

A three-mRNA status risk score has greater predictive ability compared with a lncRNA-based risk score for predicting prognosis in patients with hepatocellular carcinoma

WENXIA ZHANG¹, QIANG FU² and KANYU YAO³

¹Department of Hepatobiliary Surgery, The Affiliated Hospital of Inner Mongolia Medical University, Hohhot, Inner Mongolia 010050; ²Department of General Surgery, Erenhot Hospital, Erenhot, Inner Mongolia 011100;

³Department of Emergency Surgery, The Affiliated Hospital of Inner Mongolia Medical University, Hohhot, Inner Mongolia 010050, P.R China

Received April 24, 2019; Accepted October 25, 2019

DOI: 10.3892/ol.2020.11911

Abstract. Hepatocellular carcinoma (HCC) represents the fifth most common cause of cancer-associated mortality in men, and the seventh in women, worldwide. The aim of the present study was to identify a reliable and robust RNA-based risk score for the survival prediction of patients with hepatocellular carcinoma (HCC). Gene expression data from HCC and healthy control samples were obtained from The Cancer Genome Atlas to screen differentially expressed mRNAs and long non-coding RNAs (lncRNAs). Univariate and multivariate Cox proportional-hazards regression models and the LASSO algorithm for the Cox proportional-hazards model (LASSO Cox-PH model) were used to identify the prognostic mRNAs and lncRNAs among differentially expressed mRNAs (DEMs) and differentially expressed lncRNAs (DELs), respectively. Prognostic risk scores were generated based on the expression level or status of the prognostic lncRNAs and mRNAs, and the predictive abilities of these RNAs in TCGA and validation datasets were compared. Functional enrichment analyses were also performed. The results revealed a total of 154 downregulated and 625 upregulated mRNAs and 18 upregulated lncRNAs between tumor and control samples in TCGA dataset. A three-mRNA and a five-lncRNA expression signatures were identified using the LASSO Cox-PH model. Three-mRNA and five-lncRNA expression and status risk scores were generated. Using likelihood ratio P-values and area under the curve values from TCGA and the validation datasets, the three-mRNA status risk score was more accurate

compared with the other risk scores in predicting the mortality of patients with HCC. The three identified mRNAs, including hepatitis A virus cellular receptor 1, MYCN proto-oncogene BHLH transcription factor and stratifin, were associated with the cell cycle and oocyte maturation pathways. Therefore, a three-mRNA status risk score may be valuable and robust for risk stratification of patients with HCC. The three-mRNA status risk score exhibited greater prognostic value compared with the lncRNA-based risk score.

Introduction

Every year, >500,000 new cases of hepatocellular carcinoma (HCC) are diagnosed, and HCC is the fifth most prevalent cause of cancer-related mortality in men, and the seventh in women, worldwide (1,2). The poor outcome is partly caused by the clinical and molecular heterogeneity of HCC (3). Therefore, effective prognostic biomarkers are required to identify high-risk patients and guide personalized therapy.

Previous studies have demonstrated that long non-coding RNAs (lncRNAs), which are non-coding RNAs comprising >200 nucleotides, are crucial for HCC initiation and progression (4,5). A number of lncRNA signature-based models have been reported for the prognosis prediction of patients with HCC (6-8). In contrast with lncRNA signatures, prognostic mRNA signatures are more likely to be generated due to the large number of mRNAs. Li *et al* (9) have proposed a three-gene prognostic expression signature based on HCC data from The Cancer Genome Atlas (TCGA). In addition, an eight-gene signature comprising five mRNAs and three lncRNAs has been demonstrated to be efficient for predicting prognosis in patients with HCC (10). However, no currently available studies focus on the comparative analysis of mRNA and lncRNA signatures and their prognostic value for patients with HCC.

With an aim of improving the diagnosis, treatment and prevention of tumors, TCGA provides publicly available datasets for >30 types of human tumors (11). Using TCGA data, the present study aimed to identify mRNA and lncRNA signatures for HCC and explore potential mechanisms underlying HCC progression.

Correspondence to: Dr Kanyu Yao, Department of Emergency Surgery, The Affiliated Hospital of Inner Mongolia Medical University, 1 Tongdaobei Street, Hohhot, Inner Mongolia 010050, P.R. China
E-mail: yaokanyu@sina.cn

Key words: hepatocellular carcinoma, long non-coding RNA, mRNA, pathway, risk score

Materials and methods

Data source and preprocessing. A search for HCC microarray gene expression data in TCGA (<https://gdc-portal.nci.nih.gov/>) data portal retrieved data from 373 HCC samples and 50 healthy control samples obtained by Illumina HiSeq 2000 RNA Sequencing. Among them, clinical information was available for 366 patients with HCC, which were used in the analysis.

The E-TABM-36 dataset (12) was downloaded from the European Bioinformatics Institute (EBI) array database (<https://www.ebi.ac.uk/arrayexpress/>), which is based on GPL 96 [HG-U133A] Affymetrix GeneChip Human Genome HG-U133A and comprises 65 samples. Of these, 44 samples had overall survival (OS) information and were used as the validation dataset. For data preprocessing, median scale normalization and robust scale normalization approaches were used as described in the supplementary data of a previous study (13).

The HUGO Gene Nomenclature Committee (HGNC) database (<http://www.genenames.org/>) that contains 4,076 lncRNAs and 19,194 protein-coding genes (14) was used to identify the lncRNAs and mRNAs in TCGA and E-TABM-36 datasets.

Identification of differentially expressed lncRNAs and mRNAs. The Limma package in R (v. 3.34.7; <https://bioconductor.org/packages/release/bioc/html/limma.html>) was used to select differentially expressed lncRNAs (DElNs) and mRNAs (DEMs) between tumor and control samples from TCGA dataset. To determine whether the gene expression difference was significant, the threshold values of false discovery rate (FDR) <0.05 and \log_2 fold-change (FC) >1 were used. The pheatmap package (15) (version 1.0.8) of R was applied to perform two-way hierarchical clustering analysis based on the centered Pearson correlation algorithm for the identified DEMs and DElNs.

Development of prognosis prediction models. Using the survival package in R (16) (v. 2.41-1; <http://bioconductor.org/packages/survival/>), univariate and multivariate Cox proportional hazards regression analyses were performed to identify independent prognostic lncRNAs and mRNAs associated with OS among the aforementioned DElNs and DEMs. $P < 0.05$ was considered to indicate a statistically significant difference.

Subsequently, the L1-penalized (LASSO) Cox proportional-hazards (Cox-PH) regression model was fitted to the identified OS-associated lncRNAs and mRNAs using the penalized package (17) of R (v. 0.9.50, <https://cran.r-project.org/web/packages/penalized/index.html>). With the optimal λ value obtained through a 1,000 cross-validation check, signatures of optimal lncRNAs and mRNAs were determined.

Prognostic models to calculate risk scores were generated based on the expression status or levels of the prognostic lncRNAs or mRNAs and their Cox-PH regression coefficients. The expression status of a lncRNA or mRNA was defined according to an optimal expression cut-off value determined using the X-Tile tool (<https://medicine.yale.edu/lab/rimm/research/software.aspx>) based on patient

survival; patients with an RNA expression level greater than the cut-off value were assigned to the high expression group (expression status=1), whereas patients with an RNA expression level equal or lower than the cut-off value were assigned to the low expression group (expression status=0). Prognostic scoring models based on the expression status of prognostic lncRNAs or mRNAs were calculated as follows: Status Risk Score = $\sum \beta_{\text{RNA}_n} \times \text{Status}_{\text{RNA}_n}$; those based on the expression levels of prognostic lncRNAs or mRNAs were calculated as follows: Expression Risk Score = $\sum \beta_{\text{RNA}_n} \times \text{Exp}_{\text{RNA}_n}$, where β_{RNA_n} indicates the Cox-PH coefficient of an RNA and $\text{Status}_{\text{RNA}_n}$ or $\text{Exp}_{\text{RNA}_n}$ indicates the expression status or level of an RNA, respectively.

Based on the median risk score calculated using a prognostic scoring model as the cut-off value, patients in the TCGA and the E-TABM-36 datasets were classified into high- and low-risk groups.

Establishment of a nomogram. OS of different risk groups was estimated using Kaplan-Meier plots in the survival package of R (v. 2.41-1; <http://bioconductor.org/packages/survival/>) (18) and the weighted likelihood ratio test (19). The discriminative ability of different risk scores was evaluated and compared using receiver-operating characteristic (ROC) curve analyses (20) by calculating the area under the curve (AUC) using SPSS software (v. 24.0; IBM Corp.). Univariate and multivariate Cox regression analyses based on the survival package in R (v. 2.41-1; <http://bioconductor.org/packages/survival/>) (21) were performed to determine independent prognostic factors associated with OS, including clinical features as covariates. To further investigate the association between the independent prognostic factors and survival prognosis, nomograms incorporating risk score status and identified prognostic clinical factors were produced using the rms package (22) of R (v. 5.1-2; <https://cran.r-project.org/web/packages/rms/index.html>). Comparative analysis of actual and predicted 1- and 3-year survival probability were conducted using calibration plots (23). $P < 0.05$ was considered indicate a statistically significant difference.

