

The identification and preliminary study of lncRNA TUG1 and its related genes in hepatocellular carcinoma

Lei Li^{1,2}, Shuiping Liu², Li Peng³, Yapeng Zhang², Yaqin Zhang^{1,4}, Haiyan Zeng², Guancheng Li^{1,4}, Chaoyang Zhang^{4,5}

¹Key Laboratory of Carcinogenesis of the Chinese Ministry of Health and the Key Laboratory of Carcinogenesis and Cancer Invasion of Chinese Ministry of Education, Xiangya Hospital, Central South University, Changsha, China

²School of Basic Medical Science, Central South University, Changsha, China

³Guangdong Province Key Laboratory of Malignant Tumor Epigenetics and Gene Regulation, Research Center of Medicine, Sun Yat-Sen Memorial Hospital, Sun Yat-Sen University, Guangzhou, China

⁴Cancer Research Institute, Central South University, Changsha, China

⁵Division of Functional Genome Analysis, German Cancer Research Center (DKFZ), Heidelberg, Germany

Submitted: 3 January 2019; **Accepted:** 31 May 2019

Online publication: 12 November 2019

Arch Med Sci 2022; 18 (6): 1582–1595

DOI: <https://doi.org/10.5114/aoms.2019.89707>

Copyright © 2019 Termedia & Banach

Abstract

Introduction: Hepatocellular carcinoma (HCC) is a common malignant tumour of the digestive system, which is a threat to public health. The purpose of this study was to investigate the featured genes and pathways of HCC from a bioinformatics database, and verify their correlation with diagnosis and prognosis of HCC.

Material and methods: We downloaded the gene expression profile on HCC from the Gene Expression Omnibus (GEO), and software R was used to identify differentially expressed lncRNA (DEL). The target genes of the lncRNA were further predicted by using a cluster database and molecular interaction database. Subsequently, a combined interaction network of target genes was constructed using the Cytoscape platform with preliminary verification at the level of different databases, cell lines, and tissues. Finally, we explored the effectiveness of TUG1 and its target genes on the diagnosis and prognosis of HCC by univariate Cox analysis and survival analysis.

Results: A total of four DELs were identified and the most remarkably up-regulated lncRNA was TUG1, which included 12 high-confidence target genes. Moreover, we found that the expression changes of TUG1 and its target genes in different databases, cell lines, and liver cancer tissues were consistent with the prediction. The high expression of TUG1 and its target genes could significantly predict the shorter survival time of HCC patients, among which NCAPG, MCM6, PIGC, PEA15, and RACGAP1 have significant diagnostic value for HCC (AUC > 0.9).

Conclusions: This study provides a starting point for the screening of therapeutically relevant targets in HCC. Further experiment should be conducted to verify our findings.

Key words: hepatocellular carcinoma, TUG1, VGLL4, HMG4, PIGC.

Corresponding authors:

Guancheng Li
Key Laboratory
of Carcinogenesis
of the Chinese Ministry
of Health and the Key
Laboratory of Carcinogenesis
and Cancer Invasion
of the Chinese Ministry
of Education
Xiangya Hospital
Central South University
87 Xiangya Road
410008, Changsha, China
E-mail: ligc61@csu.edu.cn.

Chaoyang Zhang
Division of Functional
Genome Analysis
German Cancer Research
Center (DKFZ)
Neuenheimer Feld 580
69120, Heidelberg, Germany
E-mail: chaoyang.zhang@
dkfz-heidelberg.de

Introduction

Liver cancer, especially hepatocellular carcinoma (HCC) is one of the 10 most common malignant tumours in the world, with high morbidity

and mortality. The occurrence of HCC is a dynamic process of multigene involvement and multistage development. Liver cirrhosis induced by hepatitis and fatty liver is the main pathological basis [1, 2]. However, early patients with HCC often show atypical symptoms, therefore, an in-depth study on the key nodes will provide the novel strategies for the early diagnosis and targeted therapy of HCC.

Long non-coding RNAs (lncRNA), a class of non-coding RNA, participate in the regulation of gene expression in different ways, such as epigenetic, transcriptional, and post-transcriptional regulation [3]. Abnormal expression of lncRNA is reported to be involved in the control of a broad range of biological processes, thereby impacting differentiation, metabolism, and aging [4]. Meanwhile, lncRNA may act as a diagnostic marker in tumours and indicate the early prognosis and therapeutic response of tumour cells [5]. Compared with protein marker α -fetoprotein (AFP), lncRNA UCA1 has higher diagnostic efficacy in liver cancer. When combined UCA1 and JUN mRNA were used in the diagnosis of early HCC, the sensitivity and specificity were as high as 100% and 80%, respectively [6, 7].

So far, most studies show that lncRNA plays an extremely important role in HCC. However, the exact mechanism remains to be discovered. In this study, we screened for differentially expressed lncRNAs between normal and cancer samples by microarray analysis in HCC. Subsequently, we chose the most differentially expressed lncRNA TUG1 for further investigation. Our findings suggest that TUG1 and some of its target genes may contribute to the shorter survival time and poor prognosis. Our results may provide a clear direction for early diagnosis and treatment of HCC.

Material and methods

Microarray data

The GeneChip dataset of GSE14323 was obtained from the GEO database, which was employed to store second-generation sequencing and other forms of high-throughput functional genomics data. After quality inspection of the chip dataset, a total of 18 normal samples and 28 samples of HCC were retained for analysis.

A cohort from the TCGA (<https://tcga-data.nci.nih.gov/tcga/>) database containing 371 HCC patients and 50 controls as well as follow-up information, were included in our study.

Screening of differentially expressed lncRNAs

The statistical software R was used to screen for differentially expressed lncRNAs in the GSE14323 dataset. Fold change more than 2 and raw *p*-values less than 0.05 were selected as the cut-off criterion.

Identifying target genes of the differentially expressed lncRNA-TUG1

Firstly, screening and predicting target genes was performed, directly or indirectly regulated by TUG1, from common molecular interaction databases including starBase v2.0, lncBase v.2, miRTarBase, TarBase v7.0, and RAID v2.0.

Then Pearson's correlation test was applied to filter the genes with significant association with TUG1 in the GSE14323 dataset and TCGA transcriptome data. Subsequently, the differentially expressed TUG1 target genes related to prognosis were screened based on the clinical information and gene expression profile of 370 patients with HCC in the TCCA database by survival analysis. The Limma package of statistical analysis software R was used to assess the differential co-expression genes in the GSE14323 dataset. A Wayne map was conducted for the first intersection of the above results. The raw *p*-values less than 0.05 were selected as statistically significant.

Finally, to obtain more reliable target genes, the two parts were combined with each other to retrieve the differentially expressed target genes of TUG1.

Construction of lncRNA-target gene interaction network

The Cytoscape platform was used to process biological network and data integration.

Cell lines and human HCC samples

Human liver cancer cell lines HepG2, Huh7, QGY-7703, Hep3B, and human normal liver cell line QSG-7701 were preserved in our laboratory. All cell lines were maintained in DMEM (Gibco, BRL) supplemented with 10% foetal bovine serum (FBS) (Gibco, BRL) in a humidified atmosphere containing 5% CO₂ at 37°C. Fifteen cases of normal liver tissues and liver cancer tissue paraffin sections were obtained from the Department of Pathology of Xiangya Hospital. Five cases of normal liver tissues and 5 cases of liver cancer tissues were obtained from the key laboratory of tumour proteomics of the Ministry of Health of Xiangya Hospital and the Department of Hepatobiliary and Pancreatic surgery of Xiangya Hospital.

Correlation analysis of TUG1 and its target gene expression

Software R, SPSS18.0, and Graphad Prism version 5 were used for statistical analysis in the GSE14323 dataset and TCGA transcriptome data. qRT-PCR and siRNA transfection were applied in the abovementioned cells and tissues, which were detailed in supplementary materials. The primers of GAPDH, TUG1, HMGN4, VGLL4, and PIGC are

Table I. Sequences of primers and short RNA oligos

Primer name	Sequence
GAPDH	F: GGGAGCCAAAAGGGTCATC
	R: GAGTCCTTCCACGATACCAAA
TUG1	F: TAGCAGTTCCTCCCAATCCTTG
	R: CCACAAATCCCATCATTCCC
HMG14	F: CTCGCTCGTCTTCTCTGTC
	R: GGTGCTTGTCTCCTCAC
VGLL4	F: AACTGCAACTCTCGACTG
	R: GCTCGGGCTCCTTGAATTCT
PIGC	F: GCAACCTAGGAATCTCGG
	R: TTCAAAGGCAGGGGCTAGG
siTUG1	5' to 3': CUUGGCUUCUAUUCUGAAUCCUUU
Negative control	5' to 3': UUCUCCGAACGUGUCACGUTT

listed in Table I. The siTUG1 and negative control were synthesised by Genechem, and the sequences are listed in Table I. Pearson’s correlation test was conducted to assess the correlation of TUG1 and its target genes in the above cell lines and tissues; the raw *p*-values less than 0.05 were recognised as statistically significant.

