

Four Metastasis-Related mRNAs Signature Predicting the Survival of Patients With Liver Hepatocellular Carcinoma

chao chen

North Sichuan Medical College

ShiXiang Qiu

North Sichuan Medical College

Ya Li

North Sichuan Medical College

YanQun Liu

Nan Chong Central Hospital

Kang Liu

Nan Chong Central Hospital

LiMing Zhong (✉ li_ming_zhong@126.com)

Nan Chong Central Hospital

Research Article

Keywords: Liver hepatocellular carcinoma, Metastasis, Prognostic model, TCGA, GSEA.

Posted Date: April 14th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-358201/v1>

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

[Read Full License](#)

Four metastasis-related mRNAs signature predicting the survival of patients with Liver hepatocellular carcinoma

Chao Chen^{1,2,3}, ShiXiang QIU^{1,2,3}, Ya Li^{1,2,3}, YanQun Liu², Kang Liu², LiMing Zhong³

(1. School of Medical Imaging, North Sichuan Medical College, Nanchong 637000, Sichuan, China;

(2. Institute of Tissue Engineering and Stem Cell Research, Nan Chong Central Hospital, The second Clinical College of North Sichuan Medical College, Nanchong 637000, Sichuan, China;

(3. Department of Interventional Radiology, Nan Chong Central Hospital, The Second Clinical College of North Sichuan Medical College, Nan Chong 637000, Sichuan, China)

ABSTRACT

Backgrounds: Liver hepatocellular carcinoma (LIHC) is one of the most malignant tumors, of which prognosis is unsatisfactory in most cases and metastatic of LIHC often results in poor prognosis. In this study, we aimed to construct a metastasis-related mRNAs prognostic model to increase the accuracy of prediction of LIHC prognosis.

Methods: 374 LIHC samples and 50 normal samples were downloaded from TCGA database, involving transcriptomic and clinical data. Metastatic-related genes were acquired from HCMBD website at the same time. 343 samples were randomly divided into train dataset and test dataset with a proportion of 1:1 by using caret package in R. Kaplan-Meier method and univariate Cox regression analysis and lasso regression analysis were performed to obtain metastasis-related mRNAs which played significant roles in prognosis. Then, using multivariate Cox regression analysis, a prognostic prediction model was established. Transcriptome and clinical data were combined to construct a prognostic model and a nomogram for OS evaluation. Functional enrichment in high- and low-risk groups were also analyzed by GSEA. An entire set was applied to verify the model.

Results: 1895 metastasis-related mRNAs were screened and 8mRNAs were associated with prognosis. The overall survival (OS)-related prognostic model which was constructed based on 4 MRGs (MMP1, SPP1, STC2, CDCA8) significantly stratified LIHC patients into high- and low-risk groups. The AUC values of the 4-gene prognostic signature at 1 year, 2 years, and 3 years were 0.807, 0.729 and 0.673. A risk score based on the signature was a significantly independent prognostic factor (HR=1.295; 95%CI=1.167-1.436; $P<0.001$) for LIHC patients. A nomogram which incorporated the 4-gene signature and clinical features was also built for prognostic prediction. GSEA results that low- and high-risk group had an obviously difference in part of pathways. The value of this model was validated in test dataset and entire set.

Conclusion: Metastasis-related mRNAs prognostic model was verified that it had a predictable value on the prognosis of LIHC, which could be helpful for gene targeted therapy.

Keywords: Liver hepatocellular carcinoma, Metastasis, Prognostic model, TCGA, GSEA.

Introduction

Liver hepatocellular carcinoma had been the sixth most commonly diagnosed cancer and the second leading cause of cancer death worldwide in 2020, with about 905677 new cases and 830180 deaths annually ^[1]. The 5-year survival and OS rates are below 12%. Precursors of most LIHC cases include liver cirrhosis, chronic hepatitis viral infections, alcohol-related liver disease, non-alcoholic fatty liver disease, and drug-induced hepatitis ^[2]. Curing LIHC had been a complex issue for doctors since its birth. Surgery treatment was the main method of liver hepatocellular carcinoma. Interventional therapy is a new way for patients which through a minimally invasive surgery to suppress its proliferation. However, prognosis was often not very good. LIHC is a highly aggressive and heterogeneous disease ^[3]. For advanced LIHC cases, moreover, the recurrence rate is nearly 80% with the patients and the metastasis rate is nearly 30%, whose 5-year survival rate is only 25–39% ^[4]. In addition, LIHC' metastasis and recurrence led to shorter survival time and worse survival quality. Metastasis of LIHC is one of the important reasons for poor prognosis. Liver hepatocellular carcinoma usually metastasizes in liver in the early time and it is easy to invade portal vein and branches and form tumor thrombus, which will cause multiple metastases in the liver after falling off. Lung is the most common organ liver hepatocellular carcinoma metastasizes through blood. Hilar lymph nodes are the most common metastatic lymph nodes. It is because of the lack of symptoms and

metastasizing in the liver in early stage that most patients lost opportunities for surgeries ^[5]. At the time of LIHC diagnosis, only 5 to 15% of cases have the extrahepatic spread ^[6]. As a result, a demand for new markers to diagnose LIHC and predict prognosis is of great urgency.

The metastatic establishment of cancers at distant organs is largely incurable and primarily contributes to the deaths of cancer patients ^[7]. Many mRNAs had been reported that they were related to the metastasis of tumors or chemoresistance. For example, the overexpression of ACTN2 in human liver cancer cells enhanced cellular motility and invasion abilities which suggested it could be functional in liver cancers' metabolism ^[8]. Misawa's study showed that prostate cancer *HOXA11-AS* and *HOXB13*(a kind of metastasis gene) promote metastasis through CCL2/CCR2 signaling pathway in autocrine and paracrine manners ^[9]. HMGB1 was found that its suppression could be useful for CDDP sensitization ^[10]. In one hand, the abnormal of metastasis-related mRNAs had deeper relation to the metabolism and proliferation which could contribute to the prediction of tumors' development. In other hand, Cancer lethality is mainly caused by metastasis. Therefore, understanding the nature of the mRNAs involved in this process has become a priority ^[11].

We designed a study which extracting a series of metastasis-related mRNAs and combining them with clinical data to find out whether they had connections with OS. This study was aimed to find a better way to evaluate LIHC prognosis through analyzing LIHC risk score system and establishing metastasis-related lncRNA prognostic model and guide clinical treatment.

METHODS:

Data collection and processing

The overall study progress was showed in Fig. 1. Patients' transcriptome and clinical data were downloaded from TCGA database. The former includes 374 cases of LIHC and 50 normal cases and the latter involved age, gender, grade, stage and survival time, etc. The samples whose survival time was less than 30 days were deleted. We downloaded metastasis-related genes from HCBMD website and combined mRNAs which were expressed in LIHC patients and metastasis-related genes. "limma" packages in R was used to distinguish the mRNAs which were expressed differently in tumor and normal samples. FDR less than 0.05 and $|\log_2(FC)|$ higher than 1 were set as value. Then, 8 metastasis-related mRNAs had we obtained after we adopted univariate Cox regression analysis and lasso regression analysis to filter mRNAs. The gene which

p-value was <0.05 in univariate regression analysis was regarded as a candidate gene for prognosis. The LIHC data sets of 343 TCGA patients used for prognosis analysis were divided into training set (n = 172) and invalidation set (n = 171) according to the proportion of 50% and 50% by using caret package in R software.

Gene functional annotation of DE-MRGs

The Gene Ontology (GO) function ^[12] and Kyoto Encyclopedia of mRNAs and Genomes (KEGG) ^[13] pathway enrichment regarding the differentially expressed IRGs were analyzed using the “clusterprofiler” R package. The results which $p < 0.05$ were considered as statistically significant.^[12]

Construction of regulatory network and prognostic model

In training dataset, only mRNAs with $P < 0.05$ according to multivariate Cox regression analysis were considered as prognostic metastasis-related mRNAs. After deleting the patients without complete clinical information, we calculated each patient's risk score through constructing a prognostic model to evaluate LIHC patients' prognosis.

$$\text{Risk Score}(\text{patient}) = \sum_i \text{Coefficient}(\text{mRNA}_i) \times \text{Expression}(\text{mRNA}_i)$$

Patients and their mRNAs' expression were combined and divided into two groups according to risk scores. Data was classified by the median risk score threshold as high or low risk group. Also, mRNAs were divided into high and low expression group by the median expression threshold. Using the “survival” software package in R, the survival curve of different expression of metastasis-related mRNAs and different risk groups were visualized. In addition, we constructed a risk curve with package “survival” and “heatmap” in R. After that, a time-dependent ROC (receiver operating characteristic) curve for OS prediction was used to assess the sensitivity and specificity of the prognosis through utilizing package “timeROC” in R.

Clinical correlation analysis and validation of the prognostic model

Univariate and multivariate were implemented to confirm the independency of the prognostic model with or without clinical elements (age, gender, grade, stage). Thereafter, clinical characters and risk score were classified into two categories to

protracted a nomogram. Calibration curves were made to exam the nomogram. Nomogram is widely used to predict cancer prognosis ^[14]. Since the model had been constructed, validation group(n=171) and entire set(n=232) were used to evaluate the feasibility of the model. Similarly, methods mentioned before were used in both testing cohorts.