Functional annotation. TCGA dataset was divided into high- and low-risk groups by the optimal risk score, and differentially expressed genes (DEGs) between the two groups were screened using FDR <0.05 and \log_2 FC >0.263 as strict threshold values. The identified DEGs were analyzed by Gene Ontology (GO) function and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses using the clusterProfiler package (24) (v. 3.6.0) of R. $P < 0.05$ was considered to indicate a statistically significant difference.

Results

DEMs and DElNs between HCC and healthy control samples. Following annotation based on the HGNC database, a total of 18,497 mRNAs and 2,528 lncRNAs were obtained in TCGA and the E-TABM-36 datasets. Comparison between the HCC (n=373) and normal control (n=50) samples from TCGA dataset revealed 154 downregulated and 625 upregulated mRNAs, as well as 18 upregulated lncRNAs in HCC samples, which resulted in a total of 797 DEGs (Fig. 1A). Two-way hierarchical

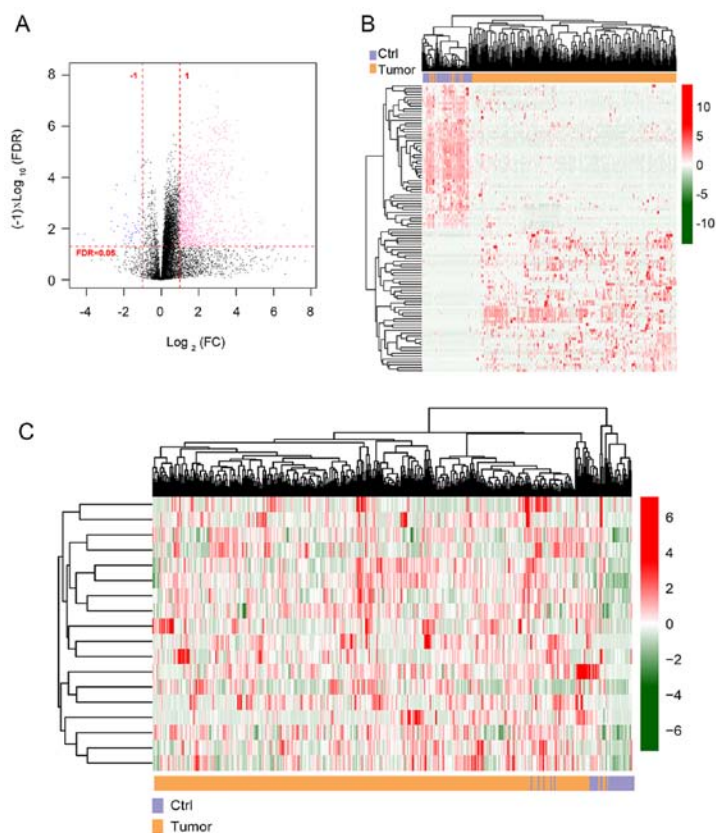


Figure 1. Differentially expressed RNAs between tumor and control samples in The Cancer Genome Atlas dataset. (A) Volcano plot of effect size ($\log_2\text{FC}$) and $-\log_{10}(\text{FDR})$ of differentially expressed RNAs. Downregulated and upregulated RNAs are represented by blue and pink dots, respectively. Black dots represent RNAs that were not differentially expressed between tumor and control samples. Vertical dashed lines indicate $|\log_2\text{FC}| > 1$; horizontal dashed line indicates $\text{FDR} < 0.05$. (B) Heatmap of tumor and control sample clustering based on differentially expressed mRNAs. (C) Heatmap of tumor and control sample clustering based on differentially expressed long non-coding RNAs. Horizontal axis indicates samples and vertical axis indicates expression levels of differentially expressed RNAs. Ctrl, control samples; tumor, hepatocellular carcinoma samples; FC, fold-change; FDR, false discovery rate.

clustering of the identified DERs demonstrated that the expression patterns of the DEMs and DELs were different between tumor and healthy control samples (Fig. 1B and C).

Three-mRNA status risk score exhibits the best results in predicting the mortality of patients with HCC. To identify the prognostic lncRNA and mRNA signatures, univariate and multivariate Cox regression analyses of the identified DELs and DEMs were performed. In univariate Cox regression analysis, 110 mRNAs and 13 lncRNAs were significantly associated with OS ($P < 0.05$, Table SI). In addition, 14 mRNAs and 8 lncRNAs were identified to be independent prognostic factors by multivariate Cox regression analysis (Table SII), and were used to fit the LASSO Cox-PH regression model. As a result, a five-lncRNA signature ($\lambda = 7.283$) and a three-mRNA signature ($\lambda = 5.646$) were obtained (Table I). The five-lncRNA signature included HLA complex group 4 (HCG4), nuclear paraspeckle assembly transcript 1 (NEAT1), RFPL1 antisense RNA 1 (RFPL1S), chromobox 2 (CBX2) and cell division cycle-associated 8 (CDCA8), whereas the three-mRNA signature comprised hepatitis A virus cellular receptor 1 (HAVCR1), MYCN proto-oncogene BHLH transcription factor (MYCN) and stratifin (SFN). These signatures were used to calculate risk scores based on their expression level or status as follows: mRNA Status Risk Score = $0.3839 \times \text{Status}_{\text{CBX2}} + 0.1625 \times \text{Status}_{\text{CDCA8}} + 0.2586 \times \text{Status}_{\text{HAVCR1}} + 0.2621 \times \text{Status}_{\text{MYCN}} +$

$0.0814 \times \text{Status}_{\text{SFN}}$; lncRNA Status Risk Score = $0.1919 \times \text{Status}_{\text{HCG4}} + 0.2966 \times \text{Status}_{\text{NEAT1}} - 0.1552 \times \text{Status}_{\text{RFPL1S}}$; mRNA Expression Risk Score = $0.3839 \times \text{Exp}_{\text{CBX2}} + 0.1625 \times \text{Exp}_{\text{CDCA8}} + 0.2586 \times \text{Exp}_{\text{HAVCR1}} + 0.2621 \times \text{Exp}_{\text{MYCN}} + 0.0814 \times \text{Exp}_{\text{SFN}}$; lncRNA Expression Risk Score = $0.1919 \times \text{Exp}_{\text{HCG4}} + 0.2966 \times \text{Exp}_{\text{NEAT1}} - 0.1552 \times \text{Exp}_{\text{RFPL1S}}$.

TCGA and the E-TABM-36 datasets were classified into high- and low-risk groups based on the status and expression risk scores of the three-mRNA and five-lncRNA signatures separately (Figs. 2 and 3). The discriminatory power of the four risk scores for patients with HCC was evaluated and compared using Kaplan-Meier analysis for OS and ROC curves (Table II; Figs. 2 and 3).

The four risk scores split TCGA dataset into two risk groups with significantly different OS rate ($P < 0.05$). However, only the three-mRNA status risk score exhibited significant prognostic value for patients with HCC in the E-TABM-36 dataset ($P = 0.031$). Among the four risk scores, the three-mRNA status risk score exhibited the lowest P-values in the Kaplan-Meier OS analysis (TCGA, $P < 0.001$; E-TABM-36, $P = 0.031$). Since the results suggested that the three-mRNA status risk score exhibited the greatest predictive value for HCC, it was selected for further analyses. Notably, HAVCR1 ($\log_2\text{FC} = 4.2553$; $\text{FDR} = 4.34 \times 10^{-14}$), MYCN ($\log_2\text{FC} = 3.1756$; $\text{FDR} = 1.59 \times 10^{-14}$) and SFN ($\log_2\text{FC} = 2.08$; $\text{FDR} = 7.05 \times 10^{-14}$) were upregulated in HCC samples compared with normal control samples.

Table I. Characteristics of the five-lncRNA signature and the three-mRNA signature.