Predicting the effects of TUG1 and its target genes on the diagnosis and prognosis of HCC

The diagnosis of HCC was performed using clinical criteria and the findings obtained by B-mode ultrasonography (US), computed tomography (CT) angiography, or magnetic resonance imaging (MRI) [8, 9].

The statistical analysis was performed by SPSS 18.0 software or Graphad Prism version 5 in the TCCA cohort. The continuous data were presented as the mean ± standard deviation (SD). The mean values of two samples were compared by independent sample *t* test. Kaplan-Meier survival analysis was carried out by the log-rank method. The survival curve was used to describe the correlation between the target gene expression and the patients’ survival. Pearson correlation analysis was applied to analyse the expression of target genes and tumour stages, and receiver operating characteristic (ROC) curve analysis was performed to describe their diagnostic value for patients with liver cancer. The raw *p*-values less than 0.05 were recognised as statistically significant.

Supplementary data

Extraction and reverse transcription of RNA

Two micrograms of RNA were extracted from normal liver cell line QSG7701 and liver cancer cell line HepG2, Huh7, Hep3B, and QGY7703 for reverse transcription. One microlitre of Oligo (dT) primer was added in combination with DEPC water to a total volume of 12 µl, placed in a PCR instrument, and treated at 65°C for 5 min. Then, 1 µl reverse transcriptase inhibitor, 2 µl dNTP Mixture, 4 µl 5 × PrimeScript Buffer, and 1 µl reverse transcriptase was added to the PCR tube. The procedure was as follows: 42°C, 1 h and 70°C, 5 min; the reverse transcription cDNA solution was stored at 4°C.

The system of qPCR reaction was 15 µl, containing 0.3 µl cDNA solution, 6.6 µl aseptic water, 0.6 µl upstream and downstream primers, and

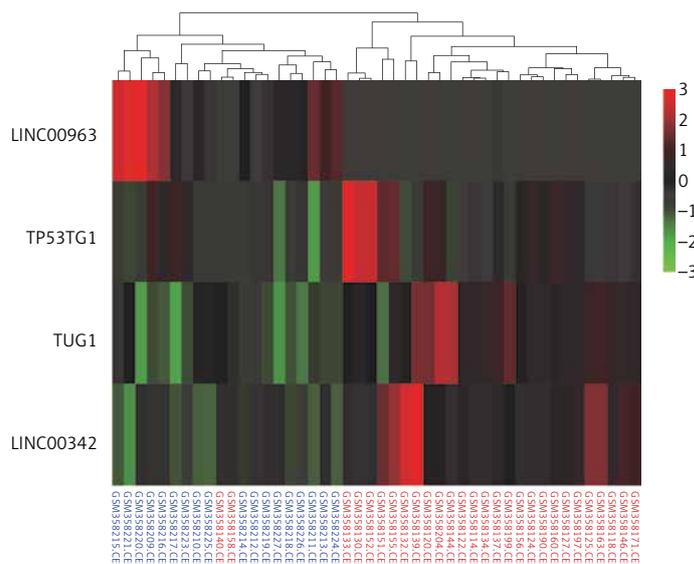


Figure 1. Heat map showing the top four DELs. The rows of the heat map represent genes, and the columns represent samples. Each column is coloured based on the expression level of that gene in that sample. Red bar: up-regulation; green bar: down-regulation; blue labelled sample: controls; red labelled sample: HCC

7.5 μ l qPCR mixture. The total reaction system was calculated according to the cell type and gene number, and the corresponding premixed solution was added to the octahedral tube, then centrifuged and placed in the qPCR instrument. The specific procedure of fluorescence quantitative PCR with two-step method was as follows: denaturation at 95°C for 5 min; denaturation at 95°C for 30 s and annealing at 60°C for 30 s, a total of 40 cycles. The differential expression of target gene was calculated by $2^{-\Delta\Delta C_T}$ method.

siRNA transfection

Hep3B and HepG2 cells were seeded in 96-well plates with 2.5×10^5 cells per well. Lipofectamine 3000 reagent was used to transfect siTUG1 to the cells when the confluence was 40–50%.

Results

Prediction of lncRNA TUG1 target genes and construction of interaction network

Differentially expressed lncRNAs

The top four differentially expressed lncRNAs (DEs) were identified in 18 normal samples

and in 28 samples of HCC from the GSE14323 dataset, of which the lncRNA LINC00963 was underexpressed and the rest were significantly

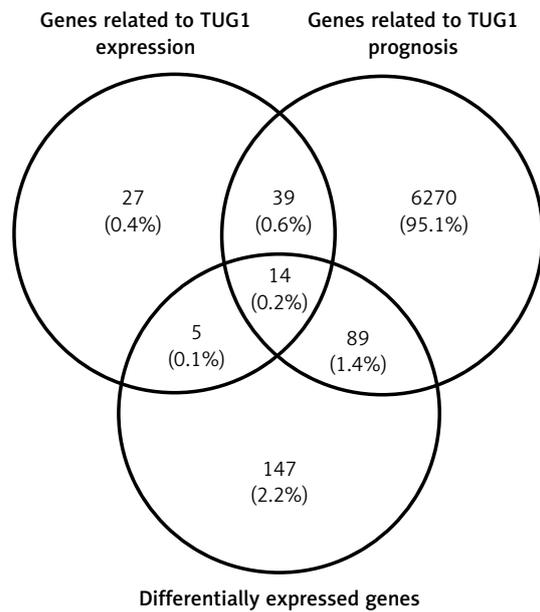


Figure 2. Differentially expressed genes related to TUG1 expression and prognosis

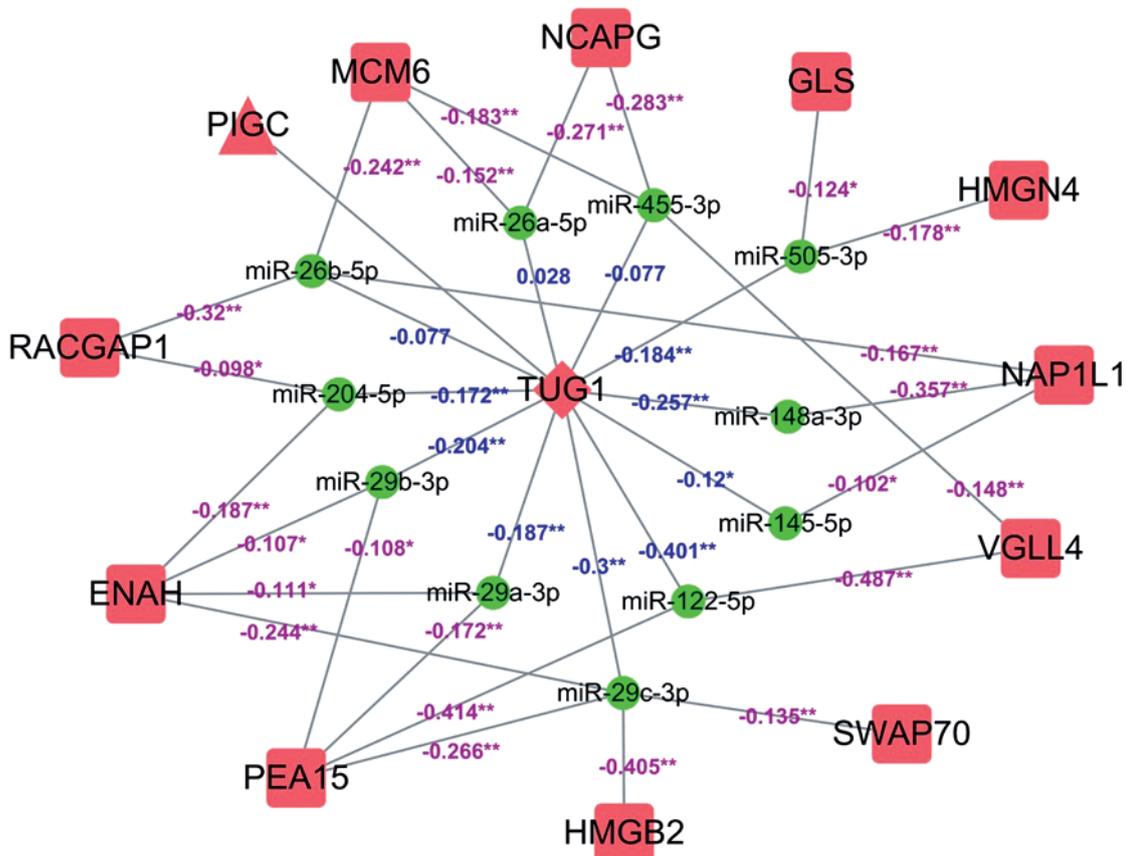


Figure 3. The molecular regulatory network for potential TUG1-driven genes. Square: mRNA; triangle: protein; round: miRNA; rhombus: lncRNA; red: significantly up-regulated expression; green: significantly down-regulated expression; purple number: correlation between miRNA and mRNA expression; blue number: correlation between TUG1 and miRNA expression; * $p < 0.05$; ** $p < 0.01$

overexpressed in HCC ($p < 0.05$). TUG1 was the only lncRNA, with more than double up-regulated expression (Figure 1).