Functional analysis

To reveal the KEGG pathways involved in high-risk and low-risk groups, GSEA analysis was performed to define the biological processes enriched in the gene rank between the two groups. GSEA with Java program was implemented to analyze the functions of the prognostic model. The random sample permutation number was set as 1,000, and the significance threshold was $P < 0.05$, adopting high risk versus low risk.

Statistical Analysis

All statistical analyses were performed using the R 4.0.3 software and Perl packages was utilized to arrange the data. Of all the process, $P < 0.05$ was recognized as statistically significant.

Results:

GO and KEGG functional annotations

GO analysis demonstrated that these DEGs were significantly enriched in biological process (BP), including epithelial cell proliferation, gland development, regulation of epithelial cell proliferation, regulation of blinding and regulation of apoptotic signaling pathway. In addition, in the cellular component (CC) analysis, these DE-MRGs were enriched in the focal adhesion, focal adhesion and cell–substrate junction. It was suggested that in molecular function (MF), cell adhesion molecule binding and receptor ligand activity were enriched mostly. (Fig. 2A). KEGG analysis results showed that these mRNAs mainly enriched in Human papillomavirus infection and PI3K–Akt signaling pathway (Fig. 2B).

Metastasis-related mRNAs prognostic model and survival analysis

1895 mRNAs were screened after metastasis-related genes and mRNAs which were expressed in LIHC patients were overlapped. By intersecting the DEGs and metastasis-related mRNAs in LIHC, 666 mRNAs were up-regulated and 73 mRNAs were down-regulated among 739 different expression mRNAs. The DEGs were showed in a heatmap and a volcano plot (Fig. 3A,3B). Whereafter, through univariate Cox regression analysis with $p < 0.05$, we extracted 134 metastasis-related mRNAs associated with prognosis. Next, lasso regression analysis (Fig.4A,4B) and multivariate Cox regression analysis were performed to screen the candidate mRNAs. As a result, we acquired 4 MRGs (MMP1, SPP1, STC2, CDCA8) which could be used to construct prognostic model. All mRNAs were considered as negative factors. Multivariate Cox regression analysis was performed to establish prognostic model and construct a risk score prognostic index. Risk score was calculated by the following formula: [Expression level of MMP1*(0.25)] + [Expression level of SPP1*(0.08)] + [Expression level of STC2*(0.34)] + [Expression level of CDCA8*(0.34)].

Through the level of Risk score, LIHC patients could be split into two groups: high-risk and low-risk group. Risk score which was higher than median would be classified as the high-risk group. Survival time of 4 mRNAs was displayed by “survival” package in R (Fig. 5A-D). Meanwhile, R was used to draw survival curve of 4 mRNAs according to different risk score (Fig. 6A-C). As is shown in the figures, with the growth of risk scores, the risk rating increases gradually. The same was true in the survival diagram and heatmap of 4 mRNAs. The survival intervals could be significantly differentiated in LIHC patients with high and low risk. Survival analysis showed that high-risk group had shorter overall survival time than low-risk group (Fig. 6D). Such results showed that risk score could be a potential index for prognosis prediction of LIHC patients. We also drew a ROC curve to assess the sensitivity and specificity of the model for 1-,2-,3-year survival, with an AUC of 0.807,0.729 and 0.673 (Fig. 6E).

Independent prognostic analysis and Clinical correlation analysis

Since constructing a risk score model, we integrated the model with clinical characters and adopted univariate and multivariate Cox regression analysis to identify whether risk score could be an independent prognostic factor. Results showed that risk score could evaluate prognosis by itself in univariate Cox regression analysis ($P < 0.05$) (Fig. 7A) and it could also be an obviously predictable factor for prognosis

after eliminating the influence of other characters (Fig. 7B). We used a method of bisection to divide the clinical traits (including age, gender, grade, stage) and risk score into two groups. Results indicated that with the growth of tumors' stage, the risk score of patients increased ($P<0.05$) and other characters didn't show significant connection. Similarly, we employed package "rms" to draw an OS nomogram at 1-, 2- and 3-year in LIHC patients. We performed subgroup analysis of the two signatures in age (<65 , ≥ 65), clinical stage (stage I-II, stage III-IV), gender (female, male) and riskScore (low, high). Results showed that shorter OS happened in subgroup of ≥ 65 , stage III-IV and high-risk. In addition, a verified calibration curve for 1-, 2- and 3-year was plotted to testify the nomogram (Fig. 8A-D).

Functional analysis of prognostic model

KEGG enriched analysis had we adopted to find out which pathway related to cancer would the mRNAs focus on and validate the biological function of the constructed model. Results (Fig. 9) showed that mRNAs which were of high score had an obvious enrichment in pathways such as "Notch signaling pathway", "VEGF signaling pathway", "WNT signaling pathway", "ERBB signaling pathway" etc. while "PPAR signaling pathway", "drug metabolism-cytochrome p450", "fatty acid metabolism" and "Glycine serine and threonine metabolism" etc. were enriched in low-risk group.

Validation of the model

Following the primary methods and same coefficients, we established two risk score models coming from testing dataset to verify the accuracy of the metastasis-related prognosis model constructed before and all of the 4 metastasis-related mRNAs were validated in TCGA testing data and entire set. Patients from the testing cohort were also divided into high or low risk group. Survival curve was supplemented in both two groups (Fig. 10A-C, 11A-C). In line with the results of the TCGA testing cohort and entire set, patients who were classified as high-risk also had significantly inferior OS than the low-risk group ($P<0.05$) (Fig. 10D, 11D). To assess the predictive performance of the 4-gene based signature, we constructed a time-dependent ROC curve in two validation groups (Fig. 10E, 11E), with an AUC of 0.785, 0.710, 0.703 for 1-, 2-, 3- year survival in testing cohort and an AUC of 0.754, 0.753, 0.769 for 1-, 2-, 3- year survival in entire set. Similarly, we found that risk score could be an independent factor through both univariate and multivariate Cox regression analysis ($P<0.05$) (Fig. 12A-D).

Discussion

Liver cancer has been one of the most diseases contributed to death and many prognostic models had been established to predict the prognosis of LIHC. For example, autophagy-related long non-coding RNAs prognostic model ^[15], immune-related long non-coding RNAs prognostic model ^[16] and metastasis-related miRNAs ^[17] etc. Here we established a metastasis-related mRNAs prognostic model to predict the prognosis of LIHC.

In this article, of all the analysis, LIHC samples were randomly classified into training cohort, and testing cohort. Training cohort was used to construct a prognostic model, while testing cohort was utilized for validation. Firstly, we analyzed the gene expression data and clinical data of LIHC patients enrolled in TCGA, discerning 881 mRNAs related to metastasis. 4 mRNAs (MMP1, SPP1, STC2, CDCA8) had we found using univariate, lasso and multivariate Cox regression analysis were detected as independent prognosis predictors in LIHC. Secondly, survival analysis was utilized to examine the availability of the prognostic model. The high expression of all the 4 mRNAs, had a positive correlation to OS which meant that with the generate of these mRNAs' expression, patients would have a longer survival time. The results suggested that metastasis-related mRNAs model was a significant prognostic factor for LIHC patients. Thirdly, the model constructed in training group was validated internally and externally, adding dependability to the outcomes.

The mRNAs mentioned before had been reported in other articles that they also had relationship with different types of cancers. A study from Gabasa et al. ^[18] found that the expression of MMP1 is necessary for induction of fibroblast senescence and consequent tumor promotion and MMP1, in combination with TGF- β 1, is sufficient to induce fibroblast senescence and consequent LCC promotion. Up-regulation of SPP1 had a relationship with increased risk of LC in patients with COPD and it could a potential therapeutic target for LC in patients with COPD ^[19]. Li et al. ^[20] found that positive expression of BIRC7 and STC2 are associated with progression and poor clinical outcomes of EHCC which would be a potential biomarker for EHCC. Moreover, high expression of CDCA8 promoted the proliferation of ovarian cancer cells in vitro and in vivo which increased the tumorigenesis, aggressiveness and chemoresistance of ovarian cancer ^[21]. Those results had represented similar conclusions as this study.

GSEA enriched manifested these mRNAs had effects on prognosis probably via several signaling pathways. For example, ERBB signaling pathway was found that it was related to hepatocellular carcinoma ^[22] while the progress of hepatocellular carcinoma had a relationship with Notch signaling pathway through a series of target spots ^[23-24]. Moreover, VEGF signaling pathway ^[25-26], WNT signaling pathway ^[27-28] and PPAR signaling pathway ^[29] suggested similar results. Such findings showed

the biological functions of mRNAs could be regarded as predictable factor for prognosis though more researches and experiments needed to be done to verify the hypothesis. These results help us to explore the mechanism of metastasis-related mRNAs.

In conclusion, through a series of bioinformation analysis, we constructed a model and set up a biomarker for predicting the prognosis of LIHC. Our study showed that patients who were in low-risk group had better OS than those in high-risk group. Such results were verified in both train and test cohort. Hence, this model had a good sensitivity and specificity on 1 year-, 2 years-, 3 years- survival time for LIHC patients, having AUC values of the models at 1 year-, 2 years-, 3 years- survival were varied from 0.673 to 0.807. This study was highly methodologically reasonable because we downloaded data from TCGA database which contained great amounts of samples and opened a new prospect for the regulation of metabolic processes and the treatment^[30] of LIHC. However, there are several deficiencies in this study. Firstly, data in TCGA may have variable degrees of errors and the amount of data included is not large, which may cause inaccuracy. Secondly, lacking of experiments *in vivo* and *in vitro* will lead to insufficient evidence for this model. Thus, future studies and more experiments should be implemented to validate the model and biomarker and ensure its robustness.

