

RESEARCH ARTICLE

Editorial Process: Submission:07/21/2021 Acceptance:12/15/2021

lncRNA Expression-Based Risk Scoring System Can Predict Survival of Tumor-Positive Patients with Hepatocellular Carcinoma

Siyao Wu, Yayan Deng, Yue Luo, Jiaxiang Ye*, Zhihui Liu*

Abstract

Background: Long non-coding RNAs (lncRNAs) play critical roles in the progression of hepatocellular carcinoma (HCC). The aim of this study was to explore whether lncRNA expression profiles can predict prognosis of HCC patients with tumors. **Methods:** Expression of lncRNAs in HCC patients based on data in The Cancer Genome Atlas (TCGA) was examined by uni- and multivariate cox analysis to identify associations between clinical features and overall survival (OS) or recurrence-free survival (RFS). Based on our finding that both were independently associated with tumor status, we examined lncRNAs differentially expressed between patients with or without tumors. An lncRNA-based risk scoring system was developed to predict OS and RFS in tumor-positive patients, and it was assessed using uni- and multivariate cox analyses. Potential functions of the prognostic lncRNAs were explored. **Results:** A risk scoring system to predict OS for HCC patients with tumors was developed based on the expression of six lncRNAs (AC090921.1, AC012640.1, AL158839.1, AL356056.1, AL359853.1 and C10orf91), and a corresponding scoring system to predict RFS was developed from nine lncRNAs (AL356056.1, AL158839.1, MIR7-3HG, AL445493.2, AP000808.1, AP003354.2, PLCE1-AS1, TH2LCRR and LINC01447). Both risk scoring systems gave areas under receiver operating characteristic curves >0.7. Uni- and multivariate cox analyses showed that both risk scoring systems independently predicted survival even after adjusting for clinical factors. The lncRNAs related to OS may be involved in complement and coagulation cascades, while those related to RFS may be involved in the cell cycle. **Conclusion:** Risk scoring system based on these lncRNAs may be useful for predicting prognosis of tumor-positive HCC patients.

Keywords: Hepatocellular carcinoma- risk scoring system- long non-coding RNA- survival

Asian Pac J Cancer Prev, 22 (12), 3741- 3753

Introduction

Hepatocellular carcinoma (HCC) is estimated to cause 782,000 death cases annually and its incidences and associated mortality are increasing, making it one of the most fatal malignant cancers in the world (Bertuccio et al., 2017; Forner et al., 2018; Siegel et al., 2020). Median survival time of patients with advanced HCC is only 7.1 months without treatment (Kulik and El-Serag, 2019). Even after treatment, tumor recurrence is a problem. For example, recurrence occurs in nearly 70% of patients by 5 years after surgery, reducing their survival (Forner et al., 2018; Xu et al., 2020). Recurrence rates after R0 resection are high even for patients with early-stage HCC, leading to unsatisfactory long-term survival (Torzilli et al., 2013; Zhong et al., 2014; Kang and Ahn, 2017). OS after ablation or radiotherapy is as poor as after surgery (Bruix et al., 2014; Hara et al., 2019). Therefore, recurrence maybe a factor that affect the prognosis. Indeed, the typical

heterogeneity of HCC is a major reason why rates of overall survival (OS) and recurrence-free survival (RFS) remain very low for HCC patients (Villanueva et al., 2013; Bruix et al., 2014). This highlights the need to identify reliable biomarkers to predict prognosis.

Long non-coding RNAs (lncRNAs), usually longer than 200 nt (Ponting et al., 2009), are associated with occurrence and progression of malignant tumors (Schmitt and Chang, 2016; Fernández-Barrena et al., 2017; Huang et al., 2018; Wu et al., 2018b). For example, the lncRNA MAGI2-AS3 may protect against HCC (Pu et al., 2019), the lncRNA HIS is thought to promote HCC proliferation and metastasis (Chen et al., 2019a), and the lncRNA DUXAP10 acts via microRNA-1914 to inhibit cell proliferation (Sun et al., 2019).

Considering that the prognosis of HCC patients can differ depending on whether they have tumors or not (Li et al., 2019; Lim et al., 2019), we wanted to explore whether some lncRNAs are differentially expressed in the presence

Department of Medical Oncology, Guangxi Medical University Cancer Hospital, 71 Hedi Road, Nanning 530021, People's Republic of China. *For Correspondence: yejiaxiang2013@163.com, lzhh101@sina.com. Siyao Wu and Yayan Deng have equal contribution in this study.

of tumors, and whether these might influence the risk of recurrence or death. We aimed to identify a risk scoring system based on lncRNA expression levels in order to predict prognosis of tumor-positive HCC patients.

Materials and Methods

Data on patients with HCC and their lncRNA expression profiles

Data sets containing lncRNA and mRNA expression as well as the corresponding clinical information for patients with HCC were obtained from the Cancer Genome Atlas (TCGA data version 09-14-2017 for HCC) via UCSC Xena (<https://xenabrowser.net/datapages/>). To be included in the present study, patients had to have been diagnosed with HCC based on histology, the patient's tumor status had to be known, complete RNA-Seq data had to be available for lncRNAs and mRNAs, and complete follow-up data including OS and RFS had to be available. In the end, data were included from 338 HCC patients, of whom 146 had tumors and 192 did not. Since the data is from the TCGA, no further approval was required from the Ethics Committee.

Association of clinical features with survival, and analysis of lncRNA expression

Univariate cox analysis was carried out to identify associations between clinical features and OS or RFS. Factors associated with P values less than 0.05 were regarded as statistically significant and analyzed further in multivariate cox regression.

Therefore we compared lncRNA expression between HCC patients with or without tumors. First, we removed lncRNAs and mRNAs that showed null expression in more than 50% of patients. Then the edgeR package in R software (version 3.4.4, <https://www.r-project.org/>) was used to identify differentially expressed lncRNAs (DELncRNAs) that showed a $|\log_2(\text{fold change})| > 1$ and false discovery rate (FDR) < 0.05 (Robinson et al., 2010). Cluster heat maps and volcano maps were generated using the gplots and heatmap packages in R.

Construction of an lncRNA expression-based risk scoring system and prognostic assessment

All analyses were conducted using R/Bioconductor (version 3.4.4, <https://www.r-project.org/>). First, the standardized expression levels from multiple tissues in the same patient were averaged, then univariate cox analysis was used to assess associations between DELncRNAs and OS or RFS. The DELncRNAs that proved significant ($P < 0.05$) in this analysis were included in backward stepwise multivariate cox regression. The prognostic lncRNAs from this model were assembled into a scoring system based on expression level and corresponding regression coefficients (β):

Risk score = (β^* expression level of lncRNA1) + (β^* expression level of lncRNA2) + (β^* expression level of lncRNA3) +

A risk score was calculated for each patient, and time-dependent receiver operating characteristic (ROC) curves were plotted to assess prognostic performance.

