Xu JW. Identification of a six-gene signature predicting overall survival for hepatocellular carcinoma. *Cancer Cell Int*. 2019;19:138.
35. Schott AF, Perou CM, Hayes DF. Genome medicine in cancer: what's in a name? *Cancer Res*. 2015;75(10):1930—1935.
36. Kamps R, Brandao RD, van den Bosch BJ, et al. Next-generation sequencing in oncology: genetic diagnosis, risk prediction and cancer classification. *Int J Mol Sci*. 2017;18(2):308.
37. Piano V, Benjamin DI, Valente S, et al. Discovery of inhibitors for the ether lipid-generating enzyme AGPS as anti-cancer agents. *ACS Chem Biol*. 2015;10(11):2589—2597.
38. Brand A, Singer K, Koehl GE, et al. LDHA-associated lactic acid production blunts tumor immunosurveillance by T and NK cells. *Cell Metab*. 2016;24(5):657—671.
39. Jin L, Chun J, Pan C, et al. Phosphorylation-mediated activation of LDHA promotes cancer cell invasion and tumour metastasis. *Oncogene*. 2017;36(27):3797—3806.
40. Kallijarvi J, Avela K, Lipsanen-Nyman M, Ulmanen I, Lehesjoki AE. The TRIM37 gene encodes a peroxisomal RING-B-box-coiled-coil protein: classification of mulibrey nanism as a new peroxisomal disorder. *Am J Hum Genet*. 2002;70(5): 1215—1228.
41. Bhatnagar S, Gazin C, Chamberlain L, et al. TRIM37 is a new histone H2A ubiquitin ligase and breast cancer oncoprotein. *Nature*. 2014;516(7529):116—120.
42. Li Y, Deng L, Zhao X, et al. Tripartite motif-containing 37 (TRIM37) promotes the aggressiveness of non-small-cell lung cancer cells by activating the NF-kappaB pathway. *J Pathol*. 2018;246(3):366—378.
43. Ding C, Fan X, Wu G. Peroxiredoxin 1 - an antioxidant enzyme in cancer. *J Cell Mol Med*. 2017;21(1):193—202.

44. Wirthschaft P, Bode J, Simon AEM, et al. A PRDX1-p38alpha heterodimer amplifies MET-driven invasion of IDH-wildtype and IDH-mutant gliomas. *Int J Cancer*. 2018;143(5):1176—1187.
45. Chen WC, Wang CY, Hung YH, Weng TY, Yen MC, Lai MD. Systematic analysis of gene expression alterations and clinical outcomes for long-chain acyl-coenzyme A synthetase family in cancer. *PLoS One*. 2016;11(5):e0155660.
46. Goudarzi A. The recent insights into the function of ACAT1: a possible anti-cancer therapeutic target. *Life Sci*. 2019;232:116592.
47. Fan J, Lin R, Xia S, et al. Tetrameric acetyl-CoA acetyltransferase 1 is important for tumor growth. *Mol Cell*. 2016;64(5):859—874.
48. Garcia-Bermudez J, Birsoy K. Drugging ACAT1 for cancer therapy. *Mol Cell*. 2016;64(5):856—857.
49. Qi L, Chen L, Li Y, et al. Critical limitations of prognostic signatures based on risk scores summarized from gene expression levels: a case study for resected stage I non-small-cell lung cancer. *Brief Bioinform*. 2016;17(2):233—242.