A, Five-lncRNA signature					
ID	β_{RNA}	Hazard ratio	95% CI	P-value	Cut-off
HCG4	0.1919	1.3055	1.0997-1.7100	P=5.30x10 ⁻³	-0.54
NEAT1	0.2966	1.2743	1.1713-1.8640	P=2.12x10 ⁻²	-0.91
RFPL1S	-0.1552	0.7951	0.6477-0.9760	P=2.80x10 ⁻³	0.15
CBX2	0.3839	3.2933	1.7548-6.1810	P=2.00x10 ⁻⁴	0.67
CDCA8	0.1625	5.1601	1.3476-19.7600	P=1.66x10 ⁻²	0.91
B, Three-mRNA signature					
ID	β_{RNA}	Hazard ratio	95% CI	P-value	Cut-off
HAVCR1	0.2586	2.1561	1.4016-3.3170	P=5.00x10 ⁻⁴	1.13
MYCN	0.2621	1.9775	1.3593-2.8770	P=4.00x10 ⁻⁴	1.69
SFN	0.0814	1.2379	1.0198-1.5030	P=3.09x10 ⁻²	0.92

lncRNA, long non-coding RNA; HCG4, HLA complex group 4; NEAT1, nuclear paraspeckle assembly transcript 1; RFPL1S, RFPL1 antisense RNA 1; CBX2, chromobox 2; CDCA8, cell division cycle-associated 8; HAVCR1, hepatitis A virus cellular receptor 1; MYCN, MYCN proto-oncogene BHLH transcription factor; SFN, stratifin; β_{RNA} , Cox proportional hazard coefficient; CI, confidence interval.

Establishment of a nomogram combining three-mRNA status risk score, pathological T stage and pathological stage. Univariate Cox regression analysis was performed to select the prognostic clinical factors. Pathological T stage, pathological stage and the three-mRNA status risk score were determined to be significantly associated with OS (P<0.05; Table III). As presented in Fig. 4, patients with HCC in TCGA dataset were successfully divided by pathological T stage (P<0.001) or pathological stage (P<0.001) into two groups with significantly different OS rates, respectively. To improve the predictive power of the three-mRNA status risk score, a nomogram based on the three-mRNA status risk score, pathological T stage and pathological stage was developed (Fig. 5A). Calibration plots demonstrated that good concordance was achieved between the predicted and actual 3- and 5-year OS probabilities (Fig. 5B).

Functional annotation of the three-mRNA signature. To determine the functional roles of the three-mRNA signature in the biology of HCC, DEGs were identified between the high- and low-risk patient groups predicted by the three-mRNA status risk score in TCGA dataset. A total of 91 downregulated and 1,300 upregulated DEGs were identified. These DEGs were functionally enriched in GO biological processes associated with the cell cycle, cell proliferation and intracellular signaling cascades (Tables IV and V). In addition, five significant KEGG pathways were identified, including ‘cell cycle progression’, ‘DNA replication’, ‘cytokine-cytokine receptor interaction’, ‘progesterone-mediated oocyte maturation’ and ‘oocyte meiosis’ (Table IV).

Discussion

HCC is a lethal cancer that accounts for 85-90% of primary liver cancers, with risk factors including hepatitis B or C

infection, alcoholic liver disease and non-alcoholic fatty liver disease (25). Considerable efforts should be devoted to developing an effective stratification model for the prognosis of patients with HCC. The current study utilized a multi-step exploration and validation strategy to identify a three-mRNA and a five-lncRNA expression signatures. Consequently, three-mRNA and five-lncRNA expression and status risk scores were developed and compared for prognostic value. According to the likelihood ratio P-values and AUC, the three-mRNA status risk score was the optimal score for OS stratification of patients with HCC. In addition, the validation dataset confirmed the prognostic value of the three-mRNA status risk score for patients with HCC. To the best of our knowledge, the present study is the first to comparatively analyze mRNA and lncRNA signatures for predicting the mortality of patients with HCC. The results of the present study suggested that the three-mRNA status risk score may be an effective and robust risk stratification tool for HCC.

The identified three-mRNA expression signature comprised HAVCR1, MYCN and SFN. HAVCR1, also known as T-cell immunoglobulin and mucin domain (TIM)-1 or kidney injury molecule (KIM)-1, is a member of the TIM gene family, which participates in the regulation of immune cell activity and kidney regeneration (26). HAVCR1 has been demonstrated to serve a role in renal cell and ovarian clear cell carcinoma progression by reducing the expression of tight junction molecules (26). A recent study has reported that HAVCR1 is upregulated at the mRNA and protein levels in gastric cancer tissues and may be a prognostic biomarker for patients with gastric cancer (27). The MYCN gene encodes the N-myc proto-oncogene protein (c-Myc), which belongs to the MYC family of transcription factors and serves an oncogenic role in various types of tumors (28,29). c-Myc is overexpressed in the majority of HCC cases and may be a potential therapeutic

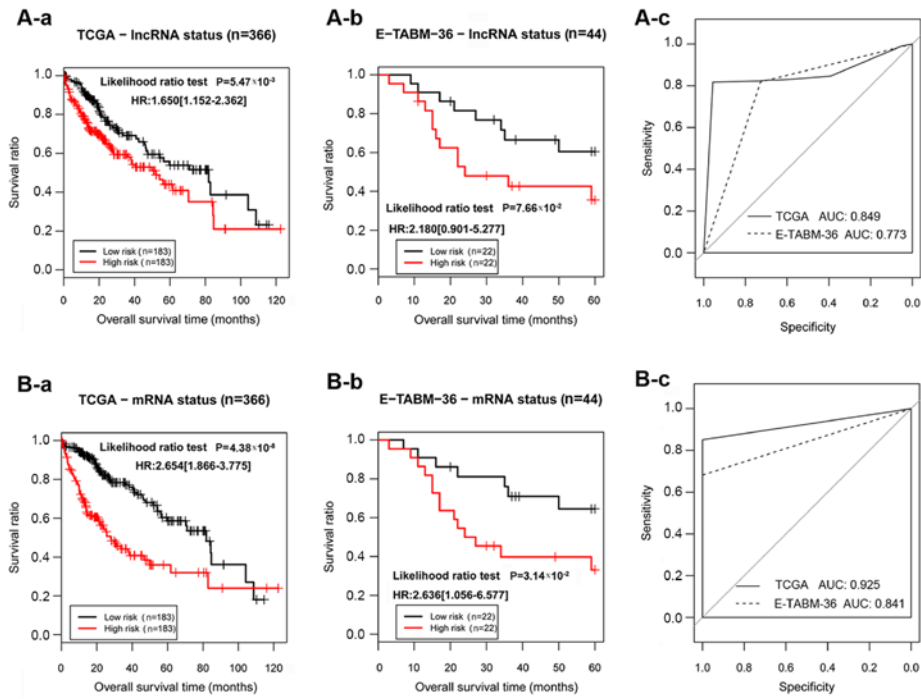


Figure 2. Stratification of patients with hepatocellular carcinoma in TCGA and the E-TABM-36 datasets using the five-lncRNA status risk score or three-mRNA status risk score. (A-a) Kaplan-Meier (KM) overall survival curves for high- and low-risk patients in TCGA; (A-b) KM overall survival curves in E-TABM-36; (A-c) receiver-operating characteristic curves using five-lncRNA status risk score. (B-a) Kaplan-Meier (KM) overall survival curves for high- and low-risk patients in TCGA; (B-b) KM overall survival curves in E-TABM-36; (B-c) receiver-operating characteristic curves using three-mRNA status risk score. TCGA, The Cancer Genome Atlas; lncRNA, long non-coding RNA; AUC, area under the curve; HR, hazard ratio.

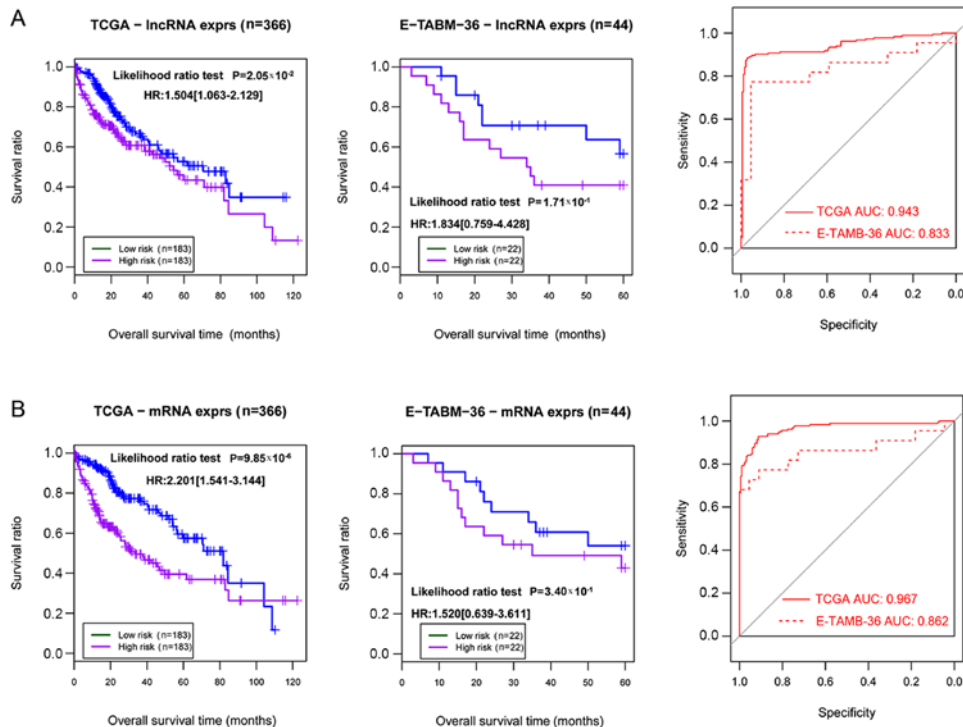


Figure 3. Stratification of patients with hepatocellular carcinoma in TCGA and the E-TABM-36 datasets using (A) the five-lncRNA expression risk score or (B) the three-mRNA expression risk score. Kaplan-Meier overall survival curves for TCGA (left) and the E-TABM-36 (middle) datasets and receiver-operating characteristic curves (right) are presented. TCGA, The Cancer Genome Atlas; expr's, expression level; lncRNA, long non-coding RNA; AUC, area under the curve; HR, hazard ratio.

target (30). Excessive activation of MYCN is observed in ~70% of HCC cases associated with human viruses and alcohol (31).