Conclusion

We constructed a 4 metastasis-related mRNAs prognostic model based on MRGs and separated LIHC patients from two groups. The survival outcomes of the two groups were statistically different, which meant that the disparate expression of MRGs may have effects on patients' prognosis. These findings may promote the development of new biomarkers and targeted therapies. Therefore, the 4 metastasis-related mRNAs and might be molecular biomarkers and therapeutic targets for the patients with liver hepatocellular carcinoma.

Conflict of Interest

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

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

All data generated or analysed during this study are included in this published article

Funding

No funding

Authors' contributions

Chao Chen wrote the main manuscript text and Chao Chen, ShiXiang Qiu, YanQun Liu, Ya Li prepared figures 1-12. All authors reviewed the manuscript.

Acknowledgements

We would like to thank the Tissue Engineering and Stem Cell Institute of Nanchong Central Hospital for providing us with a learning and experimental platform and thank NingJun Yu for accompanying me.

Reference

- [1] Latest global cancer data: Cancer burden rises to 19.3 million new cases and 10.0 million cancer deaths in 2020. Retrieved Nov 16 ,2020.

-
- [2] Gu Xinyu, Guan Jun, Xu Jia et al. Model based on five tumour immune microenvironment-related genes for predicting hepatocellular carcinoma immunotherapy outcomes.[J] . Journal of Translational Medicine, 2021, 19: 26.
- [3] Chen Bingqing, Liao Zhibin, Qi Yongqiang et al. miR-631 Inhibits Intrahepatic Metastasis of Hepatocellular Carcinoma by Targeting PTPRE.[J]. Frontiers Oncology, 2020, 10: 565266.
- [4] Li Wang, Chen Qi-Feng, Huang Tao et al. Identification and Validation of a Prognostic lncRNA Signature for Hepatocellular Carcinoma [J]. Frontiers Oncology, 2020, 10: 780.
- [5] Yoo D. J., Kim K. M., Jin Y. J., et al. Clinical outcome of 251 patients with extrahepatic metastasis at initial diagnosis of hepatocellular carcinoma: does transarterial chemoembolization improve survival in these patients? [J]. *Journal of Gastroenterology and Hepatology*. 2011;26(1):145–154.
- [6] Maluccio M, Covey A. Recent progress in understanding, diagnosing, and treating hepatocellular carcinoma[J]. A Cancer Journal For Clinicians, 2012, 62(6):394–399.
- [7] Bergers G, Fendt S M. The metabolism of cancer cells during metastasis[J]. Nature reviews. Cancer, 2021.
- [8] Lo Lilian H, Lam Coco Y, To Jeffrey C et al. Sleeping Beauty insertional mutagenesis screen identifies the pro-metastatic roles of CNPY2 and ACTN2 in hepatocellular carcinoma tumor progression.[J] .Biochemical and Biophysical Research Communication, 2021, 541: 70-77.
- [9] Misawa Aya, Kondo Yukihiro, Takei Hiroyuki et al. HOXA11-AS Long Noncoding RNA and Transcription Factor HOXB13 Modulate the Expression of Bone Metastasis-Related Genes in Prostate Cancer.[J] .Genes (Basel), 2021, 12: undefined.
- [10] Nishiguchi Yukiko, Oue Naohide, Fujiwara-Tani Rina et al. Role of Metastasis-Related Genes in Cisplatin Chemoresistance in Gastric Cancer.[J] .Int J Mol Sci, 2019, 21: undefined.
- [11] Shajari Neda, Davudian Sadaf, Kazemi Tohid et al. Silencing of BACH1 inhibits invasion and migration of prostate cancer cells by altering metastasis-related gene expression.[J] .Artificial Cells Nanomedicine Biotechnology, 2018, 46: 1495-1504.

-
- 【12】 Liu Jinhui,Meng Huangyang,Nie Sipei et al. Identification of a prognostic signature of epithelial ovarian cancer based on tumor immune microenvironment exploration.[J] .Genomics, 2020, 112: 4827-4841.
- 【13】 Liu Jinhui,Wu Zhipeng,Wang Yichun et al. A prognostic signature based on immune-related genes for cervical squamous cell carcinoma and endocervical adenocarcinoma. [J]. International Immunopharmacology, 2020, 88: 106884.
- 【14】 Iasonos Alexia,Schrag Deborah,Raj Ganesh V et al. How to build and interpret a nomogram for cancer prognosis. [J]. Journal of Clinical Oncology, 2008, 26: 1364-70.
- 【15】 Zhao jisen, Li Jinghua, Yang Jihong, et al. Establishment and analysis of prognosis model of autophagy related long non coding RNA in hepatocellular carcinoma [J]. Chinese Journal of general surgery, Vol. 29, No. 7, 2020, pp. 839-848, ISTIC PKU CSCD Ca, 2020
- 【16】 Yuan Mengqin,Wang Yanqing,Sun Qinqian et al. Identification of a Nine Immune-Related lncRNA Signature as a Novel Diagnostic Biomarker for Hepatocellular Carcinoma.[J] .Biomed Research International, 2021, 2021: 9798231.
- 【17】 Chen Yuan,Wang Guifu,Xu Hao et al. Identification of a Novel Metastasis-Related miRNAs-Based Signature for Predicting the Prognosis of Hepatocellular Carcinoma.[J] .Journal of Oncology, 2021, 2021: 6629633.
- 【18】 Gabasa Marta,Radisky Evette S,Ikemori Rafael et al. MMP1 drives tumor progression in large cell carcinoma of the lung through fibroblast senescence.[J] .Cancer Letters, 2021, 507: 1-12.
- 【19】 Miao Ti-Wei,Xiao Wei,Du Long-Yi et al. High expression of SPP1 in patients with chronic obstructive pulmonary disease (COPD) is correlated with increased risk of lung cancer.[J] .FEBS Open Bio, 2021, undefined: undefined.
- 【20】 Li Jiequn,Yang Zhulin,Huang Shengfu et al. BIRC7 and STC2 Expression Are Associated With Tumorigenesis and Poor Outcome in Extrahepatic Cholangiocarcinoma.[J] .Technology in Cancer Research and Treatment, 2020, 19: 1533033820971676
- 【21】 Qi Gonghua,Zhang Chenyi,Ma Hanlin et al. CDCA8, targeted by MYBL2, promotes malignant progression and olaparib insensitivity in ovarian cancer.[J] .American Journal of Cancer Research, 2021, 11: 389-415.

-
- 【22】 Liu Yuanhui, Calmel Claire, Desbois-Mouthon Christèle et al. Regulation of the EGFR/ErbB signaling by clathrin in response to various ligands in hepatocellular carcinoma cell lines. [J]. *Journal of Cellular and Molecular Medicine*, 2020, 24: 8091-8102.
- 【23】 Saha Subbroto Kumar, Choi Hye Yeon, Yang Gwang-Mo et al. GPR50 Promotes Hepatocellular Carcinoma Progression via the Notch Signaling Pathway through Direct Interaction with ADAM17. [J]. *Molecular Therapy Oncolytics*, 2020, 17: 332-349.
- 【24】 Luiken Sarah, Fraas Angelika, Bieg Matthias et al. NOTCH target gene HES5 mediates oncogenic and tumor suppressive functions in hepatocarcinogenesis.[J]. *Oncogene*, 2020, 39: 3128-3144.
- 【25】 Bolandi Seyed Mohammadreza, Abdolmaleki Zohreh, Assarehzadegan Mohammad-Ali, Bevacizumab regulates inflammatory cytokines and inhibits VEGFR2 signaling pathway in an ovalbumin-induced rat model of airway hypersensitivity.[J]. *Inflammopharmacology*, 2021, undefined: undefined.
- 【26】 Vion Anne-Clémence, Perovic Tijana, Petit Charlie et al. Endothelial Cell Orientation and Polarity Are Controlled by Shear Stress and VEGF Through Distinct Signaling Pathways.[J]. *Frontiers of Physiology*, 2020, 11: 623769.
- 【27】 Fiechter Renée H, de Jong Henriëtte M, van Mens Leonieke J J et al. IL-12p40/IL-23p40 Blockade With Ustekinumab Decreases the Synovial Inflammatory Infiltrate Through Modulation of Multiple Signaling Pathways Including MAPK-ERK and Wnt.[J]. *Frontiers of Immunology*, 2021, 12: 611656.
- 【28】 Xu Yingxing, Jiang Yaping, Jia Bin et al. Icariin stimulates osteogenesis and suppresses adipogenesis of human bone mesenchymal stem cells via miR-23a-mediated activation of the Wnt/ β -catenin signaling pathway.[J]. *Phytomedicine*, 2021, 85: 153485.
- 【29】 Kong Lingjian, Chen Jing, Ji Xiaoli et al. Alcoholic fatty liver disease inhibited the co-expression of Fmo5 and PPAR α to activate the NF- κ B signaling pathway, thereby reducing liver injury via inducing gut microbiota disturbance.[J]. *Journal of Experiment and Clinical Cancer Research*, 2021, 40: 18.
- 【30】 Liu JinHui, Li SiYue, Feng Gao et al. Nine glycolysis-related gene signature predicting the survival of patients with endometrial adenocarcinoma.[J]. *Cancer Cell International*, 2020, 20: 183.