The median risk score was used to classify HCC patients into high- or low-risk groups, and Kaplan–Meier analysis was performed to compare OS and RFS between the two groups.

Validation of the risk scoring system

Uni- and multivariate analyses were carried out with tumor-positive patients' data to estimate the association between various clinical variables containing the risk score and survival. If the result of these analyses did not return any significant results, the stratified analysis based on clinical variables was conducted to identify prognostic factors via the chi-squared test. All these analyses were carried out using SPSS 16.0 (IBM, Chicago, IL, USA). Differences associated with a two-sided $P < 0.05$ were considered statistically significant.

Identification of mRNAs co-expressed with prognostic lncRNAs and prediction of their function

The data sets were screened for mRNAs whose expression was strongly associated with that of the prognostic lncRNAs in our risk scoring system. Such mRNAs had to give a z-test $P < 0.01$ and two-sided | Pearson correlation coefficient | > 0.30 . Potential functions of these mRNAs were explored by performing enrichment analyses for cellular components, biological process, and molecular functions in the Gene Ontology (GO) system, as well as for pathways in the Kyoto Encyclopedia of Genes and Genomes (KEGG) system. Enrichment analyses were performed using the clusterProfiler package in R (Yu et al., 2012). $P < 0.05$ was considered to indicate statistical significance.

Results

delncRNAs in HCC

A total of 338 HCC patients were included in our analysis. Clinical and demographic characteristics of the patients are shown in Table 1. Uni- and multivariate cox analysis identified tumor status as an independent predictor of OS and RFS (Tables 2-3). DELncRNAs whose expression differed between 146 tumor-positive and 192 tumor-free patients were selected for further analysis (Figures 1-2).

Risk scoring system to predict OS of HCC patients with tumors

Univariate cox analysis based on HCC-specific lncRNAs identified nine that were significantly associated with OS: AL356056.1, AC090921.1, AL359853.1, AC139749.1, C10orf91, AC012640.1, AL158839.1, AC016205.1, AC007938.3. Of these lncRNAs, multivariate analysis identified six as independently associated with OS: AC090921.1, AC012640.1, AL158839.1, AL356056.1, AL359853.1 and C10orf91. The contribution of each lncRNA to this risk scoring system was assessed using multivariate cox regression (Table 4), such that each lncRNA was weighted appropriately to give the final formula:

Risk score = (0.2047 * AC090921.1) + (0.1440 * AC012640.1) + (0.2350 * AL158839.1) + (0.3872 *

AL356056.1) + (0.2318 * AL359853.1) + (0.1328 * C10orf91).

In this formula, higher expression of DElncRNAs is associated with worse OS ($\beta > 0$).

An OS risk score was calculated for each patient, and the median risk score was used to stratify tumor-positive patients into low- or high-risk groups (Figure 3A). Kaplan–Meier analysis showed that cumulative rates of OS were 69.3% at 3 years and 38.6% at 5 years for low-risk patients, compared to 31.6% and 14.9% for high-risk patients (Figure 4A). The area under the receiver operating characteristic curve for predicting OS was 0.76 (Figure 5A).

Next we assessed whether the risk score could predict OS independently of the following clinical features that may also influence survival: age, body mass index (BMI), ethnicity, alpha fetoprotein (AFP), sex, presence of hepatitis, alcohol consumption, cirrhosis, histology grade, new tumor event, pathology stage, tumor status, family history of cancer, presence of residual tumor, and vascular invasion. Univariate analysis showed that risk score and vascular invasion were significantly associated with OS. Of these two factors, only risk score emerged from multivariate analysis as an independent predictor of OS [hazard ratio (HR) 2.32, 95% confidence interval (CI) 1.43-3.77] (Table 5).

Risk scoring system to predict RFS of HCC patients with tumor

Using the same procedure as in section 3.2, we identified nine lncRNAs significantly associated with RFS: AL356056.1, AL158839.1, MIR7-3HG, AL445493.2, AP000808.1, AP003354.2, PLCE1-AS1, TH2LCRR and LINC01447. Weighting the contributions of individual lncRNAs (Table 6) led to the final risk scoring system:

Risk score = (0.2502 * AL356056.1) + (0.2805 * AL158839.1) + (0.2882 * MIR7-3HG) + (-0.2911 * AL445493.2) + (-0.0984 * AP000808.1) + (-0.0.2184 * AP003354.2) + (-0.4527 * PLCE1-AS1) + (0.1879 * TH2LCRR) + (-0.0751 * LINC01447).

In this formula, higher expression of AL158839.1, TH2LCRR, AL356056.1 and MIR7-3HG is related to worse RFS ($\beta > 0$), while higher expression of PLCE1-AS1, AL445493.2, AP003354.2, LINC01447 and AP000808.1 is associated with better RFS ($\beta < 0$).

Patients were stratified into low- or high-risk groups based on median risk score (Figure 3B), and their RFS was compared using Kaplan-Meier analysis (Figure 4B). RFS was 21.0% at 3 years and 6.3% at 5 years for low-risk patients, compared to 0% at both time points for high-risk patients (Figure 5B). The area under the receiver operating characteristic curve for predicting RFS was 0.861.

Univariate analysis showed that risk score, AFP, pathology stage and vascular invasion correlated with RFS, while multivariate analysis identified only risk score (HR 3.86, 95%CI 2.48-6.01) and AFP (HR 1.75, 95%CI 1.09-2.82) as independent factors (Table 7).

Table 1. Clinicopathological Characteristics of 338 HCC Patients with or without Tumor

Clinicopathological characteristics	Patients (n=338)	
	n	%
Age	≤60	164 48.52
	≥60	174 51.48
BMI	≤25	168 49.7
	≥25	143 42.31
	Not reported	27 7.99
Race	Non-Asian	179 52.96
	Asian	149 44.08
	Not reported	10 2.96
AFP	≤20	138 40.8
	≥20	119 35.2
	Not reported	81 24
Gender	Female	108 31.95
	Male	230 68.05
Hepatitis	No	178 52.66
	Yes	145 42.9
	Not reported	15 4.44
Alcohol consumption	No	214 63.31
	Yes	109 32.25
	Not reported	15 4.44
Cirrhosis	non-cirrhosis	125 36.98
	Cirrhosis	75 22.19
	Not reported	138 40.83
Histologic grade	G1-2	211 62.43
	G3-4	122 36.09
	Not reported	5 1.48
New tumor event	No	170 50.3
	Yes	159 47.04
	Not reported	9 2.66
Pathologic stage	Stage I+II	234 69.23
	Stage II+III	81 23.96
	Not reported	23 6.8
Tumor status	Tumor free	192 56.8
	With tumor	146 43.2
Family cancer history	No	193 57.1
	Yes	102 30.18
	Not reported	43 12.72
Residual tumor	R0	301 89.05
	non-R0	30 8.88
	Not reported	7 2.07
Vascular invasion	Negative	188 55.62
	Positive	97 28.7
	Not reported	53 15.68