Identification of TUG1-target genes

A total of 56 miRNAs (including 31 up-regulated and 25 down-regulated miRNAs), which were directly regulated by TUG1, were screened from lncRNA-miRNA interaction databases: RAID, starBase v2.0, and lncBase v2.1. Putative miRNA targets were predicted using miRTarBase and TarBase v7.0, and a total of 5796 TUG1-related genes were obtained from the databases.

After analysing the GSE14323 gene expression profile and TCGA database, respectively, we found 374 genes and 5476 genes with significant correlation with TUG1 ($p < 0.05$, $|R| > 0.5$). Furthermore, 85 target gene groups associated with TUG1 expression were obtained after the first intersection ($p < 0.05$, $|R| > 0.5$). Through survival analysis of 370 patients with HCC in the TCGA database, 6412 genes were screened out, which were significantly related to undefined patient survival time ($p < 0.05$). In order to find the TUG1 target genes for the diagnosis of HCC, differential expression analysis was performed on the GSE14323 data-

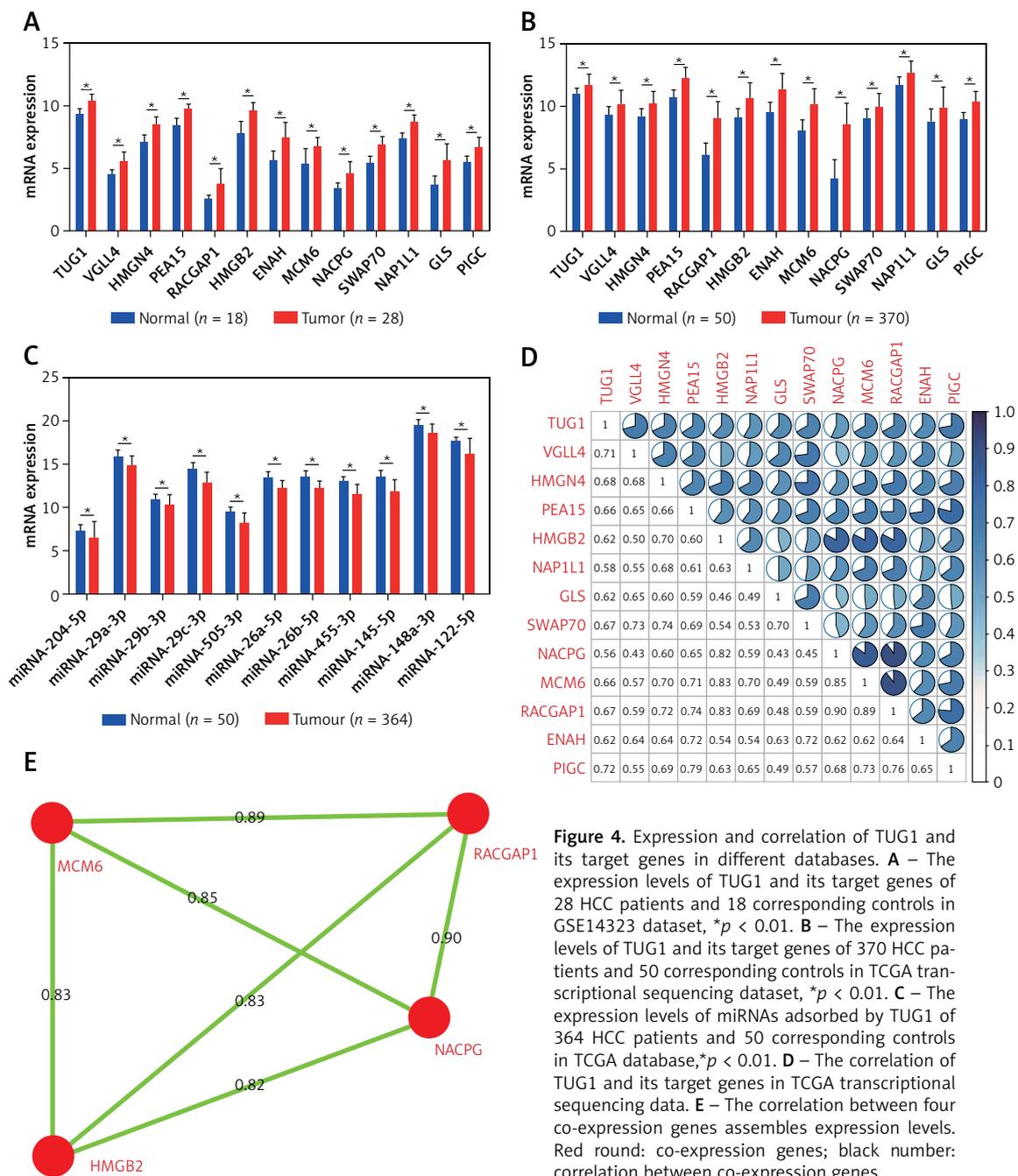


Figure 4. Expression and correlation of TUG1 and its target genes in different databases. **A** – The expression levels of TUG1 and its target genes of 28 HCC patients and 18 corresponding controls in GSE14323 dataset, $*p < 0.01$. **B** – The expression levels of TUG1 and its target genes of 370 HCC patients and 50 corresponding controls in TCGA transcriptional sequencing dataset, $*p < 0.01$. **C** – The expression levels of miRNAs adsorbed by TUG1 of 364 HCC patients and 50 corresponding controls in TCGA database, $*p < 0.01$. **D** – The correlation of TUG1 and its target genes in TCGA transcriptional sequencing data. **E** – The correlation between four co-expression genes assembles expression levels. Red round: co-expression genes; black number: correlation between co-expression genes

set and TCGA cohort. Finally, 211 co-up-regulated differentially expressed genes ($p < 0.05$ FC2) and 44 co-down-regulated differentially expressed genes (FC2) were screened out. Thereby, we constructed a Wayne map for subsequent intersection, from which 14 significant target genes were identified (Figure 2).

Construction of lncRNA TUG1 and its target gene interaction network

A total of 12 differentially expressed targeted genes related to prognosis of TUG1 were screened after the combination of 5796 and 14 genes mentioned above ($p < 0.05$, $R \geq 0.5$, $FC \geq 2$), of which 11 miRNAs were found to be down-regulated from TCGA transcriptome data ($p < 0.05$, $R < 0$). A regulatory network of TUG1 and its target genes were constructed with the help of Cytoscape software (Figure 3).

Correlation analysis of TUG1 and its target genes expression

Detecting correlation between TUG1 and its target genes in the

TUG1 database and its target genes were significantly up-regulated in 28 HCC patients from the GSE14323 dataset and 370 HCC patients in TCGA transcriptional sequencing data ($p < 0.05$, $FC \geq 2$, Figures 4 A, B). The above 11 miRNAs were

accordingly down-regulated in 364 HCC patients ($p < 0.05$, Figure 4 C).

The correlation of TUG1 and its target genes were analysed in 50 normal liver tissues and 370 liver cancer tissues from TCGA transcriptional sequencing data, and the results are shown in Figure 4 D ($p < 0.05$, $R \geq 0.4$). In addition, the top four co-expression gene assemblies including HMGB2 and NCAPG ($R = 0.82$), MCM6 ($R = 0.83$), RACGAP1 ($R = 0.83$), MCM6 and NCAPG ($R = 0.85$), RACGAP1 and NCAPG ($R = 0.9$), and MCM6 ($R = 0.89$) were also found (Figure 4 E). This suggests that HMGB2, NCAPG, MCM6, and RACGAP1 may be regulated by the same upstream regulatory factor, TUG1, and participate in the same or similar pathways in the pathogenesis of HCC.