In addition, MYCN may be a prognostic biomarker and a therapeutic target of acyclic retinoid in liver cancer stem cells

Table II. Evaluation of the prognostic power of four risk scores based on lncRNAs or mRNAs.

Type	Expression risk score				Status risk score			
	Likelihood ratio P-value		AUC value		Likelihood ratio P-value		AUC value	
	Training set	Validation set	Training set	Validation set	Training set	Validation set	Training set	Validation set
lncRNA	P=2.05x10 ⁻²	P=1.71x10 ⁻¹	0.9430	0.8330	P=5.50x10 ⁻³	P=7.66x10 ⁻²	0.8490	0.7730
mRNA	P<0.0001	P=3.40x10 ⁻¹	0.9670	0.8620	P<0.0001	P=3.14x10 ⁻²	0.9250	0.8410

lncRNA, long non-coding RNA; AUC, area under the curve.

Table III. Identification of prognostic clinical factors in TCGA set.

Clinical characteristic	TCGA set (n=366)	Univariate Cox analysis	
		HR (95% CI)	P-value
Age, years (mean ± SD)	59.66±13.34	1.012 (0.999-1.026)	P=7.90x10 ⁻²
Gender (Male/Female)	247/119	0.816 (0.573-1.164)	P=2.62x10 ⁻¹
Pathologic M (M0/M1/-)	263/3/100	4.032 (0.267-12.83)	P=1.06x10 ⁻¹
Pathologic N (N0/N1/-)	248/4/114	2.004 (0.491-8.181)	P=3.83x10 ⁻¹
Pathologic T (T1/T2/T3/T4/-)	180/92/78/13/3	1.675 (1.397-2.007)	P<0.0001 ^a
Pathologic stage (I/II/III/IV/-)	170/85/83/4/24	1.661 (1.355-2.037)	P<0.0001 ^a
Histologic grade (G1/G2/G3/G4/-)	55/176/118/12/5	1.121 (0.887-1.416)	P=3.39x10 ⁻¹
Vascular invasion (Yes/No/-)	107/205/54	1.351 (0.892-2.047)	P=1.54x10 ⁻¹
Recurrence (Yes/No/-)	140/179/47	1.375 (0.914-2.068)	P=1.25x10 ⁻¹
RS status (High/Low)	183/183	2.177 (1.524-3.110)	P<0.0001 ^a
Survival status (Dead/Alive)	130/263	-	-
Overall survival time, months (mean ± SD)	27.23±24.31	-	-

^aP<0.05. HR, hazard ratio; CI, confidence index; TCGA, The Cancer Genome Atlas; RS, Risk score.

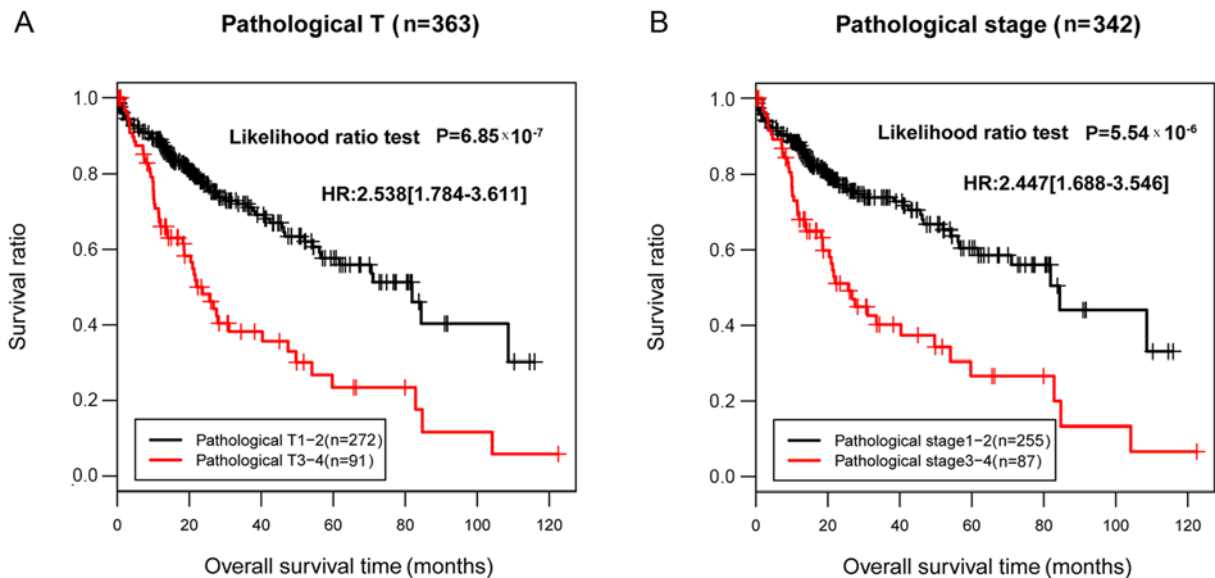


Figure 4. Overall survival of patients with HCC classified by (A) pathologic T stage and (B) pathologic stage in TCGA dataset. Patients with HCC (n=363) in TCGA dataset were classified by pathologic T stage into two subgroups (pathologic T 1-2 and 3-4) with significantly different OS rate (P=6.847x10⁻⁷). Similarly, the difference in OS between patients at pathologic stages 1-2 and 3-4 was significant (P=5.538x10⁻⁶). HCC, hepatocellular carcinoma; TCGA, The Cancer Genome Atlas; OS, overall survival; HR, hazard ratio.

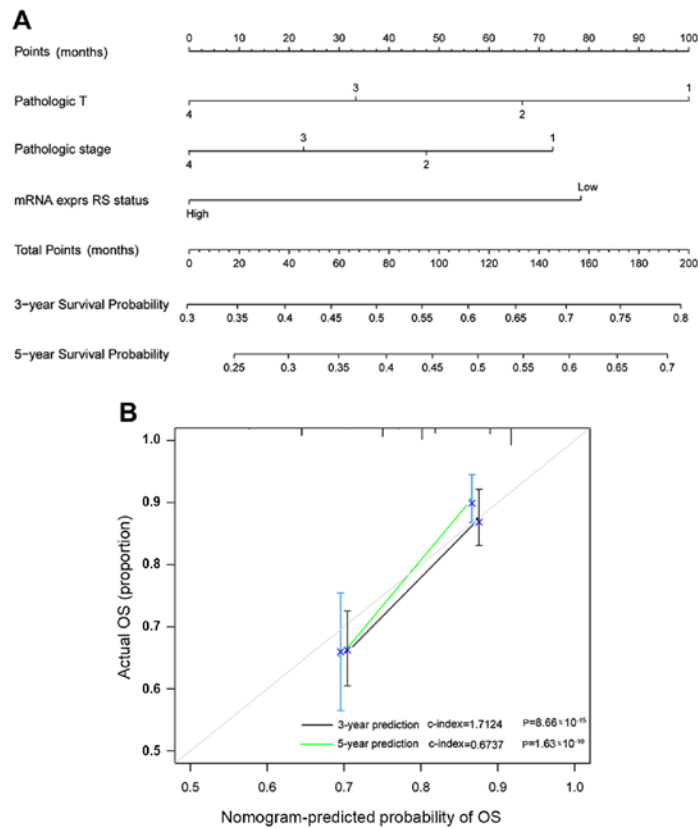


Figure 5. Nomogram with pathologic T stage, pathologic stage and three-mRNA status risk score for survival prediction in patients with hepatocellular carcinoma. (A) For each patient, a value of each variable is located on the variable axes. The number of points for each variable is determined by drawing a line upward to the Points axis. The total sum of points of all variables is located on the Total Points axis, and the 3-year and 5-year overall survival probabilities on the survival axes can be obtained. (B) Calibration plot of the nomogram for predicting the 3- and 5-year OS probabilities. The black line represents the predicted 3-year OS probability, and the green line represents the predicted 5-year OS probability. OS, overall survival; expr's, expression; RS, risk score.

in HCC (32). The SFN gene encodes stratifin, also known as 14-3-3 σ protein; several 14-3-3 σ isoforms are implicated in the progression of HCC through the regulation of cell proliferation, migration and epithelial-mesenchymal transition (33). Zhang *et al* (34) have reported that 14-3-3 σ is upregulated in HCC and has prognostic value for HCC, which is in concordance with the results of the present study. Additionally, 14-3-3 σ induced HCC cell migration and tumor progression by regulating the β -catenin/heat shock factor 1 α /70 kDa heat shock protein pathway and may serve as a promising prognostic indicator (35). The three mRNAs may represent novel potential therapeutic targets for the treatment of HCC.