Fig. 1. Overall study progress

Fig. 2. Gene functional enrichment of differentially expressed IRGs. (A) Gene ontology analysis; (B) The Kyoto Encyclopedia of MRNAs and Genomes analysis.

Fig.3. Identification of differentially expressed metastasis-related mRNAs. (A-B) Heatmap and volcano plot of differentially expressed mRNAs in EOC based on data from TCGA.

Fig. 4. (A-B) The coefficients calculated by LASSO.

Fig. 5. Prognostic gene survival curve. (A) MMP1 (B) SPP1. (C) STC2. (D) CDCA8

Fig. 6. The prognostic results of risk score in train dataset. (A) Rank of prognostic index and distribution of high and low risk groups. (B) Survival status of patients in different groups. (C) Heatmap of expression profiles of included mRNAs in high and low risk groups. (D) Survival curves of the LIHC patients with different risk scores. (E) Time-dependent ROC curves of the risk model for the 1-, 2- and 3-year survival.

Fig. 7. Cox analysis of the 4 mRNAs signature. (A)Univariate Cox regression analysis of characteristics and risk score of the LIHC patients in the train dataset. (B)Multivariate Cox regression analysis of characteristics and risk score of the LIHC patients in the train dataset.

Fig. 8. The nomogram to predict 1-,2, and 3-year OS and prognostic value of 4 mRNAs in the training cohort. (A) Nomogram for OS at 1-, 2- and 3-year in LIHC patients. (B-D) Calibration plot at 1-, 2- and 3-year for validation to predict the probability of OS.

Fig. 9. Gene Set Enrichment Analysis in TCGA database. Enrichment Map were used for visualization of the GSEA results.

Fig. 10. The prognostic results of risk score in TCGA testing dataset. (A) Rank of prognostic index and distribution of high and low risk groups. (B) Survival status of patients in different groups. (C) Heatmap of expression profiles of included mRNAs in high and low risk groups. (D) Survival curves of the LIHC patients with different risk scores. (E) Time-dependent ROC curves of the risk model for the 1-, 2- and 3-year survival.

Fig. 11. The prognostic results of risk score in the entire set. (A) Rank of prognostic index and distribution of high and low risk groups. (B) Survival status of patients in different groups. (C) Heatmap of expression profiles of included mRNAs in high and low risk groups. (D) Survival curves of the LIHC patients with different risk scores. (E) Time-dependent ROC curves of the risk model for the 1-, 2- and 3-year survival.

Fig. 12. Cox analysis of the 4 mRNAs signature in 2 testing cohorts. (A-B) Unit and Multivariate Cox regression analysis of characteristics and risk score of the HCC patients in testing cohort. (C-D) Unit and Multivariate Cox regression analysis of characteristics and risk score of the HCC patients in the entire set.

Figures

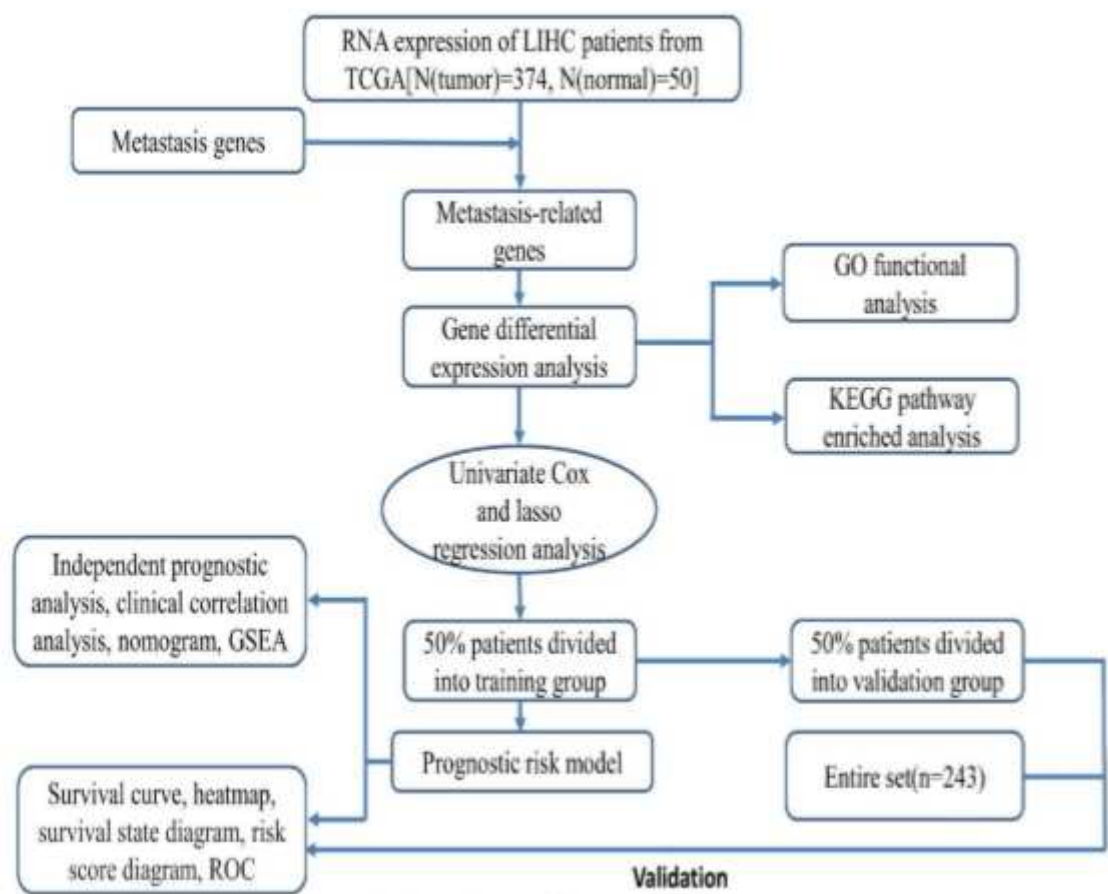


Figure 1

Overall study progress

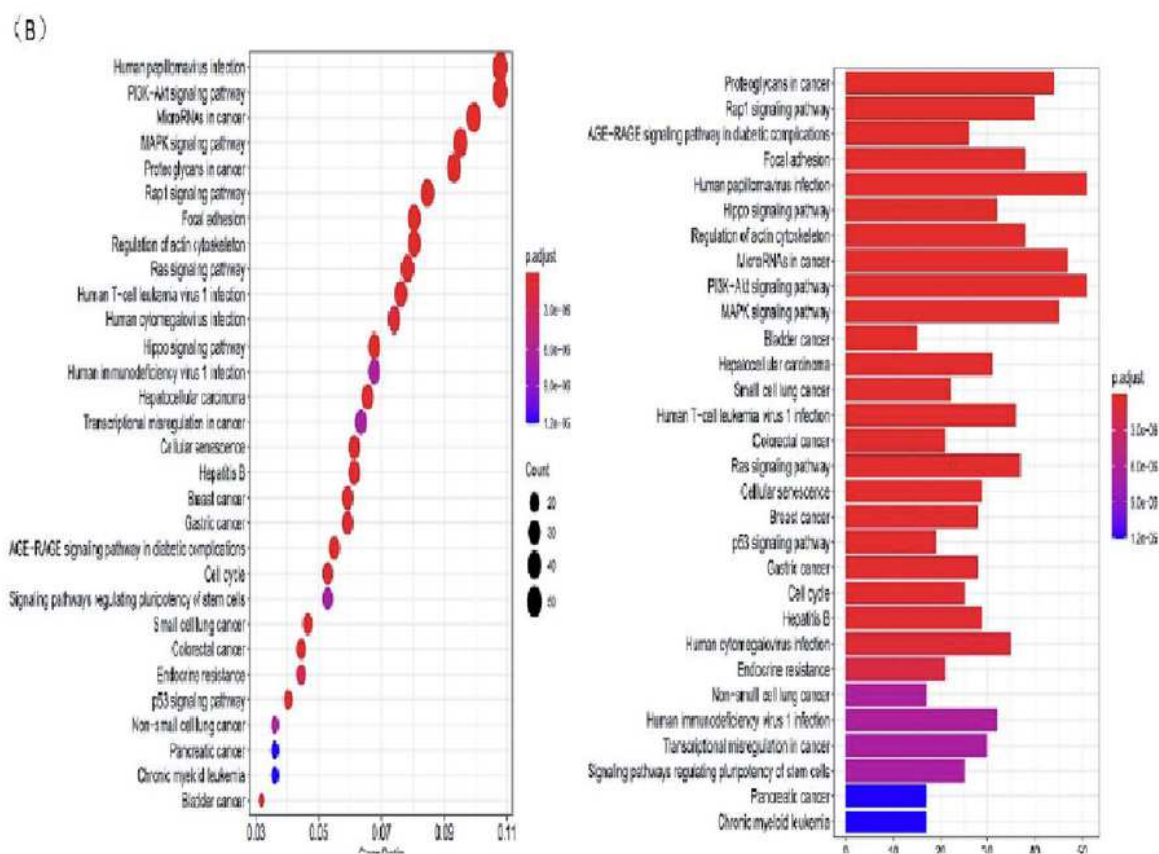
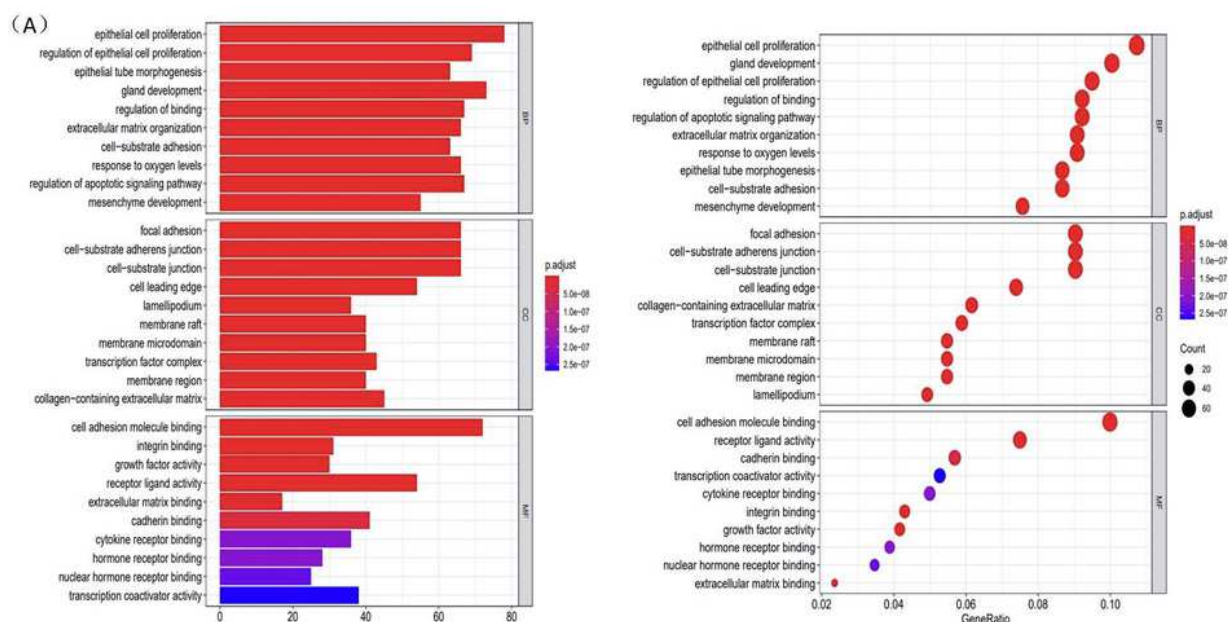