Detecting correlation between TUG1 and its target genes in cell lines

The expression levels of TUG1, VGLL4, PIGC, and HMGN4 mRNA were detected in normal liver cell line and hepatoma cell lines, respectively, and the results of qRT-PCR showed that TUG1 and VGLL4 were highly expressed in HepG2, Huh7, Hep3B, and QGY7703 compared with normal liver cell line (QSG7701). However, as shown in Figure 5 A, HMGN4 and PIGC were down-regulated in HepG2 and Huh7, but up-regulated in QGY7703 and Hep3B. To further demonstrate the interaction of TUG1 and the three target

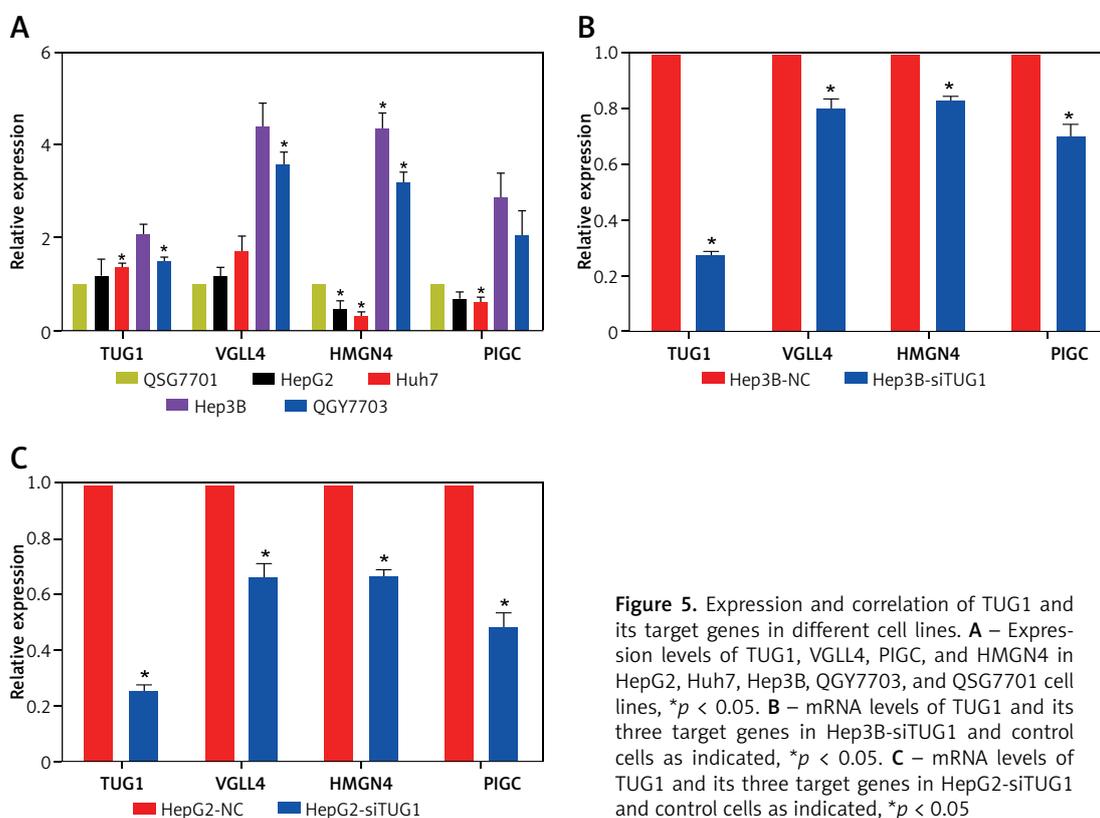


Figure 5. Expression and correlation of TUG1 and its target genes in different cell lines. **A** – Expression levels of TUG1, VGLL4, PIGC, and HMGN4 in HepG2, Huh7, Hep3B, QGY7703, and QSG7701 cell lines, $*p < 0.05$. **B** – mRNA levels of TUG1 and its three target genes in Hep3B-siTUG1 and control cells as indicated, $*p < 0.05$. **C** – mRNA levels of TUG1 and its three target genes in HepG2-siTUG1 and control cells as indicated, $*p < 0.05$

genes, siRNA was used to interfere with the expression of TUG1 in hepatoma cell lines (HepG2 and Hep3B). The results show that VGLL4, PIGC, and HMGN4 expression decreased significantly ($p < 0.05$, Figures 5 B, C), indicating that TUG1 could regulate the expression of VGLL4, PIGC, and HMGN4.

Detecting correlation between TUG1 and its target genes in tissues

To determine the effect of TUG1 over expression on the target genes, TUG1 and VGLL4, HMGN4, PIGC levels were assessed by qRT-PCR in five HCC tissues and five matched normal tissues, as shown in Figure 6 A. The expression levels of TUG1, VGLL4, and PIGC in HCC tissues were significantly higher than those in normal tissues ($p < 0.05$), while HMGN4 expression was down-regulated in HCC tissues ($p < 0.05$, Figure 6 A). Correlation analysis showed that TUG1 and VGLL4 had a higher correlation than the rest;

however, the p value was not significant, but this result requires more samples for validation (Figures 6 B–D).

The function of TUG1 target genes and their clinical significance

Prognostic and diagnostic values of TUG1 and its target genes in HCC

Based on TCGA transcriptome cohort of HCC-indicating patients, a univariate survival analysis was performed to access the clinical features (Table II). We summarised from Table II that features including age, sex, race, and body mass index (BMI) had no significant effect on survival time ($p > 0.05$), whereas TUG1 and its 12 target genes were independent prognostic risk factors according to univariate COX regression analysis (Table III). At the same time, Kaplan-Meier analysis showed that the higher expression of TUG1 and its 12 target genes could significantly predict shorter