To determine the molecular mechanisms of action of the three prognostic mRNAs in the pathogenesis of HCC, GO function and KEGG pathway enrichment analyses were performed using the 1,391 DEGs identified between the two risk groups predicted by the three-mRNA status risk score in TCGA dataset. The results demonstrated that the three mRNAs were involved in various cell cycle-related biological processes, cytokine-cytokine receptor interaction pathways and two progesterone-associated pathways. Similarly, based on HCC microarray data, a recent study has demonstrated that cytokine-cytokine receptor interactions may involve a large number of differentially co-expressed genes (36). The incidence of HCC is higher in men compared with women (37), and sex hormones including progesterone may influence this statistic. Progesterone may also act as the precursor for androgens and

estrogens, while the higher incidence in men compared with women may be attributable to the stimulatory effects of androgens and the inhibitory effects of estrogen (38). In addition, a number of studies have demonstrated that the hepatic androgen receptor promotes HCC development, whereas the estrogen receptor suppresses HCC progression (39,40). In hepatoma cells treated with epirubicin, progesterone increases apoptosis and inhibits autophagy by enhancing oxidative stress and upregulating the cell surface death receptor Fas/CD95 (41,42). Based on these results, it may be concluded that the three prognostic mRNAs in HCC identified in the present study may act through the modulation of these DEGs, which affect cell cycle progression, cytokine-cytokine receptor interaction and progesterone-associated pathways.

By conducting comprehensive bioinformatics analyses on existing microarray data from TCGA and EBI array repositories, the present study identified and validated a three-mRNA status risk score as the optimal score for predicting the prognosis of patients with HCC. The three-mRNA status risk score exhibited higher prognostic value compared with the lncRNA-based risk scores, and the prognosis of patients with HCC may be predicted with the median risk score calculated using the mRNA status-based risk score system as the cut-off value. The three mRNAs identified in the present study may be involved in the pathogenesis of HCC through effects on cell cycle progression and progesterone-associated pathways. Further study in larger patient cohorts is required to validate these findings.

Table IV. Summary of significant GO biological processes and KEGG pathways for upregulated genes.

Category	Term	n-value	P-value	FDR	Genes
GOTERM_BP_DIRECT	GO:0006260~DNA replication	43	P=1.95x10 ⁻¹⁴	7.41x10 ⁻¹¹	CLSPN, BLM, DBF4, TIPIN, POLA1, CHEK1, POLA2, MCM10, CDT1, CDC45, POLE2, ORC6, EGF, ORC1, RTEL1, EXO1, RECQL4, CDC7, CDK1, CDC6, GINS2, CCDC88A, DTL, GINS3, GINS4, BRIP1, TREX1, MCM2, WRN, CDC25C, MCM4, BRCA1, CDC25A, RBBP8, MCM6, POLD3, DNA2, RFC4, POLD1, RRM2, CHAF1B, BARD1, DSCC1
GOTERM_BP_DIRECT	GO:0051301~cell division	67	P=2.18x10 ⁻¹³	4.14x10 ⁻¹⁰	CDK19, KIFC1, CDC14A, KNTC1, PTTG1, SPICE1, CCNE2, KIF2C, CCNE1, NDE1, CDCA8, OIP5, CCNA2, CDCA4, CDCA3, KIF2A, KIF14, ANAPC1, CDC7, CDK1, CDC6, ARHGEF2, KIF11, CCNF, TPX2, UBE2C, TACC3, NCAPD3, RBBP8, NCAPD2, CDK3, MAD2L1, SPAG5, ZWINT, BUB1B, FBXL7, UBE2S, HAUS6, NEK2, TIPIN, SPC25, NCAPH, NCAPG2, NCAPG, BUB1, FBXO5, SKA1, ZWILCH, HELLS, ERCC6L, PARD6B, PSRC1, CENPF, KIF18B, NDC80, CENPE, CDC20, BIRC5, CDC25C, SMC2, CENPJ, CDC25A, SMC4, CCNB1, CCNB2, KIF20B, CDK20
GOTERM_BP_DIRECT	GO:0006270~DNA replication initiation	18	P=5.76x10 ⁻¹²	7.28x10 ⁻⁹	CDC7, CDC6, GINS4, POLA1, POLA2, MCM2, MCM10, MCM4, RPA4, MCM6, PRIM1, CCNE2, CCNE1, CDC45, POLE2, PRIM2, ORC6, ORC1
GOTERM_BP_DIRECT	GO:000082~G1/S transition of mitotic cell cycle	31	P=1.08x10 ⁻¹¹	1.03x10 ⁻⁸	BCAT1, DBF4, POLA1, PKMYT1, POLA2, MCM10, RPA4, CDT1, CCNE2, PRIM1, CCNE1, TYMS, CDC45, CDKN2A, POLE2, PRIM2, FBXO5, ORC6, ORC1, CDC7, CDC6, CDK1, GPR132, MCM2, CDKN3, MCM4, CDC25A, RBBP8, CDK3, MCM6, RRM2
GOTERM_BP_DIRECT	GO:0007067~mitotic nuclear division	47	P=1.71x10 ⁻⁹	1.30x10 ⁻⁶	HAUS6, NEK2, TIPIN, KNTC1, PKMYT1, CEP55, AURKB, PTTG1, SPC25, KIF2C, NCAPG2, OIP5, INCENP, TTYH1, BUB1, FBXO5, SKA1, ZWILCH, CCNA2, ASPM, TUBB3, HELLS, CDCA3, ERCC6L, ANAPC1, CDK1, CDC6, ARHGEF2, KIF11, CCNF, KIF15, TPX2, CENPF, NDC80, CDC20, BIRC5, PBK, CDC25C, CDC25A, RBBP8, CDK3, CCNB2, PLK1, KIF20B, BUB1B, FBXL7, CIT
GOTERM_BP_DIRECT	GO:0007062~sister chromatid cohesion	28	P=1.99x10 ⁻⁹	1.26x10 ⁻⁶	KNTC1, AURKB, SPC25, KIF2C, CDCA8, NDE1, DDX11, CENPA, INCENP, BUB1, SKA1, ZWILCH, KIF2A, ERCC6L, CENPO, CENPM, CENPO, KIF18A, CENPF, NDC80, BIRC5, CDC20, CENPE, CENPI, MAD2L1, PLK1, ZWINT, BUB1B
GOTERM_BP_DIRECT	GO:0000731~DNA synthesis involved in DNA repair	16	P=4.77x10 ⁻⁹	2.59x10 ⁻⁶	EXO1, XRCC3, RAD51AP1, XRCC2, BLM, POLA1, BRCA2, BRIP1, WRN, BRCA1, RAD51, RBBP8, POLD3, DNA2, POLD1, BARD1

Table IV. Continued.