Figure 2

Gene functional enrichment of differentially expressed IRGs. (A) Gene ontology analysis; (B) The Kyoto Encyclopedia of MRNs and Genomes analysis.

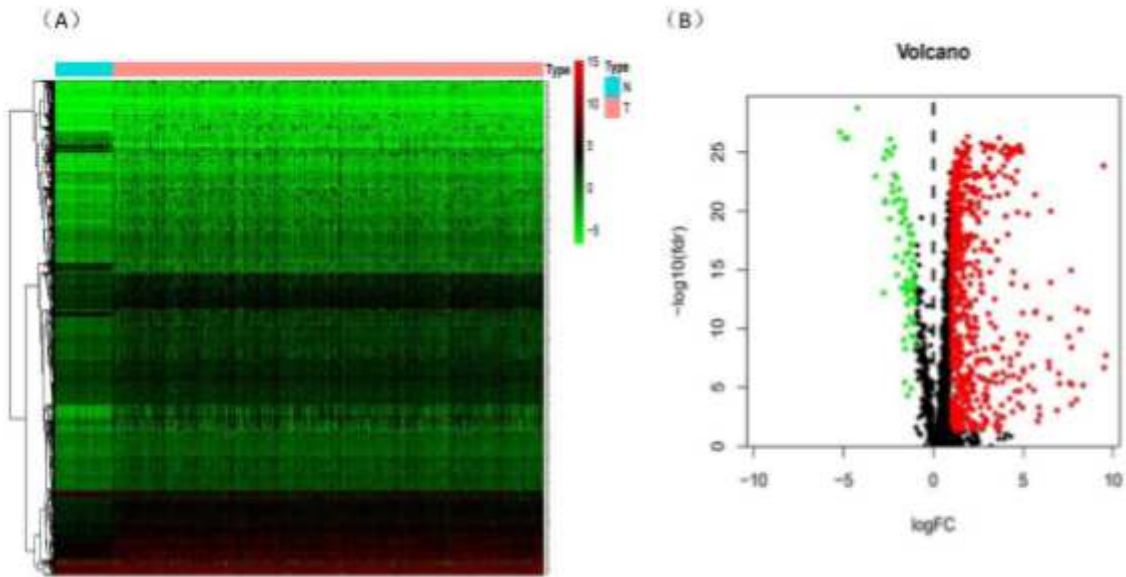


Figure 3

Identification of differentially expressed metastasis-related mRNAs. (A-B) Heatmap and volcano plot of differentially expressed mRNAs in EOC based on data from TCGA.

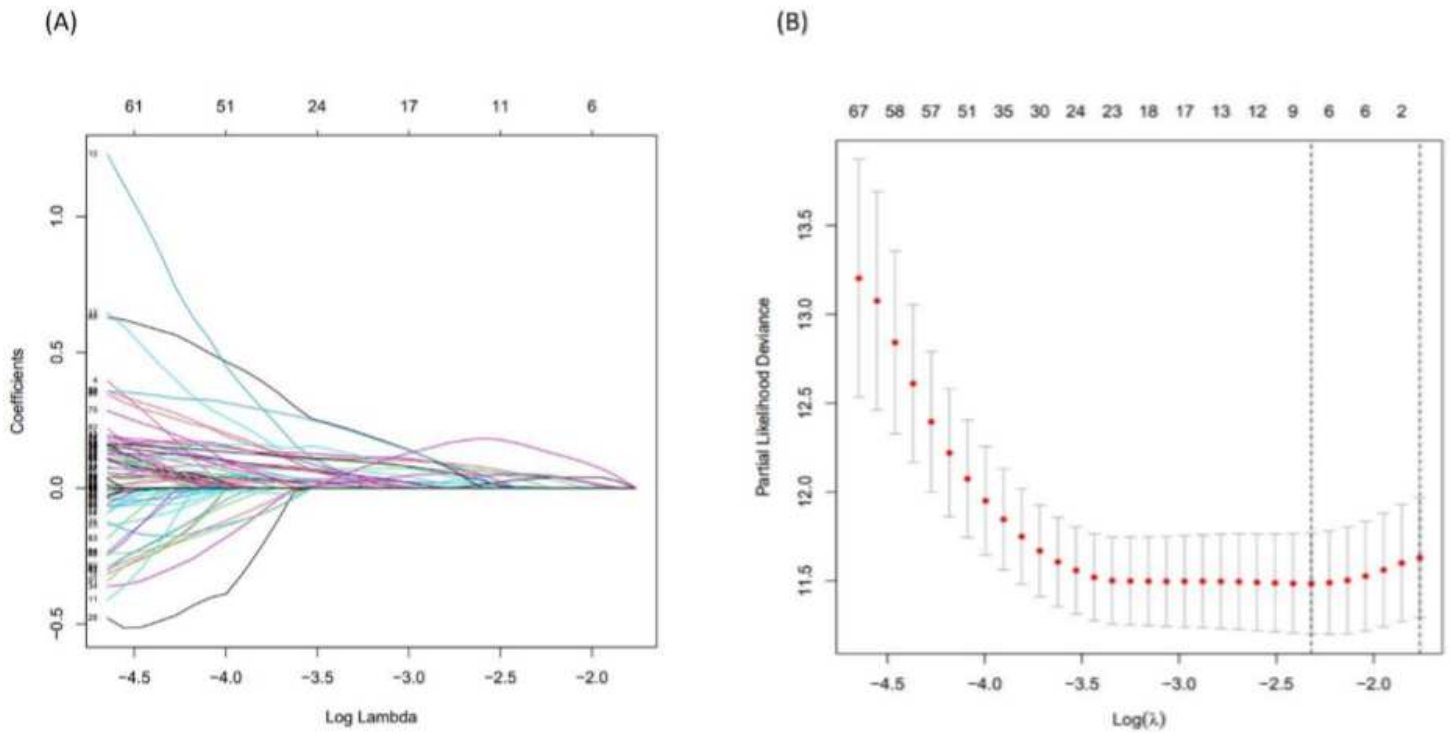


Figure 4

(A-B) The coefficients calculated by LASSO.