BMI, Body mass Index; AFP, Alpha fetoprotein; *Hepatitis B or C

Functional analysis of mRNAs strongly related to prognostic lncRNAs

The mRNAs co-expressed with DElncRNAs were identified based on Pearson's correlation coefficients

Table 2. Univariate and Multivariate Cox Regression Analysis for OS in 338 HCC Patients

Variables	Univariate Cox regression				Multivariate Cox regression			
	P-value	HR	95% CI		P-value	HR	95% CI	
Age (>60/≤60)	0.46	1.17	0.77	1.79				
BMI	0.19							
<25		Reference						
≥25		0.68	0.43	1.09				
Not reported		1.04	0.44	2.42				
Race	0.36							
Non-Asian		Reference						
Asian		0.71	0.38	1.31				
Not reported		1.58	0.52	4.8				
AFP	0.17							
≤20ng/mL		Reference						
>20ng/ml		1.46	0.88	2.43				
Not reported		1.68	0.95	2.98				
Gender (Male/Female)	0.34	1.26	0.79	2.01				
Hepatitis B or C	0.44							
No		Reference						
Yes		0.83	0.48	1.45				
Not reported		1.58	0.55	4.53				
Alcohol	0.29							
No		Reference						
Yes		0.76	0.46	1.27				
Cirrhosis	0.14							
No		Reference						
Yes		1.21	0.63	2.33				
Not reported		1.72	1.01	2.94				
Histologic grade	0.3							
G1-2		Reference						
G3-4		1.39	0.91	2.12				
Not reported		0.82	0.22	3.04				
New tumor event	0							
No		Reference				Reference		
Yes		0.88	0.47	1.64	0.35	0.76	0.42	1.37
Not reported		5.46	1.96	15.17	0	4.14	1.74	9.88
Pathologic stage	0.06							
Stage I+II		Reference						
Stage III+IV		1.73	1.09	2.74				
Not reported		1.6	0.76	3.37				
Tumor Status	0.01	2.2	1.19	4.07	0	3.24	1.81	5.8
Family cancer history	0.32							
No		Reference						
Yes		0.9	0.56	1.45				
Not reported		0.54	0.24	1.21				
Residual tumor	0.78							
R0		Reference						
Non-R0		0.92	0.45	1.9				
Not reported		1.54	0.42	5.57				
Vascular invasion	0.06							
Negative		Reference						
Positive		0.99	0.61	1.63				
Not reported		1.85	1.06	3.23				

BMI, Body mass index; AFP, Alpha fetoprotein; HR, Hazard ratio; CI, Confidence interval.

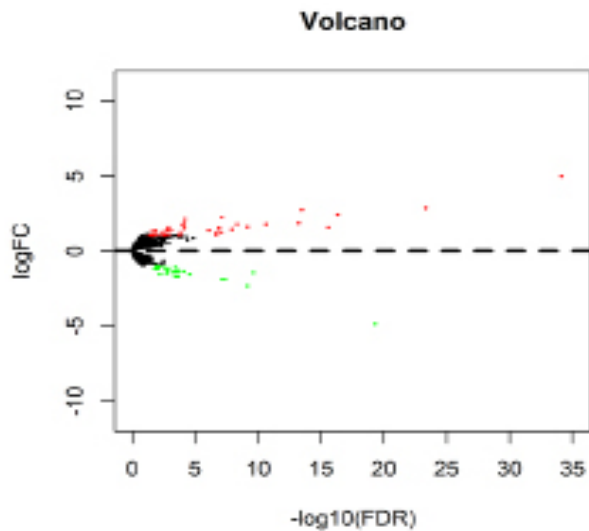


Figure 1. The Volcano Map of the Differentially Expressed lncRNAs in HCC Patients between with Tumor and without Tumor. Red dots represent upregulated genes, and green dots represent downregulated genes.

(Appendices 1-2). Pathways and functional enrichment were analyzed for protein-coding mRNAs using the GO and KEGG systems (Figures 6-7). These mRNAs were enriched mainly in complement and coagulation cascades,

the peroxisome, and the cell cycle.

Discussion

HCC is the most common and malignant type of liver cancer, and it is associated with high morbidity and poor prognosis. Carcinogenesis in HCC is a multistep process involving sustained inflammatory damage. The molecular heterogeneity of the disease has thwarted attempts at classification, making it difficult to predict prognosis and recurrence (Forner et al., 2018). The presence of a tumor in patients may be linked to poor outcomes (Li et al., 2019; Lim et al., 2019), so we hypothesized that we might be able to develop a reliable prognostic signature for predicting OS or RFS among HCC patients with tumors.

Many studies have reported that tumor size in HCC affects prognosis (Wu et al., 2018a; Chen et al., 2019b; Zeng et al., 2019), but we unaware of studies that have identified tumor status as a factor that independently affects OS or RFS and that have used lncRNA expression to create a risk scoring system to predict prognosis specifically for tumor-positive patients. Using data from 388 HCC patients in the TCGA, we found through uni- and multivariate analysis that tumor status was significantly associated with survival. After comprehensive analysis of DELncRNAs, we selected 15 HCC-specific lncRNAs