Category	Term	n-value	P-value	FDR	Genes
GOTERM_BP_DIRECT	GO:0000732~strand displacement	14	P=4.99x10 ⁻⁹	2.36x10 ⁻⁶	EXO1, XRCC3, RAD51A1, XRCC2, BLM, BRCA2, BRIP1, WRN, BRCA1, RAD51, RBBP8, DNA2, RTEL1, BARD1
GOTERM_BP_DIRECT	GO:0006271~DNA strand elongation involved in DNA replication	11	P=7.05x10 ⁻⁹	2.97x10 ⁻⁶	POLD3, GINS1, PRIM1, GINS2, RFC4, POLD1, GINS3, GINS4, PRIM2, POLA1, POLA2
GOTERM_BP_DIRECT	GO:0051726~regulation of cell cycle	30	P=8.79x10 ⁻⁹	3.34x10 ⁻⁶	CDK19, E2F2, E2F5, FOXM1, PRR11, PKMYT1, MYBL2, SRC, CCNE2, CCNE1, OVOL2, PLCB1, MX2, MAK, DTL, CCNF, L3MBTL1, RBL1, CENPF, CDC25C, MECOM, CDKL5, CDC25A, CDK3, CCNB1, TRIM36, CCNB2, CDK1, PLK1, CDK20
GOTERM_BP_DIRECT	GO:000086~G2/M transition of mitotic cell cycle	29	P=3.32x10 ⁻⁷	1.14x10 ⁻⁶	CEP72, HAUS6, NEK2, FOXM1, PKMYT1, CHEK1, HMMR, NDE1, CDKN2B, CEP250, CEP290, PLCB1, CDK1, CEP135, CEP192, TPX2, BIRC5, SFI1, CDC25C, CEP152, CENPJ, CDC25A, CCNB1, PLK4, CCNB2, PLK1, FBXL7, CIT, MELK
GOTERM_BP_DIRECT	GO:0000070~mitotic sister chromatid segregation	12	P=4.05x10 ⁻⁷	1.28x10 ⁻⁴	KIFC1, MAD2L1, CENPA, PLK1, NEK2, SPAG5, ZWINT, NUSAP1, KIF18B, NDC80, ESPL1, SMC4
GOTERM_BP_DIRECT	GO:0000722~telomere maintenance via recombination	13	P=9.64x10 ⁻⁷	2.81x10 ⁻⁴	POLD3, PRIM1, DNA2, XRCC3, RFC4, POLE2, POLD1, PRIM2, POLA1, BRCA2, POLA2, WRN, RAD51
GOTERM_BP_DIRECT	GO:0007018~microtubule-based movement	20	P=3.01x10 ⁻⁶	8.15x10 ⁻⁴	KIF14, KIF23, KIFC1, KIF4A, KIF24, KIF11, KIF5A, KIF15, KIF18A, KIF18B, CENPE, RACGAP1, KIF3C, KIF2C, DYNC2H1, KIF20B, KIF21B, KIF26B, KIF20A, KIF2A
GOTERM_BP_DIRECT	GO:0043547~positive regulation of GTPase activity	69	P=1.39x10 ⁻⁵	3.52x10 ⁻³	EZH2, BTC, ARHGAP22, TIAM2, GIT2, CSF2RB, ARHGAP11A, AGAP1, DOCK10, PLCB1, CSF2RA, ARHGEF2, ARHGEF18, GRIN2A, RAB3IL1, ARTN, SIPA1L3, DEPDC1, ECT2, CDKL5, ANKRD27, NCK2, PLCE1, SH2D3A, ARHGAP33, OPHN1, RIN1, SEMA4D, DENND3, RAPGEFL1, EPS8L1, SPTB, FGFR2, FGFR1, CSF2, RAPIGAP, CYTH4, ASAP1, KITLG, LLGL1, ADAP1, RASAL1, DOCK2, CCL20, GMP, PLEKHG6, IL2RG, EGF, RASA3, RAP1GAP2, FGD6, IQSEC2, PIK3R2, OBSCN, IL2RA, ABR, VAV3, ARHGEF38, RGS17, RACGAP1, VAV1, RGS20, RGS1, CHML, RGS5, JAK3, XCL1, SH3BP1, ARAP2
GOTERM_BP_DIRECT	GO:0000083~regulation of transcription involved in G1/S transition of mitotic cell cycle	10	P=1.51x10 ⁻⁵	3.57x10 ⁻³	CCNE1, CDK1, CDC6, TYMS, CDC45, RRM2, POLA1, FBXO5, ORC1, CDT1

Table IV. Continued.

Category	Term	n-value	P-value	FDR	Genes
GOTERM_BP_DIRECT	GO:0008283~cell proliferation	49	P=3.00x10 ⁻⁵	6.67x10 ⁻³	BCAT1, STIL, CDK5R1, IL9R, CDC14A, EZF8, NR6A1, UCHL1, POLA1, KITLG, EHF, AURKB, FER, MCM10, TRAP, SRC, TGFB2, GPC4, EDNRA, TYMS, KIF2C, BUB1, TGFA, BHLHE41, IL1A, OCA2, SYK, CDK1, IL2RA, MKI67, DLGAP5, KIF15, TPX2, SIX2, CENPF, ITGA2, TNFSF9, TACC3, CDC25C, CDC25A, CDK3, SH2D2A, PLCE1, S100B, PLK1, SIX1, BUB1B, FBXL7, MELK
GOTERM_BP_DIRECT	GO:0032508~DNA duplex unwinding	13	P=4.04x10 ⁻⁵	8.49x10 ⁻³	RECQL4, GINS1, GINS2, DNA2, CDC45, DDX11, BLM, GINS4, BRIP1, RAD54B, WRN, RTEL1, CHD3
GOTERM_BP_DIRECT	GO:0007165~signal transduction	119	P=4.66x10 ⁻⁵	9.26x10 ⁻³	IL9R, CRABP2, TLR1, TNFSF15, IL15, TLR6, ITPKA, ADORA1, CASP8AP2, UNC5B, ANK2, PDE4A, CHRNA5, CSF3R, IL1B, TYRO3, STK24, PIK3CD, LRRCL1, DEPDC1, PDE4C, PLAUR, GABRR2, LILRB1, TNFAIP6, CCR6, IGSF1, LRP12, LILRB4, PDE5A, OPHN1, MST1R, CLEC5A, PTK7, OPN1SW, SOX9, SRC, RASAL1, VDR, NPHP4, IL12RB1, EGF, TRAF5, NPHP1, GABRE, ABR, GNRH1, SPHK1, CD300C, SALL2, RGS1, HIVEP3, XCL1, GRB7, CHRNG, GPR85, CLDN4, TAS2R4, RASSF8, PDCC1, SIRPB1, EDNRA, ARHGAP22, TNFRSF11A, PTGES, CSF2RB, MICAL1, CNTNAP1, ARHGAP11A, DLG5, PLCB1, LTA, MPP2, LIMK1, MPP3, ARTN, CDS1, SH2D2A, HUNK, DOK1, NCK2, TNFRSF10C, ACVR2B, ARHGAP3, CLIC3, ZDHHC13, RIN1, PDE9A, INPP4A, GNB3, RAPIGAP, CXCL5, GULP1, CTNND2, TNFRSF8, KITLG, SFN, TRAP, IGF1R, CCL20, FCGR1A, DGKG, BCL11A, ARR3, IL2RG, SH2B2, DKKL1, RASA3, HPGDS, PIK3R2, EPO, TNFSF4, LRRN2, TNFSF9, RACGAP1, LRP8, SH3BP1, ARAP2, CHRNA10
GOTERM_BP_DIRECT	GO:0001932~regulation of protein phosphorylation	11	P=8.80x10 ⁻⁵	1.65x10 ⁻²	PHIP, CCDC88A, CCDC88C, ARR3, EZH2, CELSR3, FER, TLR7, KDM4D, TLR8, RAD51
GOTERM_BP_DIRECT	GO:0007059~chromosome segregation	15	P=2.62x10 ⁻⁴	4.63x10 ⁻²	KIF11, NEK2, CENPF, NDC80, CENPE, BRCA1, SPC25, NDE1, HJURP, SPAG5, OIP5, INCENP, SKA1, TOP2A, TOP3B
KEGG_PATHWAY	hsa04110:Cell cycle	38	P=2.06x10 ⁻¹⁴	5.08x10 ⁻¹²	E2F1, E2F2, E2F3, CDC14A, E2F5, DBF4, PKMYT1, TTK, CHEK1, SFN, PTTG1, TGFB2, CCNE2, CCNE1, CDC45, CDKN2A, CDKN2B, BUB1, ORC6, CCNA2, ORC1, CDC7, ANAPC1, CDK1, CDC6, RBL1, ESPL1, CDC20, MCM2, CDC25C, MCM4, CDC25A, MCM6, CCNB1, MAD2L1, CCNB2, PLK1, BUB1B

Table IV. Continued.