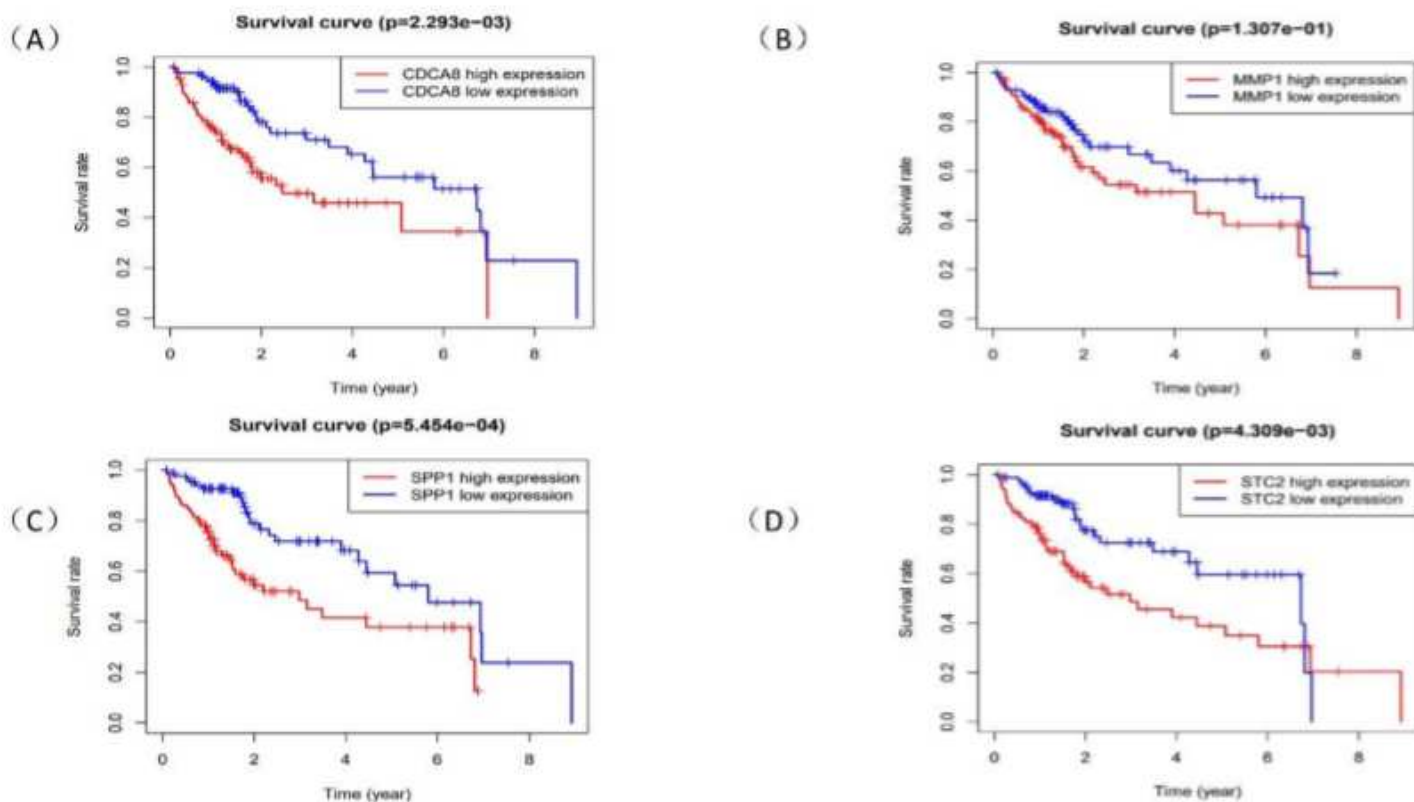


Figure 5

Prognostic gene survival curve. (A) MMP1 (B) SPP1. (C) STC2. (D) CDCA8

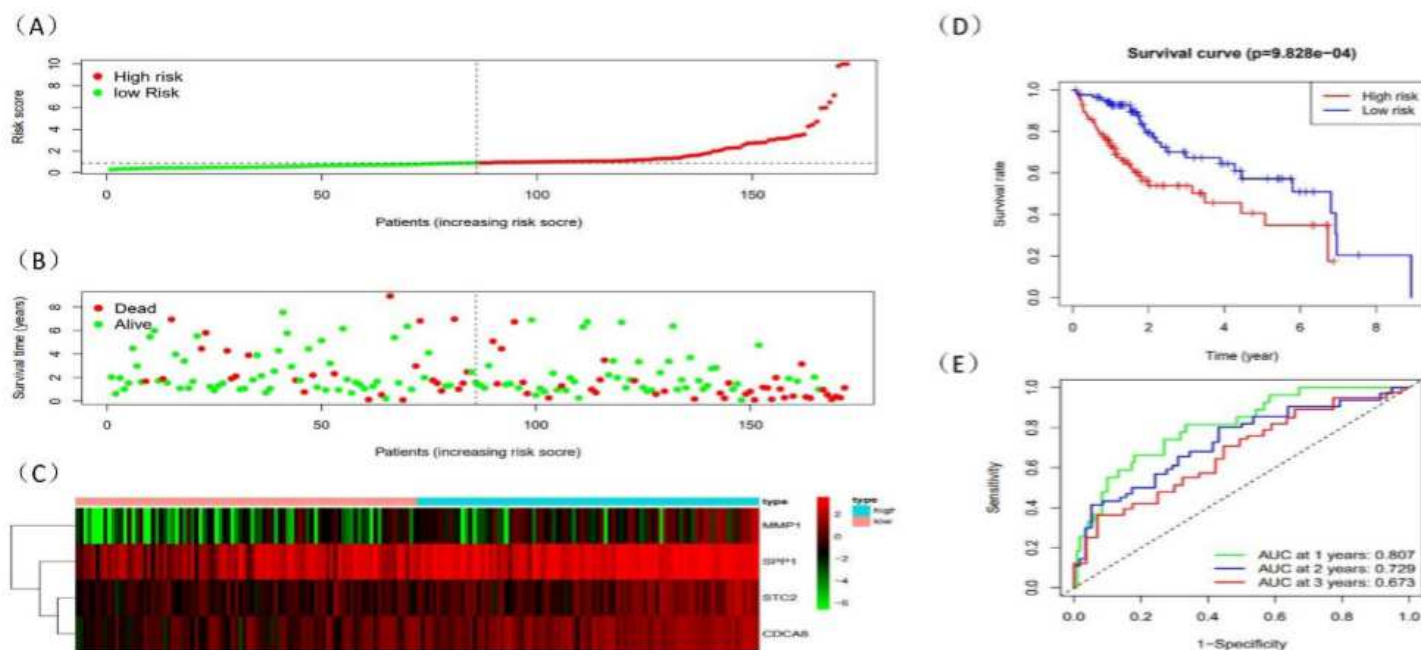


Figure 6

The prognostic results of risk score in train dataset. (A) Rank of prognostic index and distribution of high and low risk groups. (B) Survival status of patients in different groups. (C) Heatmap of expression

profiles of included mRNAs in high and low risk groups. (D) Survival curves of the LIHC patients with different risk scores. (E) Time-dependent ROC curves of the risk model for the 1-, 2- and 3-year survival.

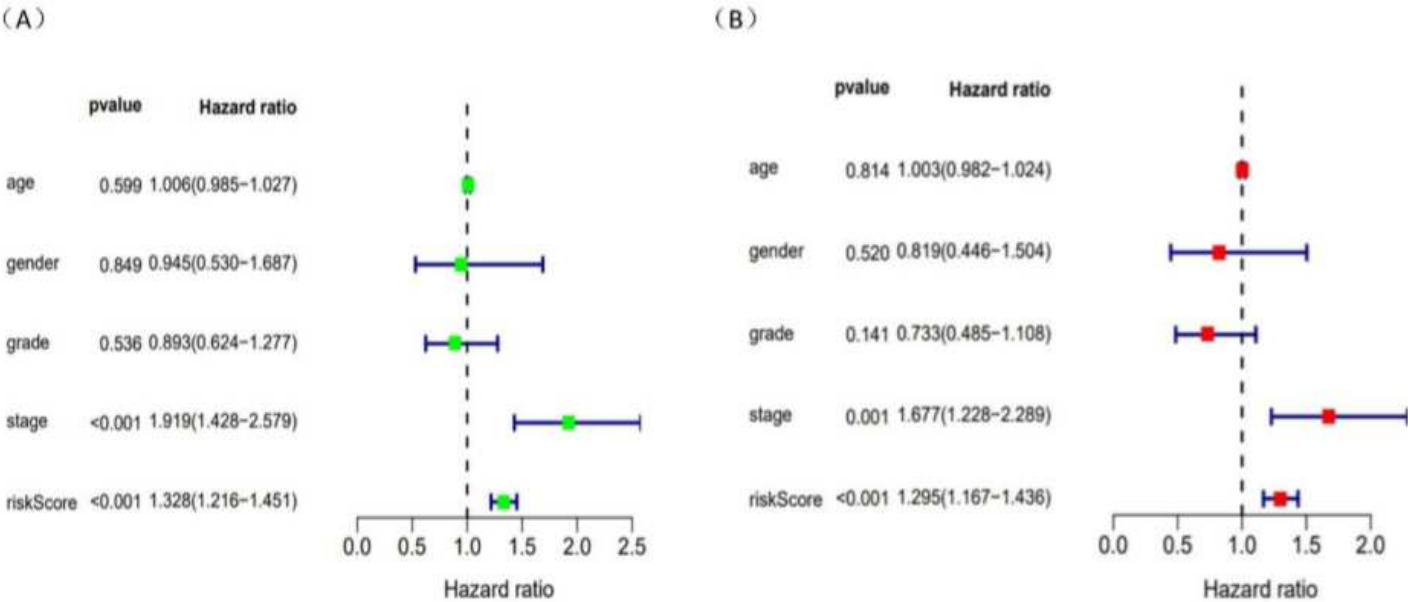


Figure 7

Cox analysis of the 4 mRNAs signature. (A)Univariate Cox regression analysis of characteristics and risk score of the LIHC patients in the train dataset. (B)Multivariate Cox regression analysis of characteristics and risk score of the LIHC patients in the train dataset.