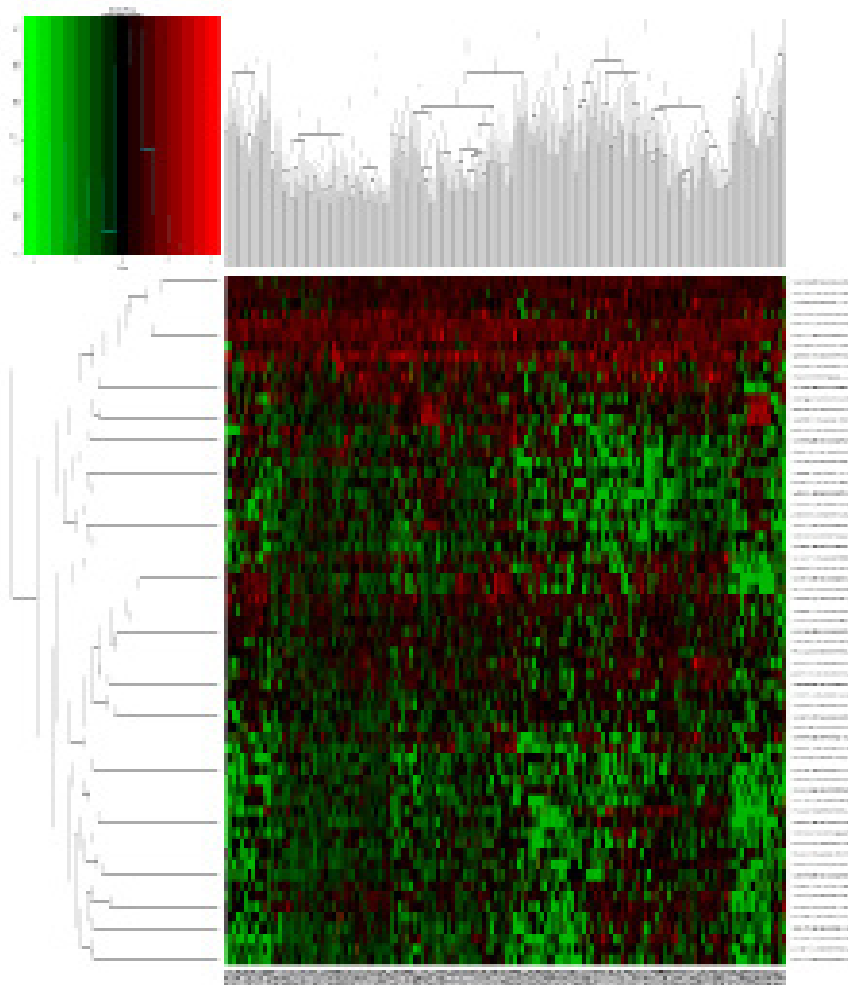


Figure 2. Heatmap Based on the Differentially Expressed lncRNAs in HCC Patients between with Tumor and without Tumor.

Table 3. Univariate and Multivariate Cox Regression Analysis for RFS in 338 HCC Patients

Variables	Univariate Cox regression				Multivariate Cox regression			
	P-value	HR	95% CI		P-value	HR	95% CI	
Age (>60/≤60)	0.09	0.68	0.44	1.06				
BMI	0.78							
<25		Reference						
≥25		1.15	0.68	1.95				
Not reported		0.9	0.34	2.38				
Race	0.13							
Non-Asian		Reference						
Asian		1.26	0.66	2.39				
Not reported		3.97	1	15.68				
AFP	0							
≤20ng/mL		Reference				Reference		
>20ng/ml		2.81	1.65	4.76	0.04	1.55	1.01	2.35
Not reported		2.09	1.06	4.11	0.35	1.25	0.78	1.98
Gender (Male/Female)	0.42	1.26	0.72	2.2				
Hepatitis B or C	0.61							
No		Reference						
Yes		0.75	0.42	1.35				
Not reported		0.97	0.35	2.7				
Alcohol	0.36							
No		Reference						
Yes		0.77	0.44	1.34				
Cirrhosis	0.32							
No		Reference						
Yes		1.4	0.8	2.42				
Not reported		1.5	0.77	2.91				
Histologic grade	0.35							
G1-2		Reference						
G3-4		0.86	0.57	1.3				
Not reported		2.46	0.58	10.39				
New tumor event	0.75	4.75E+05	0	2.91E+40				
Pathologic stage	0							
Stage I+II		Reference				Reference		
Stage III+IV		2.42	1.45	4.03	0	1.99	1.32	3.02
Not reported		0.36	0.14	0.96	0.08	0.54	0.27	1.08
Tumor Status	0	2.66	1.49	4.75	0	15.21	9.19	25.16
Family cancer history	0.25							
No		Reference						
Yes		0.98	0.58	1.66				
Not reported		0.49	0.21	1.17				
Residual tumor	0.68							
R0		Reference						
Non-R0		0.91	0.42	2				
Not reported		2.47	0.29	20.87				
Vascular invasion	0.06							
Negative		Reference						
Positive		1.33	0.8	2.2				
Not reported		0.58	0.28	1.19				

BMI, Body mass index; AFP, Alpha fetoprotein; HR, Hazard ratio; CI, Confidence interval.

Table 4. Six lncRNAs Correlated with OS of HCC Patients with Tumor in the Best Statistical Model.

lncRNA	β	HR	z	P-value
AC090921.1	0.2047	1.2271	2.88	0.0039
AC012640.1	0.144	1.1549	1.5	0.1325
AL158839.1	0.235	1.2649	2.31	0.0211
AL356056.1	0.3872	1.4728	4.2	2.70E-05
AL359853.1	0.2318	1.2609	2.57	0.0101
C10orf91	0.1328	1.142	2.57	0.0102

HR, Hazard ratio

that we analyzed in greater depth, and we found that six were related to OS and nine to RFS. We used these subsets of lncRNAs to develop risk scoring systems that were significantly associated with OS and RFS independently of other clinical characteristics. The scoring systems

showed good prognostic performance for tumor-positive patients, giving AUCs of 0.76 for OS and 0.861 for RFS. In addition, we found that AFP was a risk factor for poor RFS among patients with tumors, in accordance with previous studies (Jin et al., 2017; Notarpaolo et al., 2017). Enrichment analysis of genes co-expressed with the prognostic lncRNAs revealed that genes associated with OS of HCC patients with tumors may be related to complement and coagulation cascades, while those associated with RFS may be related to the cell cycle.

The lncRNA expression-based risk system is simple to calculate, and it quantifies mortality risk as a number easily understood by patients. Our system is based on lncRNAs differentially expressed between patients with or without tumors, so it may be more specific to HCC patients with tumors than other risk scoring systems based on lncRNAs differentially expressed between HCC and healthy liver

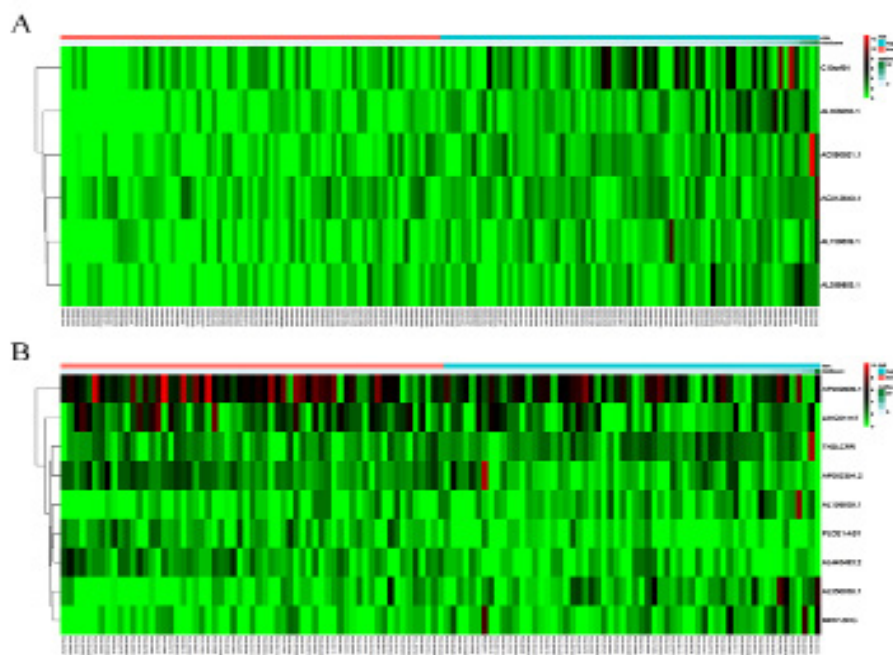


Figure 3. The Non-Cluster Risk Heat Map of lncRNA Expression-Based Risk Score System for OS(A) or RFS(B) in HCC Patients with Tumor. The value of risk rises gradually from left to right.