Category	Term	n-value	P-value	FDR	Genes
KEGG_PATHWAY	hsa03030:DNA replication	13	P=4.28x10 ⁻⁶	5.26x10 ⁻⁴	POLD3, PRIM1, DNA2, RFC4, POLE2, POLD1, PRIM2, POLA1, MCM2, POLA2, MCM4, RPA4, MCM6
KEGG_PATHWAY	hsa03440:Homologous recombination	11	P=1.95x10 ⁻⁵	1.60x10 ⁻³	POLD3, XRCC3, XRCC2, BLM, POLD1, BRCA2, RAD54B, RAD54L, RPA4, RAD51, TOP3B
KEGG_PATHWAY	hsa04640:Hematopoietic cell lineage	19	P=3.55x10 ⁻⁵	2.18x10 ⁻³	CSF2, IL9R, IL2RA, KITLG, ITGA2, CD1B, CD1A, ITGA3, ITGA4, ITGAM, FCGR1A, CD22, IL1B, CSF3R, CD24, IL1A, CSF2RA, EPO, CD7
KEGG_PATHWAY	hsa04060:Cytokine-cytokine receptor interaction	35	P=1.20x10 ⁻⁴	5.87x10 ⁻³	CSF2, IL9R, CXCL5, IL21R, TNFSF15, TNFRSF8, IL15, TGFB2, LIF, TNFRSF11A, IL12RB1, IL17B, CCL20, CLCF1, CCR10, CSF3R, CSF2RB, IL1B, IL2RG, LTA, CSF2RA, IL1A, EPO, TNFSF4, IL2RA, TNFSF9, CCR8, AMH, TNFRSF10C, ACVR2B, CCR6, IL20RA, IL12A, EDA, XCL1
KEGG_PATHWAY	hsa05200:Pathways in cancer	48	P=3.17x10 ⁻⁴	1.29x10 ⁻²	E2F1, FGFR2, E2F2, FGFR1, E2F3, ADCY7, MMP9, ARNT2, MITF, EGLN3, KITLG, LPAR2, BDKRB1, MMP1, TGFB2, GLI1, EDNRA, CCNE2, CCNE1, IGFIR, CDKN2A, CDKN2B, PAX8, SLC2A1, TGFA, CSF3R, EGF, PLCB1, TRAF5, CSF2RA, PIK3R2, CTBP2, PTGER4, MSH2, PIK3CD, CBL, FZD1, ITGA2, BRCA2, ITGA3, BIRC5, FZD2, MECOM, FZD7, RAD51, COL4A5, LAMC2, GNB3, ANAPC1, CDK1, ADCY7, PIK3CD, PKMYT1, CDC25C, CDC25A, CCNB1, IGFIR, CCNB2, MAD2L1, PLK1, MAPK13, BUB1, CCNA2, PIK3R2
KEGG_PATHWAY	hsa04914:Progesterone-mediated oocyte maturation	16	P=1.21x10 ⁻³	4.19x10 ⁻²	

GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; FDR, false discovery rate.

Table V. Summary of significant GO biological processes and KEGG pathways for down-regulated genes.

Category	Term	n-value	P-value	FDR	Genes
GOTERM_BP_DIRECT	GO:0055114~oxidation-reduction process	17	P=5.41x10 ⁻⁸	3.79x10 ⁻⁵	CYP3A4, STEAP4, CYP11A1, CYP1A1, AASS, CYP26A1, CYP1A2, CYP2E1, ACADL, RDH5, CYP3A43, CYP39A1, CYP4A22, CYP2A6, CYP2A7, SRD5A2, HPD
GOTERM_BP_DIRECT	GO:0006805~xenobiotic metabolic process	8	P=1.54x10 ⁻⁷	5.38x10 ⁻⁵	CYP3A4, GLYAT, NRII2, SULT1B1, NAT2, CYP26A1, CYP2E1, CYP1A2
GOTERM_BP_DIRECT	GO:0019373~epoxygenase P450 pathway	5	P=1.95x10 ⁻⁶	4.54x10 ⁻⁴	CYP1A1, CYP2A6, CYP2A7, CYP2E1, CYP1A2
GOTERM_BP_DIRECT	GO:0008202~steroid metabolic process	6	P=2.74x10 ⁻⁶	4.79x10 ⁻⁴	CYP3A4, NRII2, CYP1A1, SULT1B1, CYP2A6, CYP2E1
GOTERM_BP_DIRECT	GO:0017144~drug metabolic process	5	P=1.08x10 ⁻⁵	1.51x10 ⁻³	CYP3A4, CYP1A1, CYP2A6, CYP2E1, CYP1A2
GOTERM_BP_DIRECT	GO:0042738~exogenous drug catabolic process	4	P=2.86x10 ⁻⁵	3.33x10 ⁻³	CYP3A4, NRII2, CYP2A6, CYP1A2
GOTERM_BP_DIRECT	GO:0006706~steroid catabolic process	3	P=2.63x10 ⁻⁴	2.59x10 ⁻²	CYP3A4, SRD5A2, CYP1A2
GOTERM_BP_DIRECT	GO:0015721~bile acid and bile salt transport	4	P=3.59x10 ⁻⁴	3.09x10 ⁻²	SLCO1B3, ABCB11, SLC27A5, SLC10A1
GOTERM_BP_DIRECT	GO:0016098~monoterpenoid metabolic process	3	P=3.93x10 ⁻⁴	3.01x10 ⁻²	CYP3A4, CYP2E1, CYP1A2
GOTERM_BP_DIRECT	GO:0046483~heterocycle metabolic process	3	P=3.93x10 ⁻⁴	3.01x10 ⁻²	CYP3A4, CYP2E1, CYP1A2
GOTERM_BP_DIRECT	GO:0006810~transport	9	P=4.42x10 ⁻⁴	3.05x10 ⁻²	ABCA8, AR, AQP9, ABCB11, SLC6A12, FABP4, SEC14L2, ABCB4, SLC10A1
GOTERM_BP_DIRECT	GO:0042908~xenobiotic transport	3	P=5.48x10 ⁻⁴	3.43x10 ⁻²	ABCA8, NRII2, ABCB4
KEGG_PATHWAY	hsa04976:Bile secretion	7	P=1.61x10 ⁻⁵	1.61x10 ⁻³	SLCO1B3, AQP9, ABCB11, SLC27A5, SLC22A1, ABCB4, SLC10A1
KEGG_PATHWAY	hsa05204:Chemical carcinogenesis	7	P=3.78x10 ⁻⁵	1.89x10 ⁻³	CYP3A43, CYP3A4, CYP1A1, NAT2, CYP2A6, CYP2E1, CYP1A2
KEGG_PATHWAY	hsa00140:Steroid hormone biosynthesis	6	P=9.01x10 ⁻⁵	3.00x10 ⁻³	CYP3A4, CYP1A1, CYP11A1, CYP2E1, SRD5A2, CYP1A2
KEGG_PATHWAY	hsa00830:Retinol metabolism	6	P=1.45x10 ⁻⁴	3.61x10 ⁻³	CYP3A4, CYP1A1, CYP26A1, CYP2A6, CYP1A2, RDH5
KEGG_PATHWAY	hsa00232:Caffeine metabolism	3	P=6.18x10 ⁻⁴	1.23x10 ⁻²	NAT2, CYP2A6, CYP1A2
KEGG_PATHWAY	hsa03320:PPAR signaling pathway	5	P=1.93x10 ⁻³	3.17x10 ⁻²	FABP4, ACADL, SLC27A5, ACSL6, PKC1
KEGG_PATHWAY	hsa00980:Metabolism of xenobiotics by cytochrome P450	5	P=2.78x10 ⁻³	3.90x10 ⁻²	CYP3A4, CYP1A1, CYP2A6, CYP2E1, CYP1A2

GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; FDR, false discovery rate.

Acknowledgements

Not applicable.

Funding

No funding was received.

Availability of data and materials

All the data used in the study were derived from public databases including TCGA (<https://gdc-portal.nci.nih.gov/>), EBI (<https://www.ebi.ac.uk/arrayexpress/>) and HGNC (<http://www.genenames.org/>).

Authors' contributions

WZ, QF and KY all contributed to data analysis and manuscript writing and revision. All authors read and approved the final version of the manuscript.