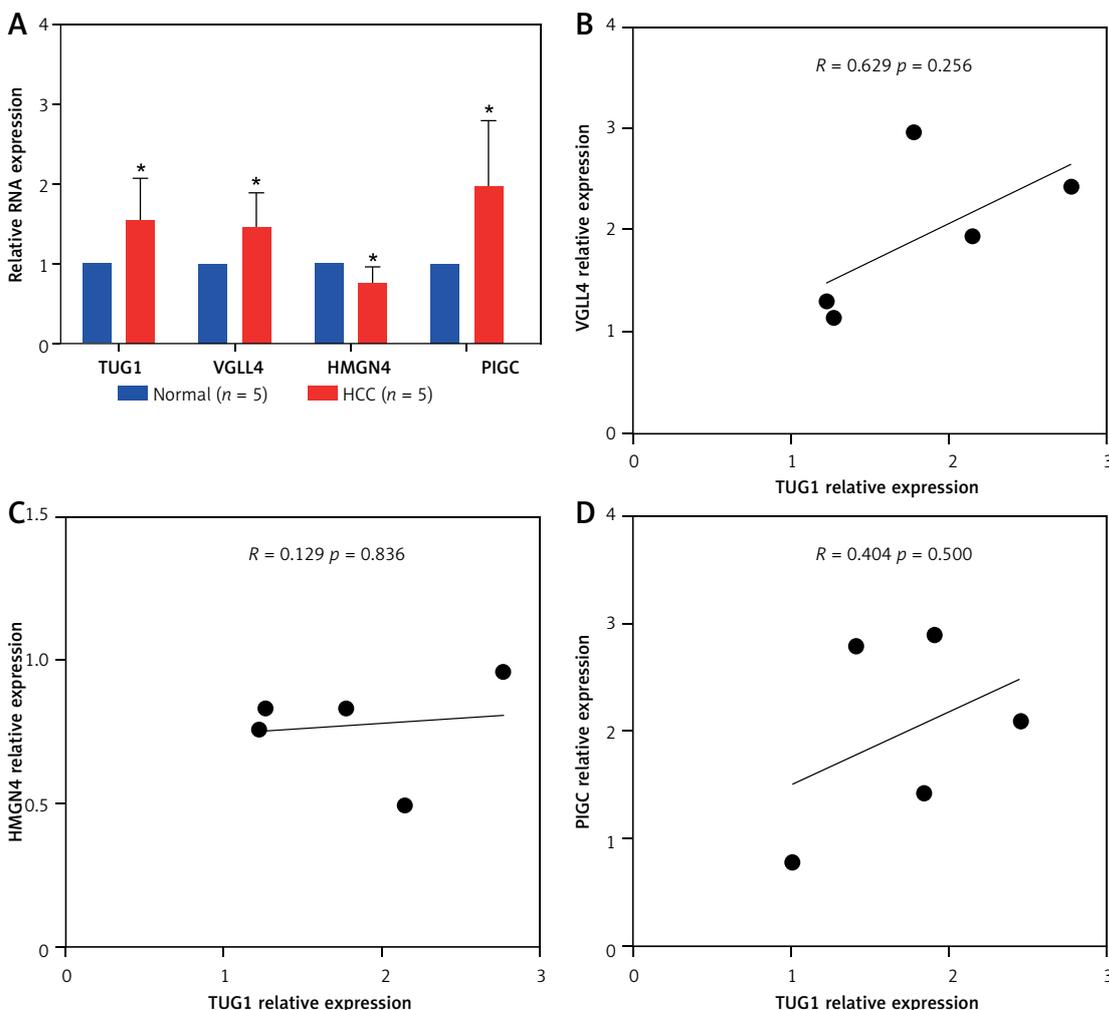


Figure 6. Expression and correlation of TUG1 and its target genes in HCC. **A** – Expression levels of TUG1, VGLL4, HMGN4, and PIGC in 5 HCC tissues and 5 matched normal tissues, * $p < 0.05$. **B** – The correlation between TUG1 and VGLL4 in 5 HCC tissues, $R = 0.629$; $p = 0.265$. **C** – The correlation between TUG1 and HMGN4 in 5 HCC tissues, $R = 0.129$; $p = 0.836$. **D** – The correlation between TUG1 and PIGC in 5 HCC tissues, relative expression

Table II. Single-variable survival analysis of clinical features of HCC

Variable	Category	N	Survival time [month]	P-value	Five-year survival rate (%)	P-value
Gender	Male	81	65.40 ±5.30	0.276	3.70	< 0.05
	Female	51	55.54 ±4.98		13.73	
Age	≤ 60	54	69.27 ±5.75	0.289	5.56	> 0.5
	> 60	76	57.09 ±3.60		9.21	
Race	White	77	55.86 ±4.19	0.278	11.69	< 0.01
	Asian	44	64.84 ±3.44		0.00	
	Black or African American	6	44.66 ±9.77		14.29	
BMI	Underweight (≤ 18.5)	6	77.35 ±13.08	0.577	16.67	> 0.5
	Normal (18.5–24.99)	55	59.12 ±4.72		5.45	
	Overweight (25–29.99)	29	67.22 ±7.44		6.90	
	Obese (≥ 30)	24	58.90 ±5.28		12.50	
Tumour stage	Stage I	44	73.05 ±5.33	< 0.001	9.09	> 0.5
	Stage II	26	68.58 ±6.98		7.69	
	Stage III	45	44.19 ±5.73		6.67	
	Stage IV	3	19.02 ±6.82		0.00	

survival time ($p < 0.05$, $n = 130$, Figures 7 A–M). We also found an interesting fact: that sex and race were significantly associated with the 5-year survival rate of HCC, revealing that the females had higher survival rates than the males and the Asians tended to have the worst 5-year survival rates ($p < 0.05$, Table II). In order to analyse the correlation between the target genes and tumour stages, we filtered 171 patients with stage I, 85 with stage II, 85 with stage III, and five with stage IV of HCC for Pearson's correlation analysis, and found that the expression of *SWAP70* had no significant correlation with tumour stage, while the expression of TUG1 and its other 11 target genes were positively correlated with tumour stages ($p < 0.05$, Figure 8), which suggested that TUG1 and its target genes could not only participate in tumour development but also promote the progression of malignancy, thus affecting the survival time of patients with liver cancer.

In addition, we performed ROC curve analysis to clarify the clinical diagnostic value of TUG1 and its target genes in HCC from the TCGA cohort. As shown in Figures 9 A–D, we found that the values of area under the curve (AUC) were higher than that of AFP (AUC = 0.663), among which AUC of *PIGC* (AUC = 0.931), *PEA15* (AUC = 0.931), *RACGAP1* (AUC = 0.961), *MCM6* (AUC = 0.920), and *NCAPG* (AUC = 0.967) were all higher than 0.9, suggesting their important roles in diagnosis of HCC.

Discussion

Liver cancer is a common malignancy of the digestive system. The annual incidence of new liver

Table III. Univariate COX regression analysis of TUG1 and its target genes

Variable	Overall Survival		Cut-off value (log2)
	HR	P-value	
TUG1	1.843	0.001	11.711
VGLL4	2.646	< 0.001	11.209
HMG4	2.111	< 0.001	10.895
PEA15	2.213	< 0.001	11.860
ENAH	1.901	< 0.001	11.892
SWAP70	1.603	0.008	10.097
MCM6	2.463	< 0.001	10.344
PIGC	1.853	0.001	10.313
NCAPG	2.188	< 0.001	9.074
RACGAP1	2.039	< 0.001	8.707
HMGB2	1.958	< 0.001	11.076
NAP1L1	2.709	< 0.001	13.273
GLS	1.595	0.021	9.152

cancer cases is about 748,000 cases worldwide, which has always been a serious threat to public health [10]. Although the medical diagnosis and treatment technology has made significant progress, its incidence is still on the rise. The prognosis of HCC patients is not satisfactory, with a 5-year survival rate of less than 20% and 5-year recurrence and metastasis rate as high as 80% [11–14]. The invasion of HCC is the main cause of the lower