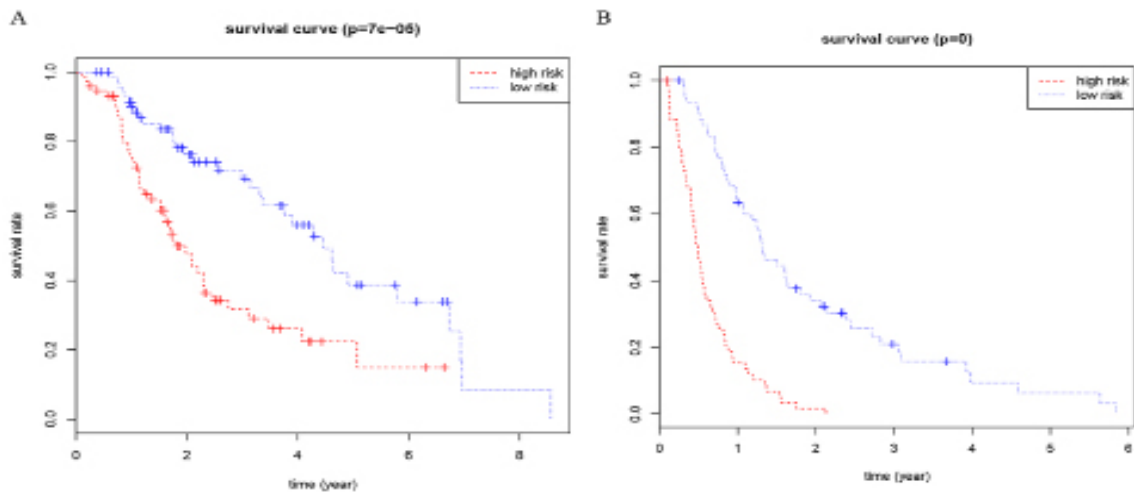


Figure 4. Kaplan-Meier Survival Curves for OS(A) or RFS(B) in HCC Patients with Tumor According to the Risk Cutoff point.

Table 5. Univariate and Multivariate Cox Regression Analysis for OS in HCC Patients with Tumor

Variables	P-value	Univariate Cox regression			Multivariate Cox regression			
		HR	95% CI		P-value	HR	95% CI	
Risk score(high/low)	0	2.59	1.38	4.87	0	2.32	1.43	3.77
Age (>60/≤60)	0.46	0.81	0.45	1.44				
BMI	0.79							
<25		Reference						
≥25		1	0.53	1.88				
Not reported	0.54	1.43	0.45	4.55				
Race	0.77							
Non-Asian		Reference						
Asian		1.1	0.48	2.55				
AFP	0.49	2.11	0.26	17.31				
≤20ng/mL		Reference						
>20ng/ml		1.77	0.84	3.73				
Not reported		2.42	1.04	5.63				
Gender (Male/Female)	0.14	1.63	0.85	3.14				
Hepatitis B or C	1							
No		Reference						
Yes		0.99	0.46	2.12				
Not reported		0	0	4.02E+251				
Alcohol	0.31							
No		Reference						
Yes		0.7	0.35	1.4				
Cirrhosis	0.52							
No		Reference						
Yes		1.6	0.61	4.22				
Not reported		1.39	0.65	2.95				
Histologic grade	0.5							
G1-2		Reference						
G3-4		1.05	0.55	1.99				
Not reported		0.31	0.04	2.28				
New tumor event	0.25							
No		Reference						
Yes		0.95	0.32	2.85				
Not reported		6.5	0.53	79.9				
Pathologic stage	0.12							
Stage I+II		Reference						
Stage III+IV		2	1.02	3.94				
Not reported		1.87	0.72	4.88				
Family cancer history	0.12							
No		Reference						
Yes		1.44	0.76	2.73				
Not reported		0.45	0.13	1.55				
Residual tumor	0.14							
R0		Reference						
Non-R0		0.3	0.09	1				
Not reported		0.59	0.05	6.42				
Vascular invasion	0							
Negative		Reference				Reference		
Positive		0.94	0.46	1.93	0.96	0.99	0.54	1.81
Not reported		3.23	1.51	6.91	0	2.57	1.53	4.32

BMI, Body mass index; AFP, Alpha fetoprotein; HR, Hazard ratio; CI, Confidence interval.

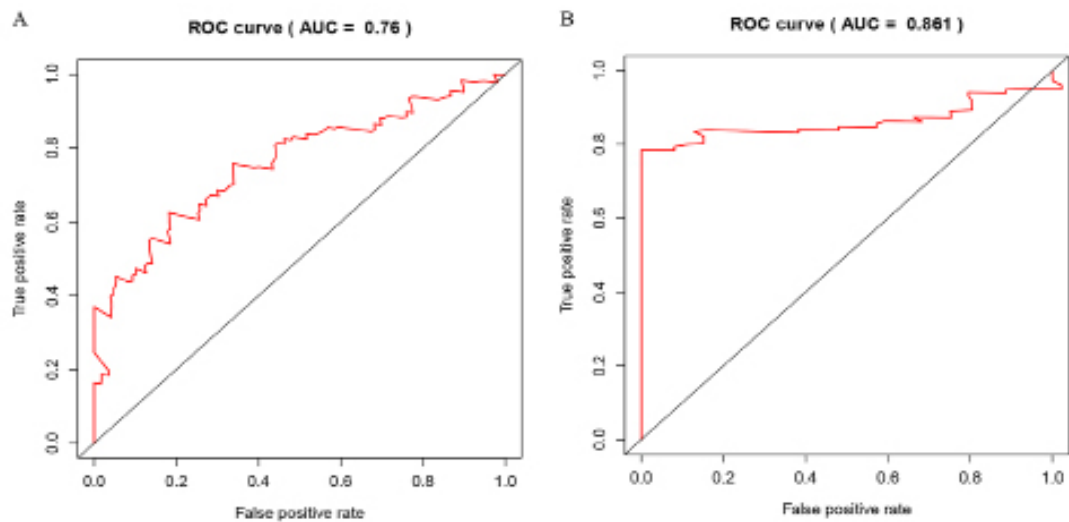


Figure 5. ROC Curves Analysis of the lncRNA Expression-Based Risk Score System for OS (A) or RFS (B) in HCC Patients with Tumor