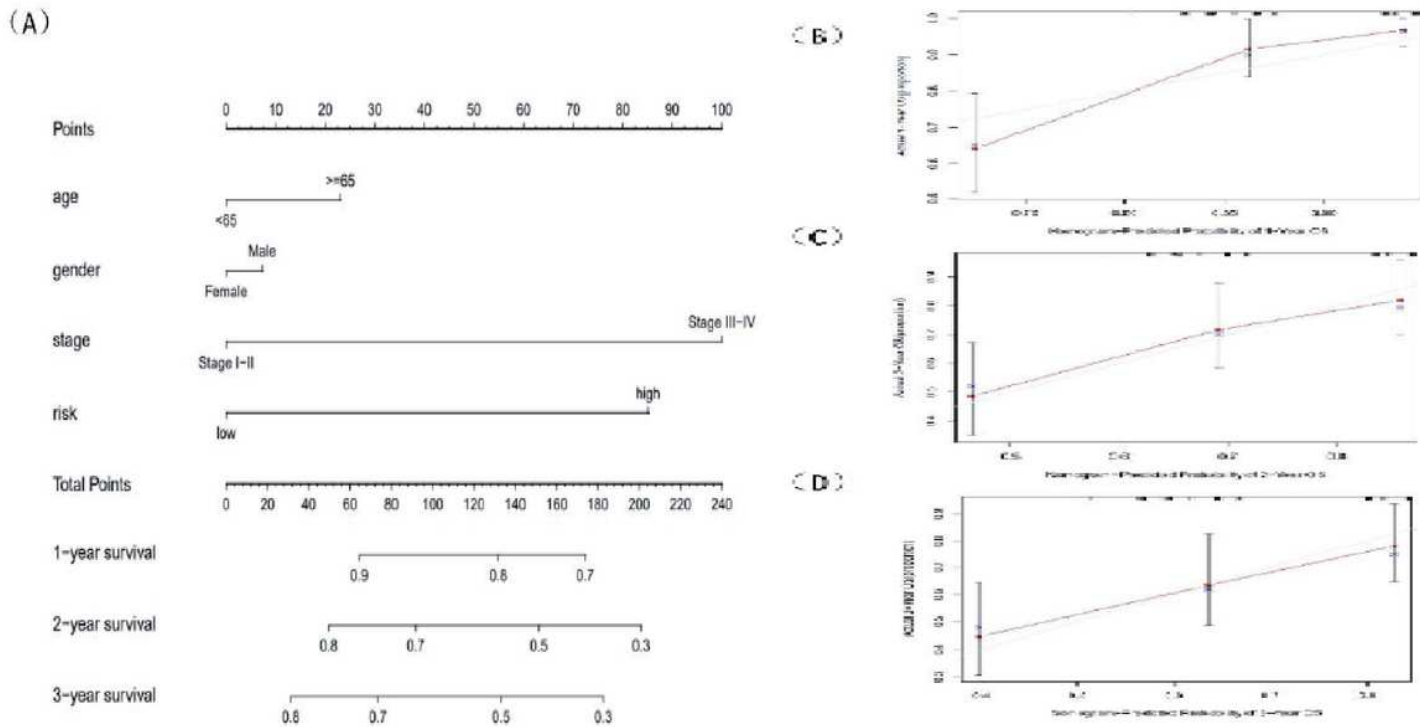


Figure 8

The nomogram to predict 1, 2, and 3-year OS and prognostic value of 4 mRNAs in the training cohort. (A) Nomogram for OS at 1-, 2- and 3-year in LIHC patients. (B-D) Calibration plot at 1-, 2- and 3-year for validation to predict the probability of OS.

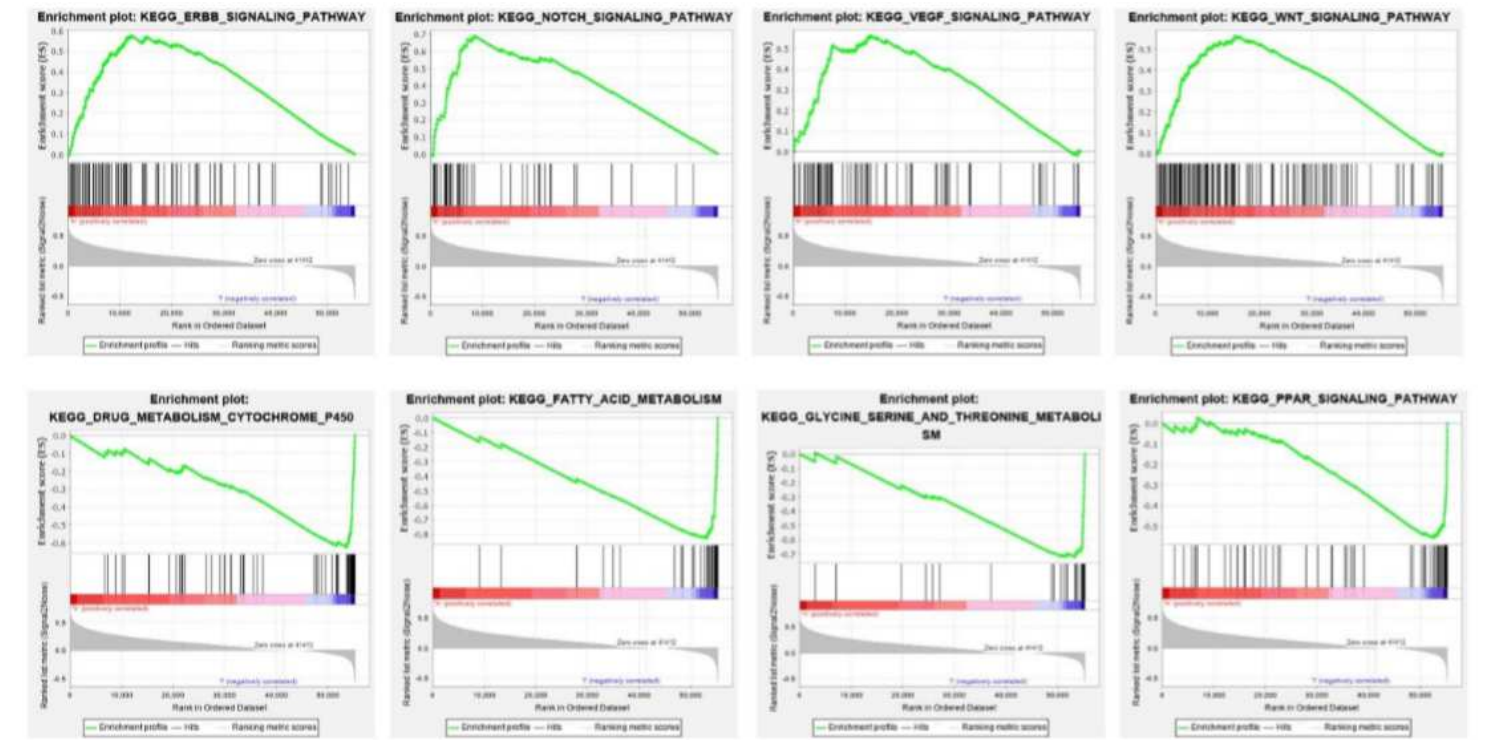


Figure 9

Gene Set Enrichment Analysis in TCGA database. Enrichment Map were used for visualization of the GSEA results.