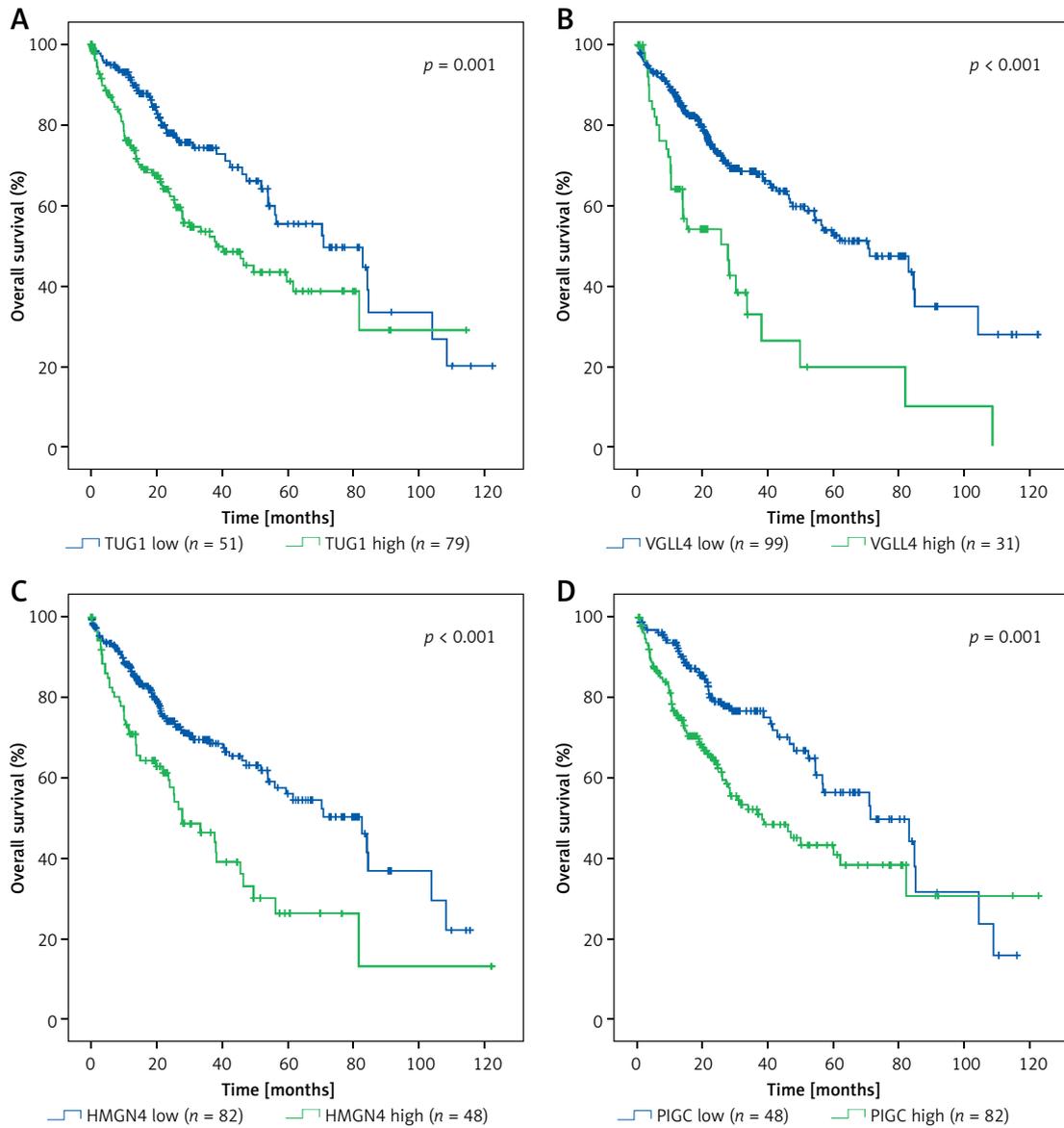


Figure 7. Kaplan–Meier analysis of TUG1 and its target genes in TCGA transcriptional sequencing dataset of HCC. The high expression of TUG1 and its 12 target genes could significantly predict shorter survival time of HCC patients. **A** – TUG1, $p = 0.001$. **B** – VGLL4, $p < 0.001$. **C** – HMGN4, $p < 0.001$. **D** – PIGC, $p = 0.001$

five-year survival rate and higher recurrence rate of metastasis. In recent years, there have been a lot of studies on liver cancer, but so far, no highly sensitive and reliable biomarkers have been found for the early diagnosis and prognosis evaluation, nor has it sufficiently clarified the occurrence, growth, and invasion of liver cancer [15]. Therefore, using combined bioinformatics will help us to better understand the relationship between aberrant lncRNA expression and cancers, and more efforts are needed to identify more efficient and sensitive molecular targets for HCC.

In our study, we identified the most significantly differentially expressed lncRNA TUG1 in the GSE14323 dataset. Twelve highly credible target genes of TUG1, screened out based on genet-

ic database and molecular interaction database, were used to construct a TUG1 regulatory network. Seven of these genes had been reported to have been associated with cancer previously. For example, RACGAP1 is an important regulatory factor of cytokinesis [16], involved in the growth and differentiation of haematopoietic cells. Loss of its expression correlates with loss of migration and invasiveness of HCC cells [17]. Patients with RACGAP1 protein expression at the invasive front had significantly poorer prognosis than those without it; RACGAP1 is increasingly recognised as a potential biomarker for identifying patients with lymph node metastasis recurrence and poor prognosis [18, 19]. Researchers have proved that MCM6 is a tumour suppressor affecting the

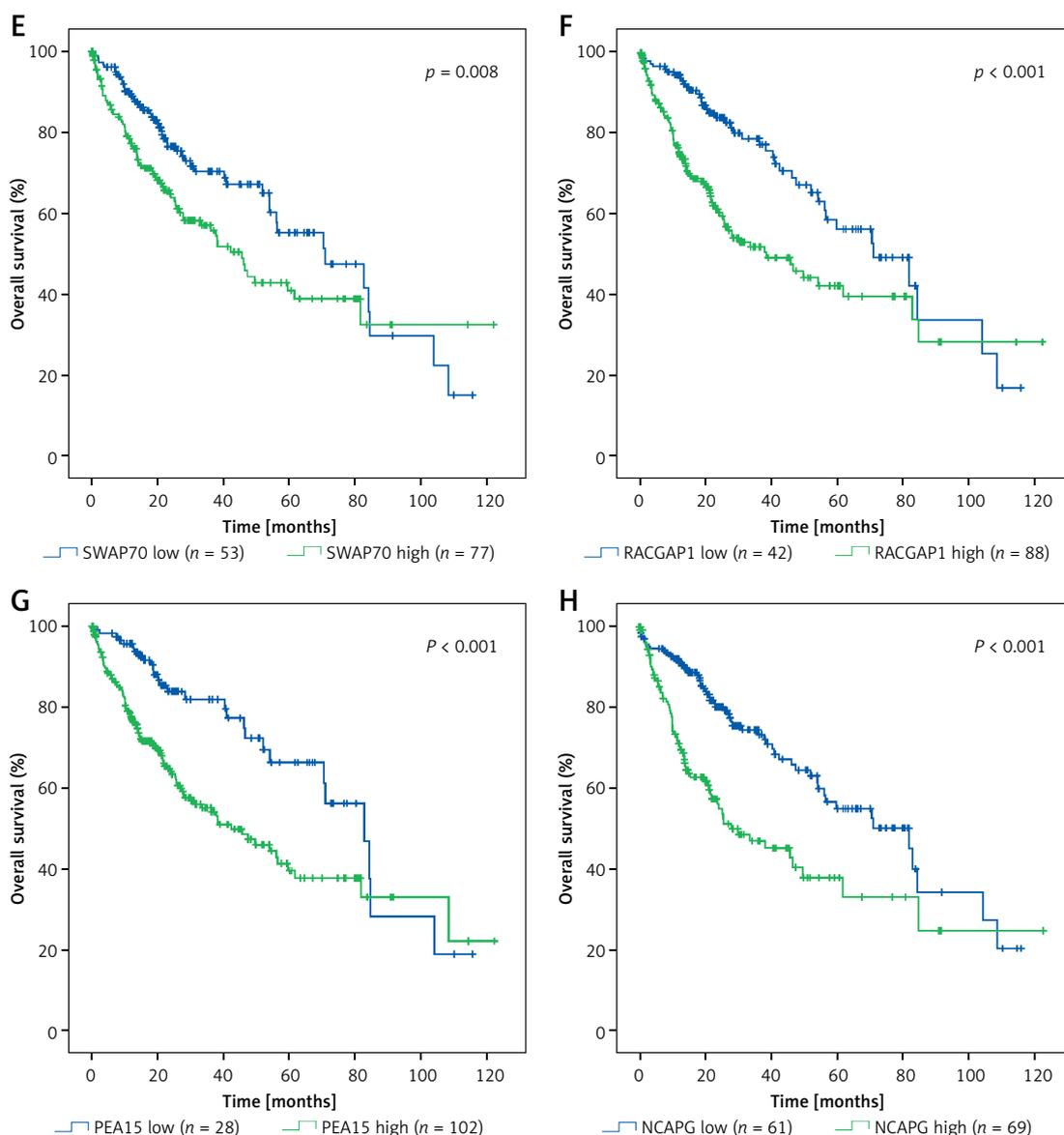


Figure 7. Cont. E – SWAP70, $p = 0.008$. F – RACGAP1, $p < 0.001$. G – PEA15, $p < 0.001$. H – NCAPG, $p < 0.001$.

growth and metastasis of HCC. Additionally, it was reported that HMGB2 is regarded as a new marker for the diagnosis and recurrence of HCC [20–22], which is involved in cell migration, localisation, and immunoreactions. High expression of HMGB2 could promote the proliferation and inhibit the apoptosis of HCC cells, as well as predicting poor prognosis [23]. Researchers have demonstrated that NCAPG is associated with the growth of hepatoma cells, apoptosis, and epithelial mesenchymal transformation (EMT) [24]; NAP1L1 was found to be a nucleosome assembly protein in response to nucleosome remodelling through CSB protein and chromatin remodelling [25], promoting tumour progression as a chromatin regulatory factor in liver cancer [26]; PEA15, an essential prognostic marker for HCC, is involved in negative regu-

lation of apoptosis signalling pathway mediated by death receptor, thereby inhibiting apoptosis of hepatoma cells. In addition, PEA15 also participates in transmembrane receptor protein tyrosine kinase signalling pathway, which mediates cell communication and promotes the migration of liver cancer cells, and the high level of ENAH expression promotes the occurrence and metastasis of liver cancer [27–29].