Table 6. Nine lncRNAs Correlated with RFS of 338 HCC Patients with Tumor in the Best Statistical Model

lncRNA	β	HR	z	P-value
PLCE1-AS1	-0.4527	0.6359	-3.57	0.00035
AL158839.1	0.2805	1.3237	3.04	0.00233
AL445493.2	-0.2911	0.7474	-2.62	0.0089
TH2LCRR	0.1879	1.2068	2.05	0.04045
AP003354.2	-0.2184	0.8038	-2.23	0.02576
LINC01447	-0.0751	0.9276	-1.53	0.12538
AP000808.1	-0.0984	0.9062	-2.46	0.01379
AL356056.1	0.2502	1.2843	3.03	0.00246
MIR7-3HG	0.2882	1.3341	3.45	0.00056

HR, Hazard ratio

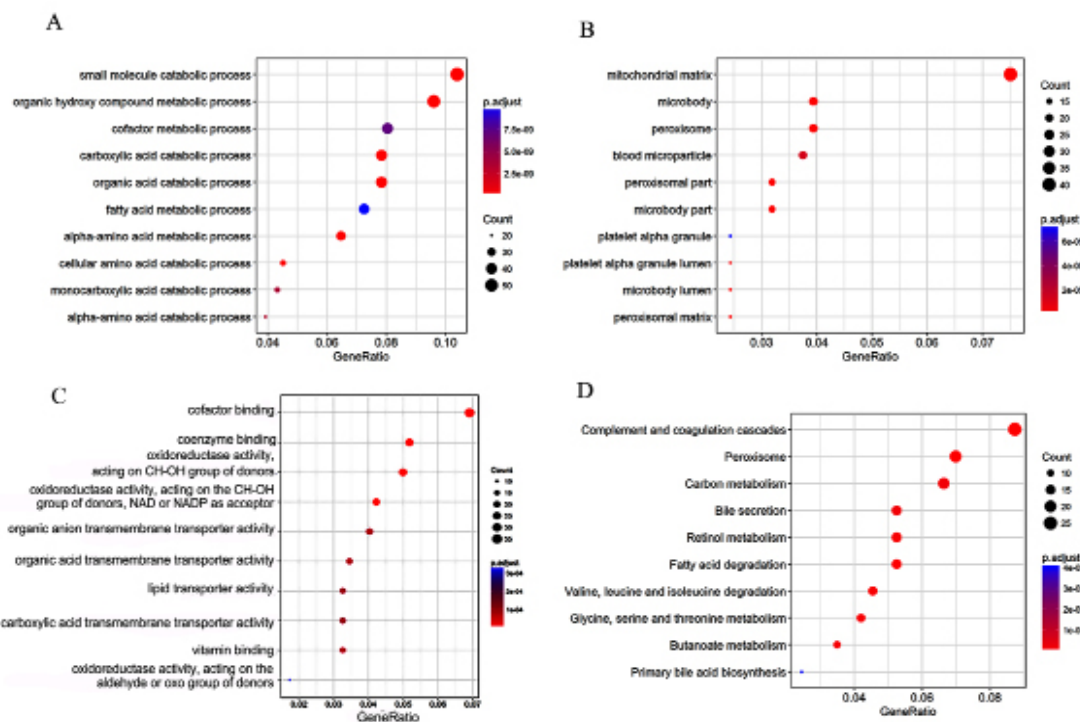


Figure 6. The Top 10 Significantly Enriched GO Annotation and Enriched KEGG Pathways of the Highly Related Genes of the lncRNAs of the Risk Score System for OS in HCC Patients with Tumor: (A) biological process, (B) cellular component, (C) molecular function. (D) KEGG pathway analysis.

Table 7. Univariate and Multivariate Cox Regression Analysis for RFS in HCC Patients with Tumor

Variables	Univariate Cox regression				Multivariate Cox regression			
	P-value	HR	95% CI		P-value	HR	95% CI	
Risk score(high/low)	0	3.45	2.05	5.82	0	3.86	2.48	6.01
Age (>60/≤60)	0.03	0.57	0.34	0.95				
BMI	0.74							
<25		Reference						
≥25		0.85	0.47	1.54				
Not reported		0.68	0.24	1.95				
Race	0.32							
Non-Asian		Reference						
Asian		1.21	0.6	2.43				
Not reported		3.54	0.64	19.68				
AFP	0.02							
≤20ng/mL		Reference				Reference		
>20ng/ml		2.22	1.21	4.09	0.02	1.75	1.09	2.82
Not reported		1.92	0.8	4.64	0.24	1.35	0.82	2.21
Gender (Male/Female)	0.92	0.97	0.54	1.76				
Hepatitis B or C	0.83							
No		Reference						
Yes		1.2	0.61	2.35				
Not reported		1.37	0.39	4.83				
Alcohol	0.75							
No		Reference						
Yes		0.91	0.5	1.65				
Cirrhosis	0.88							
No		Reference						
Yes		1.02	0.49	2.1				
Not reported		1.21	0.56	2.6				
Histologic grade	0.2							
G1-2		Reference						
G3-4		0.73	0.45	1.19				
Not reported		2.51	0.57	11.01				
New tumor event	0.96	3.53E+06	0	1.20E+273				
Pathologic stage	0.01							
Stage I+II		Reference				Reference		
Stage III+IV		1.8	1.02	3.19	0.13	1.44	0.9	2.29
Not reported		0.43	0.15	1.23	0.06	0.51	0.25	1.03
Family cancer history	0.05							
No		Reference						
Yes		0.84	0.46	1.52				
Not reported		0.33	0.13	0.81				
Residual tumor	0.89							
R0		Reference						
Non-R0		1.09	0.48	2.46				
Not reported		1.66	0.18	15.11				
Vascular invasion	0.01							
Negative		Reference				Reference		
Positive		1.96	1.09	3.49	0.37	1.23	0.78	1.94
Not reported		0.63	0.26	1.48	0.39	0.75	0.4	1.43

BMI, Body mass index; AFP, Alpha fetoprotein; HR, Hazard ratio; CI, Confidence interval.