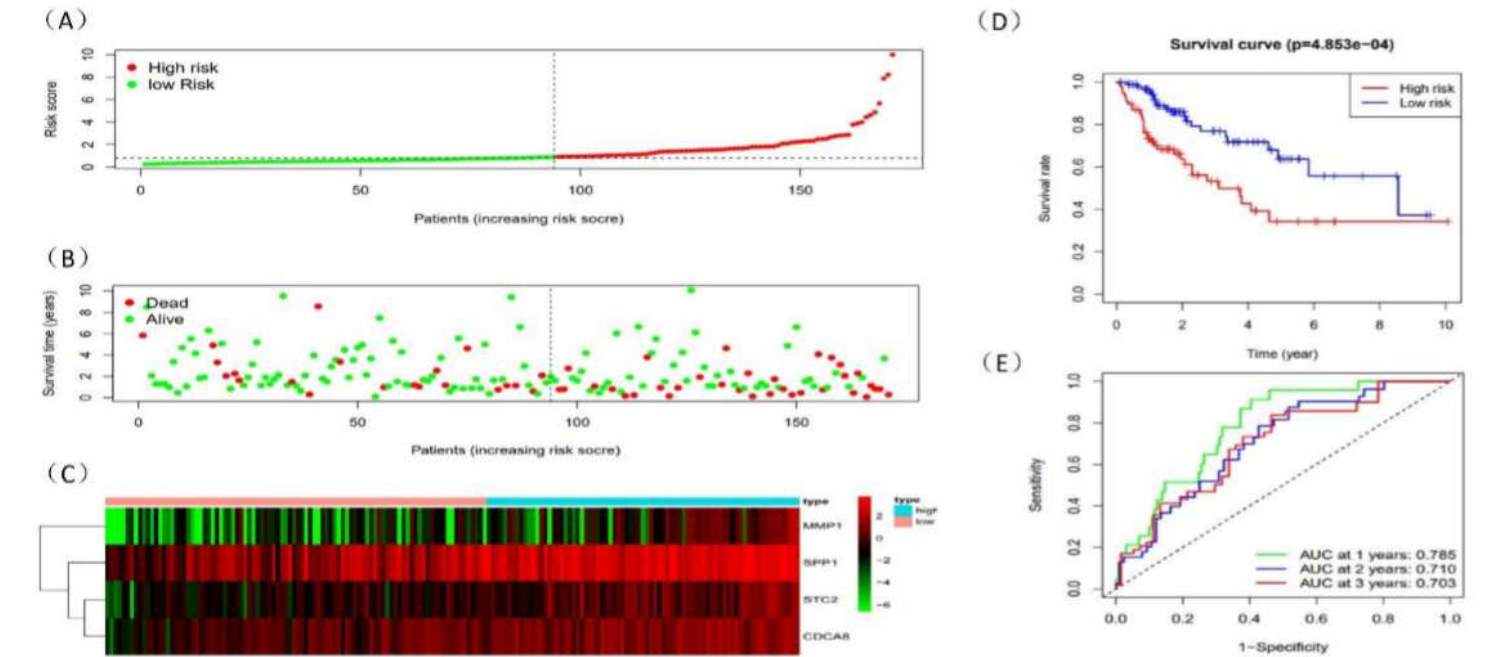


Figure 10

The prognostic results of risk score in TCGA testing dataset. (A) Rank of prognostic index and distribution of high and low risk groups. (B) Survival status of patients in different groups. (C) Heatmap of expression profiles of included mRNAs in high and low risk groups. (D) Survival curves of the LIHC patients with different risk scores. (E) Time-dependent ROC curves of the risk model for the 1-, 2- and 3-year survival.

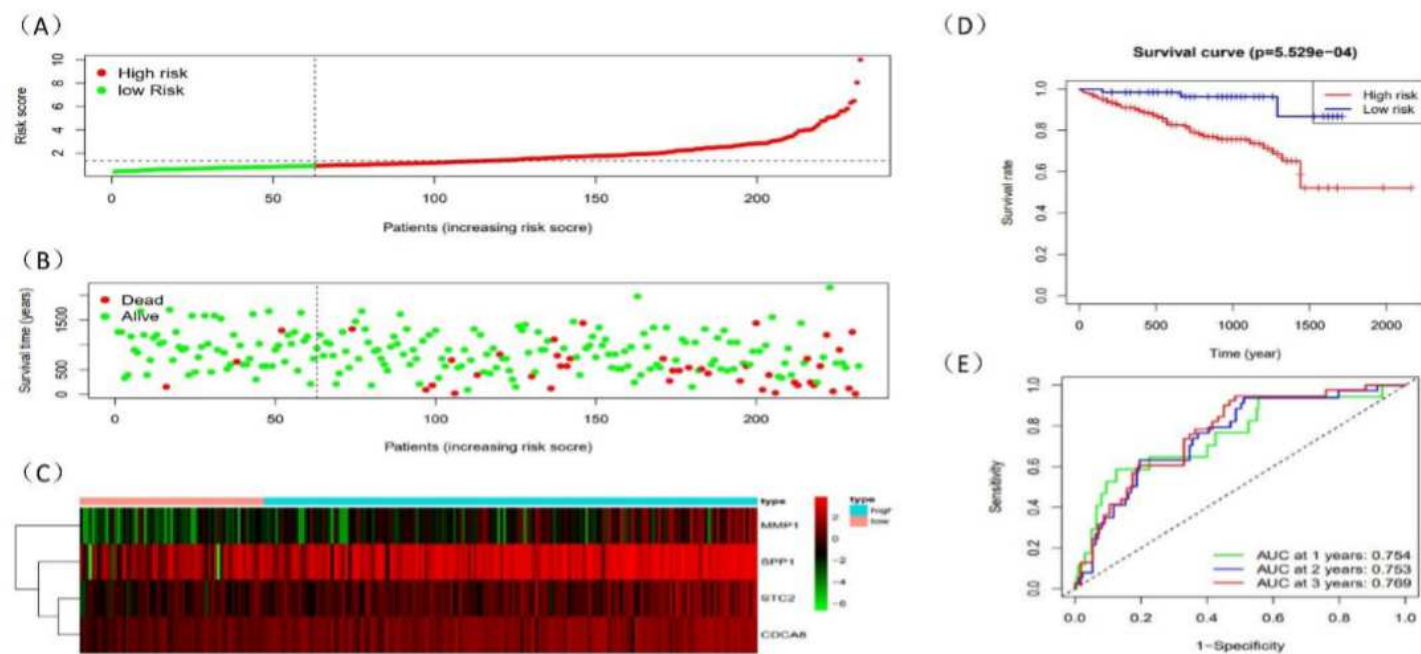


Figure 11

The prognostic results of risk score in the entire set. (A) Rank of prognostic index and distribution of high and low risk groups. (B) Survival status of patients in different groups. (C) Heatmap of expression profiles of included mRNAs in high and low risk groups. (D) Survival curves of the LIHC patients with different risk scores. (E) Time-dependent ROC curves of the risk model for the 1-, 2- and 3-year survival.

Fig. 12. Cox analysis of the 4 mRNAs signature in 2 testing cohorts. (A-B) Unit and Multivariate Cox regression analysis of characteristics and risk score of the HCC patients in testing cohort. (C-D) Unit and Multivariate Cox regression analysis of characteristics and risk score of the HCC patients in the entire set.

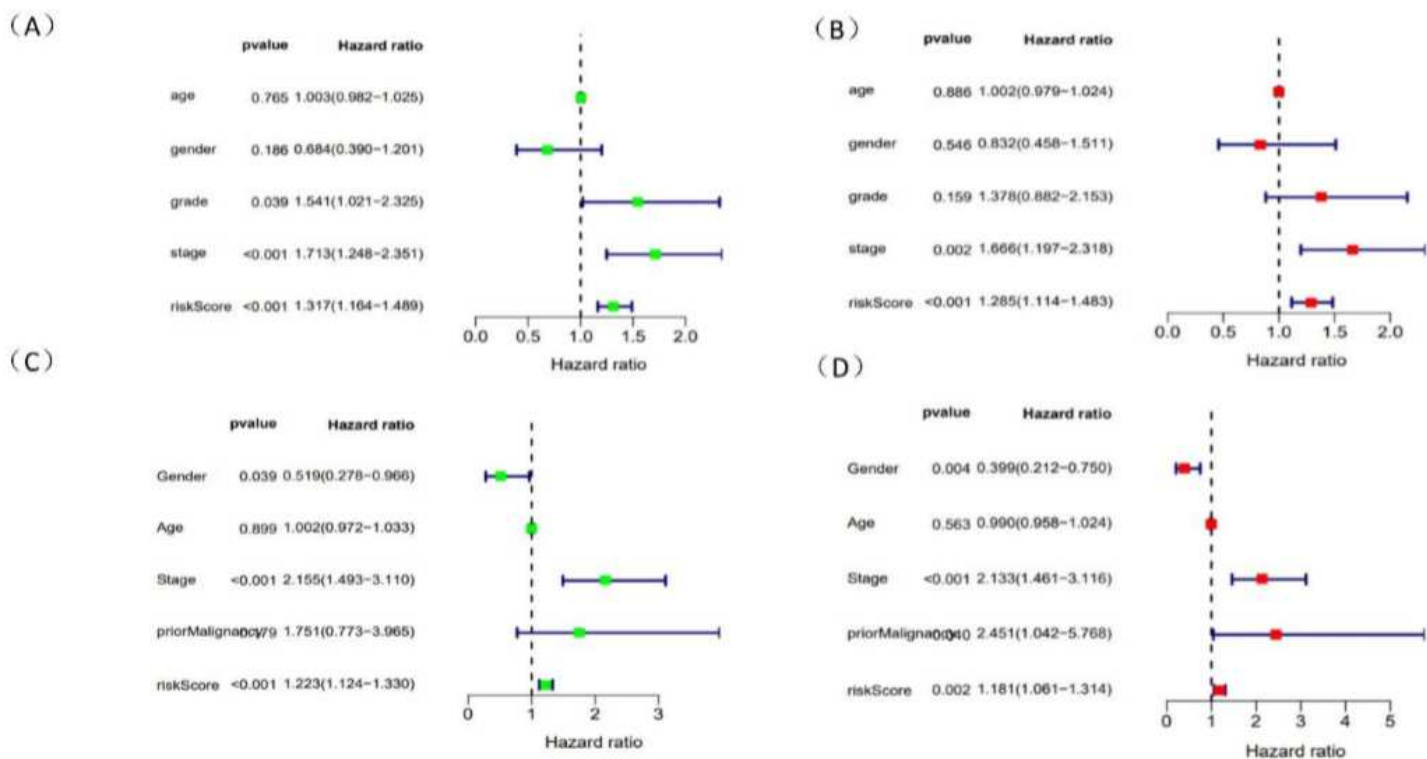


Figure 12

Cox analysis of the 4 mRNAs signature in 2 testing cohorts. (A-B) Unit and Multivariate Cox regression analysis of characteristics and risk score of the HCC patients in testing cohort. (C-D) Unit and Multivariate Cox regression analysis of characteristics and risk score of the HCC patients in the entire set.