Ethics approval and consent to participate

Not applicable.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

- Siegel RL, Miller KD and Jemal A: Cancer statistics, 2017. *CA Cancer J Clin* 67: 7-30, 2017.
- Mittal S and El-Serag HB: Epidemiology of hepatocellular carcinoma: Consider the population. *J Clin Gastroenterol* 47 (Suppl): S2-S6, 2013.
- Wallace MC, David P, Jeffrey GP and Adams LA: The evolving epidemiology of hepatocellular carcinoma: A global perspective. *Expert Rev Gastroenterol Hepatol* 9: 765-779, 2015.
- Yang Y, Chen L, Gu J, Zhang H, Yuan J, Lian Q, Lv G, Wang S, Wu Y, Yang YT, *et al*: Recurrently deregulated lncRNAs in hepatocellular carcinoma. *Nat Commun* 8: 14421, 2017.
- Huo X, Han S, Wu G, Latchoumanin O, Zhou G, Hebbard L, George J and Qiao L: Dysregulated long noncoding RNAs (lncRNAs) in hepatocellular carcinoma: Implications for tumorigenesis, disease progression, and liver cancer stem cells. *Mol Cancer* 16: 165, 2017.
- Gu JX, Zhang X, Miao RC, Xiang XH, Fu YN, Zhang JY, Liu C and Qu K: Six-long non-coding RNA signature predicts recurrence-free survival in hepatocellular carcinoma. *World J Gastroenterol* 25: 220-232, 2019.
- Wang Z, Wu Q, Feng S, Zhao Y and Tao C: Identification of four prognostic lncRNAs for survival prediction of patients with hepatocellular carcinoma. *PeerJ* 5: e374, 2017.
- Yan J, Zhou C, Guo K, Li Q and Wang Z: A novel seven-lncRNA signature for prognosis prediction in hepatocellular carcinoma. *J Cell Biochem* 120: 213-223, 2019.
- Li B, Feng W, Luo O, Xu T, Cao Y, Wu H, Yu D and Ding Y: Development and validation of a three-gene prognostic signature for patients with hepatocellular carcinoma. *Sci Rep* 7: 5517, 2017.
- Shi YM, Li YY, Lin JY, Zheng L, Zhu YM and Huang J: The discovery of a novel eight-mRNA-lncRNA signature predicting survival of hepatocellular carcinoma patients. *J Cell Biochem* 2018 (Online ahead of print).
- Tomczak K, Czerwińska P and Wiznerowicz M: The cancer genome Atlas (TCGA): An immeasurable source of knowledge. *Contemp Oncol (Pozn)* 19: A68-A77, 2015.
- Kim SM, Leem SH, Chu IS, Park YY, Kim SC, Kim SB, Park ES, Lim JY, Heo J, Kim YJ, *et al*: Sixty-five gene-based risk score classifier predicts overall survival in hepatocellular carcinoma. *Hepatology* 55: 1443-1452, 2012.
- Chaudhary K, Poirion OB, Lu L and Garmire LX: Deep learning based multi-omics integration robustly predicts survival in liver cancer. *Clin Cancer Res* 24: 1248-1259, 2018.
- Li S, Chen X, Liu X, Yu Y, Pan H, Haak R, Schmidt J, Ziebolz D and Schmalz G: Complex integrated analysis of lncRNAs-miRNAs-mRNAs in oral squamous cell carcinoma. *Oral Oncol* 73: 1-9, 2017.
- Liu JY, Yao J, Li XM, Song YC, Wang XQ, Li YJ, Yan B and Jiang Q: Pathogenic role of lncRNA-MALAT1 in endothelial cell dysfunction in diabetes mellitus. *Cell Death Dis* 5: e1506, 2014.
- Wang P, Wang Y, Hang B, Zou X and Mao JH: A novel gene expression-based prognostic scoring system to predict survival in gastric cancer. *Oncotarget* 7: 55343-55351, 2016.
- Goeman JJ: L1 penalized estimation in the Cox proportional hazards model. *Biom J* 52: 70-84, 2010.
- Anevski D: Functional central limit theorems for the Nelson-Aalen and Kaplan-Meier estimators for dependent stationary data. *Statistics Probability Lett* 124: 83-91, 2017.
- Zhi S, Liu Y, Li Z and Zhang J: A weighted likelihood ratio test-based chart for monitoring process mean and variability. *J Statistical Computation Simulation* 88: 1-22, 2018.
- Obuchowski NA and Bullen JA: Receiver operating characteristic (ROC) curves: Review of methods with applications in diagnostic medicine. *Phy Med Biol* 63 07TR01: 2018.
- Khaksar E, Askarishahi M, Hekmatimoghaddam S and Vahedian-Ardakani H: Cox regression and parametric models comparison of how they determine factors influencing survival of patients with non-small cell lung carcinoma. *Asian Pac J Cancer Prev* 18: 3389-3393, 2017.
- Eng KH, Emily S and Kayla M: On representing the prognostic value of continuous gene expression biomarkers with the restricted mean survival curve. *Oncotarget* 6: 36308-36318, 2015.
- Wu J, Zhou L, Huang L, Gu J, Li S, Liu B, Feng J and Zhou Y: Nomogram integrating gene expression signatures with clinicopathological features to predict survival in operable NSCLC: A pooled analysis of 2164 patients. *J Exp Clin Cancer Res* 36: 4, 2017.
- Huang G, Zhao G, Xia J, Wei Y, Chen F, Chen J and Shi J: FGF2 and FAM201A affect the development of osteonecrosis of the femoral head after femoral neck fracture. *Gene* 652: 39-47, 2018.
- Balogh J, Victor D III, Asham EH, Burroughs SG, Boktour M, Saharia A, Li X, Ghobrial RM and Monsour HP Jr: Hepatocellular carcinoma: A review. *J Hepatocell Carcinoma* 3: 41-53, 2016.
- Telford EJ, Jiang WG and Martin TA: HAVCr-1 involvement in cancer progression. *Histol Histopathol* 32: 121-128, 2017.
- Liu L, Song Z, Zhao Y, Li C, Wei H, Ma J and Du Y: HAVCR1 expression might be a novel prognostic factor for gastric cancer. *PLoS One* 13: e0206423, 2018.
- Lee JK, Phillips JW, Smith BA, Park JW, Stoyanova T, McCaffrey EF, Baertsch R, Sokolov A, Meyerowitz JG, Mathis C, *et al*: N-myc drives neuroendocrine prostate cancer initiated from human prostate epithelial cells. *Cancer Cell* 29: 536-547, 2016.
- Rickman DS, Schulte JH and Eilers M: The expanding world of N-MYC-driven tumors. *Cancer Discov* 8: 150-163, 2018.
- Lin CP, Liu CR, Lee CN, Chan TS and Liu HE: Targeting c-Myc as a novel approach for hepatocellular carcinoma. *World J Hepatol* 2: 16-20, 2010.
- Schlaeger C, Longrich T, Schiller C, Bewerunge P, Mehrabi A, Toedt G, Kleeff J, Ehemann V, Eils R, Lichter P and Bernhard PS: Etiology-dependent molecular mechanisms in human hepatocarcinogenesis. *Hepatology* 47: 511-520, 2008.
- Qin XY, Suzuki H, Honda M, Okada H, Kaneko S, Inoue I, Ebisui E, Hashimoto K, Carninci P, Kanki K, *et al*: Prevention of hepatocellular carcinoma by targeting MYCN-positive liver cancer stem cells with acyclic retinoid. *Proc Natl Acad Sci USA* 115: 4969-4974, 2018.
- Wu YJ, Jan YJ, Ko BS, Liang SM and Liou JY: Involvement of 14-3-3 proteins in regulating tumor progression of hepatocellular carcinoma. *Cancers (Basel)* 7: 1022-1036, 2015.
- Zhang Y, Li Y, Lin C, Ding J, Liao G and Tang B: Aberrant upregulation of 14-3-3σ and EZH2 expression serves as an inferior prognostic biomarker for hepatocellular carcinoma. *PLoS One* 9: e107251, 2014.

35. Liu CC, Jan YJ, Ko BS, Wu YM, Liang SM, Chen SC, Lee YM, Liu TA, Chang TC, Wang J, *et al*: 14-3-3 σ induces heat shock protein 70 expression in hepatocellular carcinoma. *BMC Cancer* 14: 425, 2014.
36. Wang Y, Jiang T, Li Z, Lu L, Zhang R, Zhang D, Wang X and Tan J: Analysis of differentially co-expressed genes based on microarray data of hepatocellular carcinoma. *Neoplasma* 64: 216-221, 2017.
37. Hainaut P, Amadou A and Gormally E: Cancer prevention and control: Hepatocellular carcinoma. *Ecancermedicalscience* 13: 949, 2019.
38. Yeh YT, Chang CW, Wei RJ and Wang SN: Progesterone and related compounds in hepatocellular carcinoma: Basic and clinical aspects. *Biomed Res Int* 2013: 290575, 2013.
39. Ma WL, Hsu CL, Yeh CC, Wu MH, Huang CK, Jeng LB, Hung YC, Lin TY, Yeh S and Chang C: Hepatic androgen receptor suppresses hepatocellular carcinoma metastasis through modulation of cell migration and anoikis. *Hepatology* 56: 176-185, 2012.
40. Keng VW, Largaespada DA and Augusto V: Why men are at higher risk for hepatocellular carcinoma? *J Hepatol* 57: 453-454, 2012.
41. Chang WT, Hsieh BS, Cheng HL, Lee KT and Chang KL: Progesterone augments epirubicin-induced apoptosis in HA22T/VGH cells by increasing oxidative stress and upregulating Fas/FasL. *J Surg Res* 188: 432-441, 2014.
42. Chang WT, Cheng HL, Hsieh BS, Chiu CC, Lee KT and Chang KL: Progesterone increases apoptosis and inversely decreases autophagy in human hepatoma HA22T/VGH cells treated with epirubicin. *ScientificWorldJournal* 2014: 567148, 2014.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.