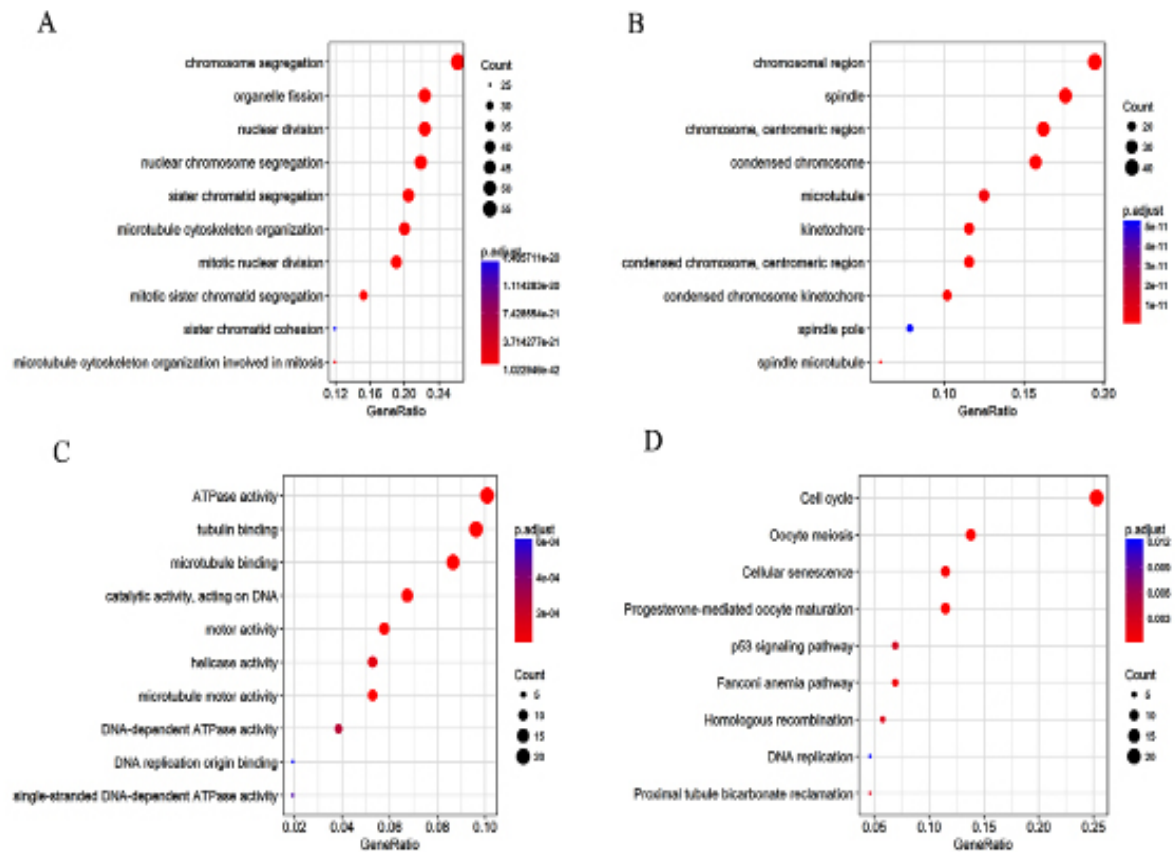


Figure 7. The Top 10 Significantly Enriched GO Annotation and Enriched KEGG Pathways of the Highly Related Genes of the lncRNAs of the Risk Score System for RFS in HCC Patients with Tumor: (A) biological process, (B) cellular component, (C) molecular function. (D) KEGG pathway analysis.

tissues(Gu et al., 2018; Liao et al., 2018; Ma et al., 2018; Shi et al., 2018; Sui et al., 2018; Wu et al., 2018c; Zhao et al., 2018; Yan et al., 2019). Indeed, the DELncRNAs in our risk scoring systems differ from those in the previously published systems. To our knowledge, the present study is the first to construct a risk scoring system to predict survival in HCC patients with tumors.

Nevertheless, our work should be interpreted and applied carefully because of several limitations. First, the publicly available data did not report on HCC treatments, so we could not take into account their potential confounding effects on OS and RFS. Second, our sample was small and lncRNAs were screened and validated for risk scoring within the same TCGA data set, so our system still needs to be validated externally. The scoring system should be assessed in other patient populations. Third, we did not divide up patients into training and testing data sets, nor did we perform Gene Set Enrichment Analysis (GSEA). Fourth, we did not perform experiments with tissues or cells to validate the differential expression or involvement of prognostic lncRNAs in HCC. These limitations need to be addressed in future studies.

In conclusion, We have derived a novel lncRNA expression-based risk scoring system for predicting the survival of HCC patients with tumors, which may be a clinically useful tool to individualize HCC therapy.

Author Contribution Statement

Siyao Wu: Formal analysis, Data Curation and Writing - Original Draft; Yayan Deng: Visualization and Writing - Original Draft; Yue Luo: Data Curation and Visualization; Zhihui Liu: Methodology, Validation and Writing - Review & Editing, Supervision; Jiexiang Ye: Conceptualization, Methodology, Writing - Review & Editing, Supervision and Funding acquisition

Acknowledgment

This work was supported by Guangxi Natural Science Foundation (2020GXNSFBA297059 and 2020GXNSFBA297024), and Guangxi First-class Discipline Project for Basic Medical Sciences (GXFCDP-BMS-2018).The funders had no role in the design of the study, collection or analysis of data, decision to publish, or preparation of the manuscript. This work was not approved as a part of an student thesis by any scientific Body.

Appendix

Appendix 1: The results of lncRNA-mRNA co-expression analysis for predicting OS in tumor positive patients.

Appendix 2: The results of lncRNA-mRNA co-expression analysis or predicting RFS in tumor positive patients.

Ethical Statement

Since the data is from the TCGA, our study did not require an ethical board approval.

Declaration of Competing Interest

The authors have no conflicts of interest to declare.

Data Availability Statement

Data are provided within the article.