Based on the bioinformatics database, hepatoma cell line, and liver cancer tissue, we detected and analysed the expression and correlation of TUG1 and its target genes, and obtained four highly consistent TUG1 target genes (*HMGB2*, *NCAPG*, *MCM6*, and *RACGAP1*). We found that the expression levels of TUG1 and its target gene *VGLL4* in HCC cell lines were higher than

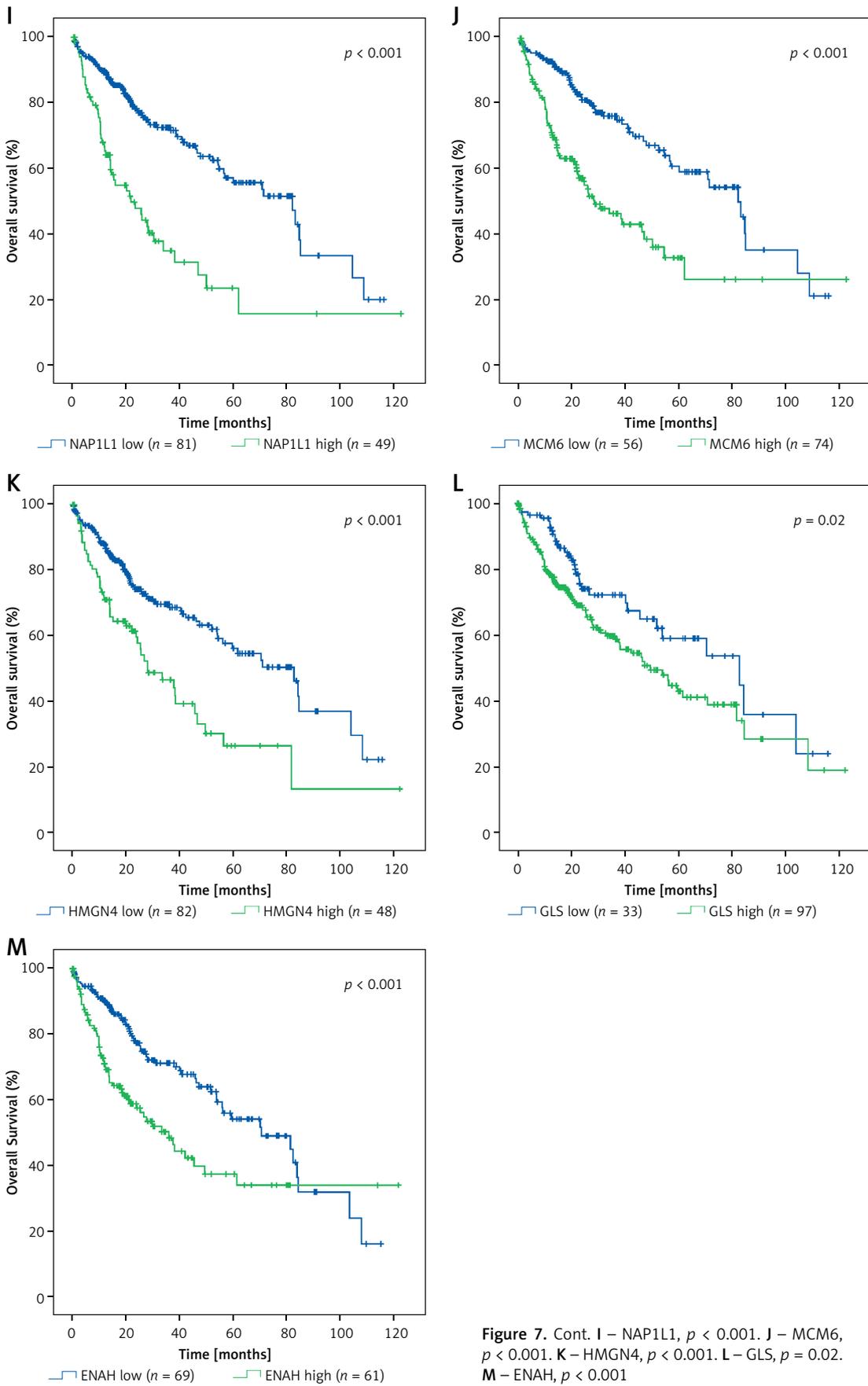


Figure 7. Cont. I – NAP1L1, $p < 0.001$. J – MCM6, $p < 0.001$. K – HMGN4, $p < 0.001$. L – GLS, $p = 0.02$. M – ENAH, $p < 0.001$

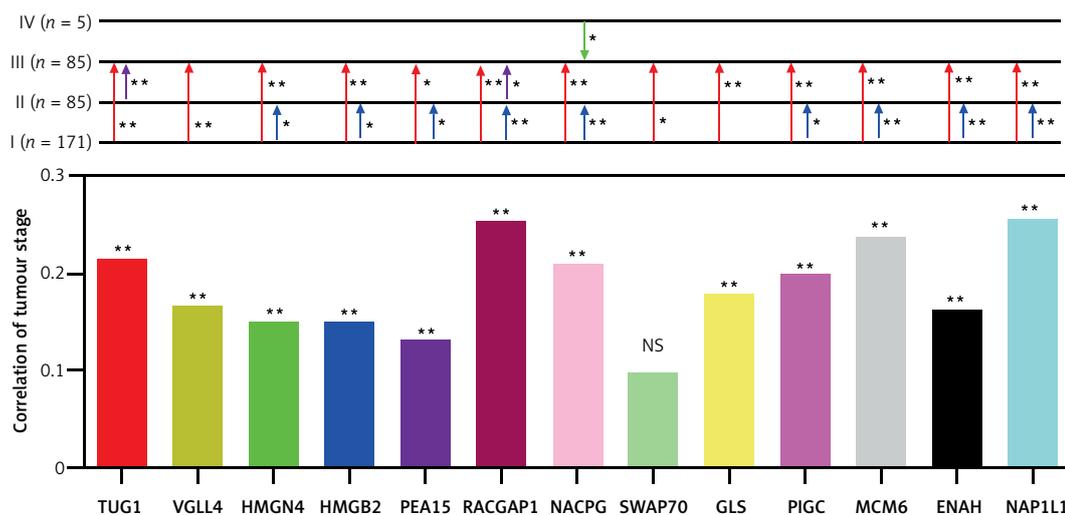


Figure 8. Differential expression and correlation of TUG1 and its target genes among the different stages of HCC. Red arrowhead: Up-regulation of the indicating genes expression from stage I to stage II; purple arrowhead: up-regulation of the indicating genes expression from stage II to stage III; blue arrowhead: up-regulation of the indicating gene expression from stage III to stage IV; green arrowhead: down-regulation of the indicating gene expression from stage IV to stage III; * $p < 0.05$ and ** $p < 0.01$; NS – no significance

those in normal hepatoma cell line, which was consistent with the results predicted by bioinformatics method, but the expressions of PIGC and HMGN4 in different hepatoma cell lines were not consistent. These results suggest that the expression levels of these two genes are related to the genetic background and malignancy of different hepatoma cell lines. In addition, after interfering with TUG1, VGLL4, PIGC, and HMGN4 were all down-regulated, but the percentages of down-regulated expression were relatively low, indicating that TUG1 had a weak regulatory effect on VGLL4, PIGC, and HMGN4, and there were other molecules involved in their expression regulation. Next, it was found that the RNA levels of TUG1 and its target genes VGLL4, PIGC, PEA15, and ENAH were significantly increased in HCC tissues ($p < 0.05$), while the expression of VGLL4 protein was significantly higher in HCC tissues ($p < 0.05$), indicating that the target genes regulated by TUG1 played a role in the promotion of HCC, and PEA15 and ENAH were regulated by TUG1 to some extent. However, the expression of HMGN4 was inconsistent with our previous prediction, indicating the high heterogeneity of HMGN4 in different HCC patients, which needs to be further investigated by increasing the sample size. Through the correlation analysis, we found that the expression of TUG1 and its target genes VGLL4, PEA15, and ENAH were highly correlated in HCC tissues, but the correlations between TUG1 and PIGC and HMGN4 were weak.

Finally, we studied the diagnostic and prognostic values of TUG1 and its target genes based on TCGA database. The analysis of transcriptome sequencing data and clinical data from the TCGA

database showed that the high expression of TUG1 and its 12 target genes could significantly predict the shorter survival time. Results obtained from this study are consistent with other studies; however, existing studies have only reported that NCAPG, HMGB2, and ENAH have significant prognostic effects. For instance, the knockdown of NCAPG expression reduced HCC cell viability but induced apoptosis and cell cycle arrest at S phase, and suppressed tumour cell EMT. Moreover, high levels of NCAPG expression are associated with the clinicopathological characteristics, recurrence, time of recurrence, metastasis, differentiation, and TNM stage in human HCC. Therefore, NCAPG can be a potential target for the diagnosis and therapy of patients with HCC [23, 24, 27]. We found that TUG1 target genes NCAPG, VGLL4, HMGN4, PIGC, MCM6, NAP1L1, PEA15, GLS, HMGB2, PEA15, and ENAH were positively correlated with tumour stages ($p < 0.05$), suggesting that these target genes not only play a role in the development of liver cancer but also participate in tumour progression. For example, MCM6, PEA15, RACGAP1, and ENAH have been reported to promote the metastasis of HCC [17, 20–30]. The diagnostic effect analysis shows that TUG1 and its target genes have higher diagnostic efficacy than the traditional marker AFP in clinics, so further clinical trials are needed.