References

- Bertuccio P, Turati F, Carioli G, et al (2017). Global trends and predictions in hepatocellular carcinoma mortality. *J Hepatol*, **67**, 302-9.
- Bruix J, Gores GJ, Mazzaferro V (2014). Hepatocellular carcinoma: clinical frontiers and perspectives. *Gut*, **63**, 844-55.
- Chen Z, Yu W, Zhou Q, et al (2019). A novel lncRNA IHS promotes tumor proliferation and metastasis in HCC by regulating the ERK- and AKT/GSK-3beta-signaling pathways. *Mol Ther Nucleic Acids*, **16**, 707-20.
- Chen ZX, Jian ZW, Wu XW, et al (2019). Clinical conditions and treatment requirements for long-term survival among hepatitis B-related hepatocellular carcinoma initially treated with chemoembolization. *Cancer Med*, **8**, 5097-107.
- Fernández-Barrena M, Perugorria M, Banales J (2017). Novel lncRNA T-UCR as a potential downstream driver of the Wnt/ β -catenin pathway in hepatobiliary carcinogenesis. *Gut*, **66**, 1177-8.
- Forner A, Reig M, Bruix J (2018). Hepatocellular carcinoma. *Lancet (London, England)*, **391**, 1301-14.
- Gu J, Zhang X, Miao R, et al (2018). A three-long non-coding RNA-expression-based risk score system can better predict both overall and recurrence-free survival in patients with small hepatocellular carcinoma. *Aging*, **10**, 1627-39.
- Hara K, Takeda A, Tsurugai Y, et al (2019). Radiotherapy for hepatocellular carcinoma results in comparable survival to radiofrequency ablation: A Propensity Score Analysis. *Hepatology*, **69**, 2533-45.
- Huang Y, Xiang B, Liu Y, et al (2018). LncRNA CDKN2B-AS1 promotes tumor growth and metastasis of human hepatocellular carcinoma by targeting let-7c-5p/NAP1L1 axis. *Cancer Lett*, **437**, 56-66.
- Jin J, Zhang X, Shi J, et al (2017). Application of AFP whole blood one-step rapid detection kit in screening for HCC in Qidong. *Am J Cancer Res*, **7**, 1384-8.
- Kang KJ, Ahn KS (2017). Anatomical resection of hepatocellular carcinoma: A critical review of the procedure and its benefits on survival. *World J Gastroenterol*, **23**, 1139-46.
- Kulik L, El-Serag HB (2019). Epidemiology and management of hepatocellular carcinoma. *Gastroenterology*, **156**, 477-91 e1.
- Li Y, Ma B, Yin Z, et al (2019). Competing endogenous RNA network and prognostic nomograms for hepatocellular carcinoma patients who underwent R0 resection. *J Cell Physiol*, **234**, 20342-53.
- Liao X, Yang C, Huang R, et al (2018). Identification of potential prognostic long non-coding RNA biomarkers for predicting survival in patients with hepatocellular carcinoma. *Cell Physiol Biochem*, **48**, 1854-69.
- Lim C, Salloum C, Lahat E, et al (2019). Impact of narrow margin and R1 resection for hepatocellular carcinoma on the salvage liver transplantation strategy. An intention-to-treat analysis. *HPB (Oxford)*, **21**, 1295-302.
- Ma Y, Luo T, Dong D, et al (2018). Characterization of long non-coding RNAs to reveal potential prognostic biomarkers in hepatocellular carcinoma. *Gene*, **663**, 148-56.
- Notarapalo A, Layese R, Magistri P, et al (2017). Validation of the AFP model as a predictor of HCC recurrence in patients with viral hepatitis-related cirrhosis who had received a liver transplant for HCC. *J Hepatol*, **66**, 552-9.
- Ponting CP, Oliver PL, Reik W (2009). Evolution and functions of long noncoding RNAs. *Cell*, **136**, 629-41.
- Pu J, Wang J, Wei H, et al (2019). lncRNA MAGI2-AS3 prevents the development of HCC via recruiting KDM1A and promoting H3K4me2 demethylation of the RACGAP1 promoter. *Mol Ther Nucleic Acids*, **18**, 351-62.
- Robinson MD, McCarthy DJ, Smyth GK (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, **26**, 139-40.
- Schmitt AM, Chang HY (2016). Long noncoding RNAs in cancer pathways. *Cancer Cell*, **29**, 452-63.
- Shi Y, Li Y, Lin J, et al (2018). The discovery of a novel eight-mRNA-lncRNA signature predicting survival of hepatocellular carcinoma patients. *J Cell Biochem*, **120**, 7539-50.
- Siegel RL, Miller KD, Jemal A (2020). Cancer statistics, 2020. *CA Cancer J Clin*, **70**, 7-30.
- Sui J, Miao Y, Han J, et al (2018). Systematic analyses of a novel lncRNA-associated signature as the prognostic biomarker for Hepatocellular Carcinoma. *Cancer Med*, **7**, 3240-56.
- Sun L, Wang L, Chen T, et al (2019). microRNA-1914, which is regulated by lncRNA DUXAP10, inhibits cell proliferation by targeting the GPR39-mediated PI3K/AKT/mTOR pathway in HCC. *J Cell Mol Med*, **23**, 8292-304.
- Torzilli G, Belghiti J, Kokudo N, et al (2013). A snapshot of the effective indications and results of surgery for hepatocellular carcinoma in tertiary referral centers: is it adherent to the EASL/AASLD recommendations?: an observational study of the HCC East-West study group. *Ann Surg*, **257**, 929-37.
- Villanueva A, Hernandez-Gea V, Llovet JM (2013). Medical therapies for hepatocellular carcinoma: a critical view of the evidence. *Nat Rev Gastroenterol Hepatol*, **10**, 34-42.
- Wu J, Zhu P, Zhang Z, et al (2018). A new tumor-associated antigen prognostic scoring system for spontaneous ruptured hepatocellular carcinoma after partial hepatectomy. *Cancer Biol Med*, **15**, 415-24.
- Wu L, Pan C, Wei X, et al (2018). LncRNA KRAL reverses 5-fluorouracil resistance in hepatocellular carcinoma cells by acting as a ceRNA against miR-141. *Cell Commun Signal*, **16**, 47.
- Wu Y, Wang PS, Wang BG, et al (2018). Genomewide identification of a novel six-lncRNA signature to improve prognosis prediction in resectable hepatocellular carcinoma. *Cancer Med*, **7**, 6219-33.
- Xu J, Lu Y, Liu Q, et al (2020). Long noncoding RNA GMAN promotes hepatocellular carcinoma progression by interacting with eIF4B. *Cancer Lett*, **473**, 1-12.
- Yan J, Zhou C, Guo K, et al (2019). A novel seven-lncRNA signature for prognosis prediction in hepatocellular carcinoma. *J Cell Biochem*, **120**, 213-23.
- Yu G, Wang LG, Han Y, He QY (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*, **16**, 284-7.
- Zeng Z, Dong J, Li Y, et al (2019). The expression level and clinical significance of lncRNA X91348 in hepatocellular carcinoma. *Artif Cells Nanomed Biotechnol*, **47**, 3067-71.
- Zhao QJ, Zhang J, Xu L, Liu FF (2018). Identification of a five-long non-coding RNA signature to improve the prognosis prediction for patients with hepatocellular carcinoma. *World J Gastroenterol*, **24**, 3426-39.
- Zhong JH, Ke Y, Gong WF, et al (2014). Hepatic resection associated with good survival for selected patients with

intermediate and advanced-stage hepatocellular carcinoma.
Ann Surg, **260**, 329-40.



This work is licensed under a Creative Commons Attribution-Non Commercial 4.0 International License.