In the present study, we identified differentially expressed lncRNA–TUG1 and its target genes (VGLL4, HMGN4, PIGC, MCM6, PEA15, RACGAP1, GLS, NAP1L1, HMGB2, NCAPG, ENAH, and SWAP70) by using bioinformatics methods.

The high expression of TUG1 and its target genes can significantly predict shorter survival

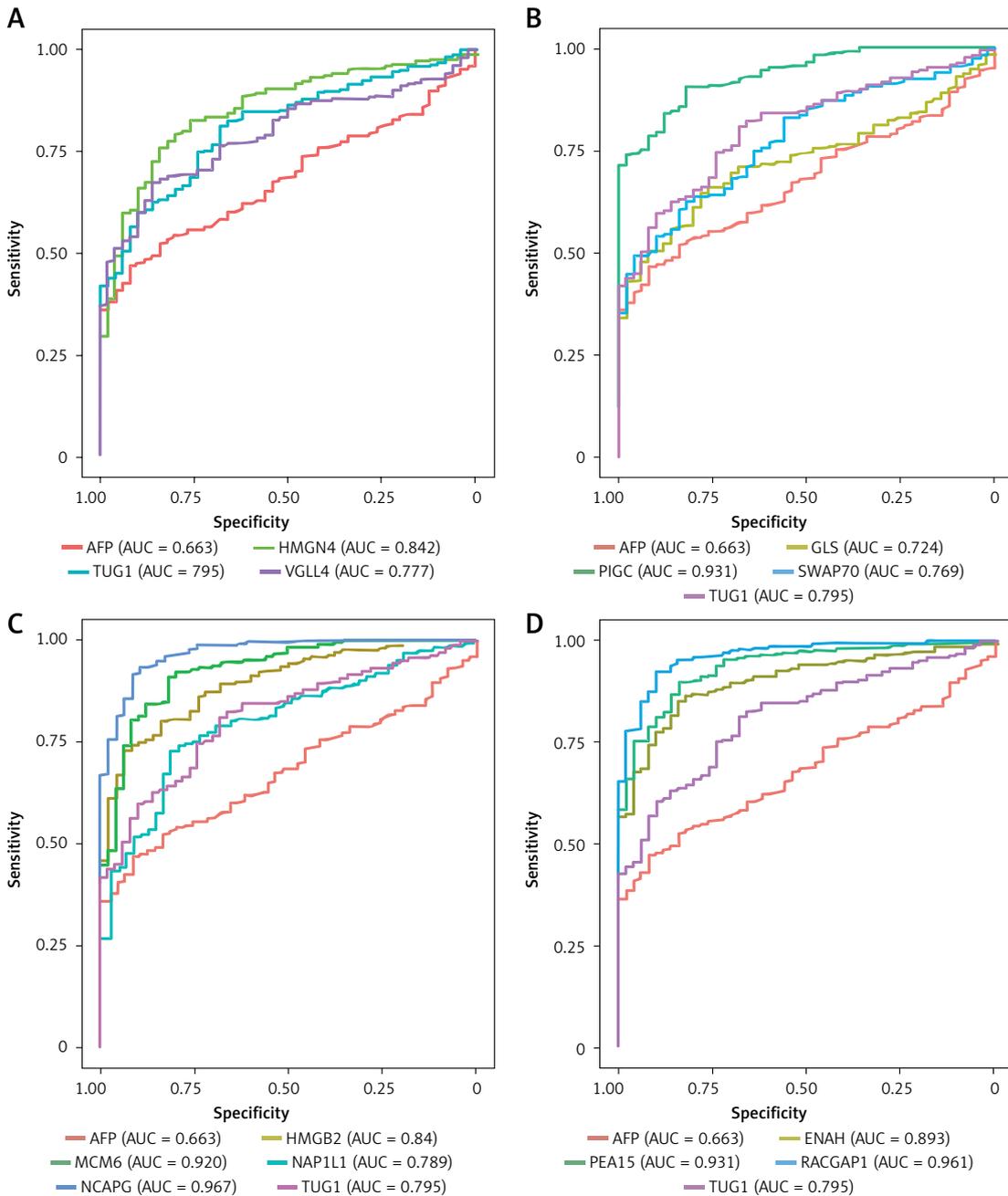


Figure 9. Diagnostic values of TUG1 and its target genes in HCC. **A** – Diagnostic efficacy of AFP (AUC = 0.663), TUG1 (AUC = 0.795), HMGN4 (AUC = 0.842) and VGLL4 (AUC = 0.777) in HCC. **B** – Diagnostic efficacy of GLS (AUC = 0.724), PIGC (AUC = 0.931), SWAP70 (AUC = 0.769) in HCC. **C** – Diagnostic efficacy of HMGB2 (AUC = 0.84), MCM6 (AUC = 0.92), NAP1L1 (AUC = 0.789) and NCAPG (AUC = 0.967) in HCC. **D** – Diagnostic efficacy of ENAH (AUC = 0.893), PEA15 (AUC = 0.931) and RACGAP1 (AUC = 0.961) in HCC

time of HCC patients and can be used as potentially effective prognostic biomarkers. NCAPG, MCM6, PIGC, PEA15, and RACGAP1 could be useful in the diagnosis of HCC.

Acknowledgments

This study was supported in part by the Fundamental Research Funds for the Central Universities of Central South University (Program number: 506021714) and Hunan Natural Science Fund (Program number: 2018 jj2554).

Conflict of interest

The authors declare no conflict of interest.

References

1. Peng L, Yuan XQ, Zhang CY, et al. High TGF-beta1 expression predicts poor disease prognosis in hepatocellular carcinoma patients. *Oncotarget* 2017; 8: 34387-97.
2. Liu H, Liu J, Zhang Y, et al. Contrast-enhanced ultrasound and computerized tomography perfusion imaging of a liver fibrosis-early cirrhosis in dogs. *J Gastroenterol Hepatol* 2016; 31: 1604-10.

3. Feng W, Li L, Xu X, Jiao Y, Du W. Up-regulation of the long non-coding RNA RMRP contributes to glioma progression and promotes glioma cell proliferation and invasion. *Arch Med Sci* 2017; 13: 1315-21.
4. Zhang C, Peng L, Zhang Y, et al. The identification of key genes and pathways in hepatocellular carcinoma by bioinformatics analysis of high-throughput data. *Med Oncol* 2017; 34: 101.
5. Gu LP, Jin S, Xu RC, et al. Long non-coding RNA PCAT-1 promotes tumor progression by inhibiting miR-129-5p in human ovarian cancer. *Arch Med Sci* 2019; 15: 513-21.
6. Kamel MM, Matboli M, Sallam M, Montasser IF, Saad AS, El-Tawdi AHF. Investigation of long noncoding RNAs expression profile as potential serum biomarkers in patients with hepatocellular carcinoma. *Transl Res* 2016; 168: 134-45.
7. Peng L, Yuan XQ, Zhang CY, et al. The emergence of long non-coding RNAs in hepatocellular carcinoma: an update. *J Cancer* 2018; 9: 2549-58.
8. Kudo M. Imaging diagnosis of hepatocellular carcinoma and premalignant/borderline lesions. *Semin Liver Dis* 1999; 19: 297-309.
9. Torzilli G, Minagawa M, Takayama T, et al. Accurate preoperative evaluation of liver mass lesions without fine-needle biopsy. *Hepatology* 1999; 30: 889-93.
10. McGuire S. World Cancer Report 2014. Geneva, Switzerland: World Health Organization, International Agency for Research on Cancer, WHO Press, 2015. *Adv Nutr* 2016; 7: 418-9.
11. Lee YY, McKinney KQ, Ghosh S, et al. Subcellular tissue proteomics of hepatocellular carcinoma for molecular signature discovery. *J Proteome Res* 2011; 10: 5070-83.
12. Ng KM, Yan TD, Black D, Chu FC, Morris DL. Prognostic determinants for survival after resection/ablation of a large hepatocellular carcinoma. *HPB (Oxford)* 2009; 11: 311-20.
13. Iizuka N, Oka M, Yamada-Okabe H, et al. Oligonucleotide microarray for prediction of early intrahepatic recurrence of hepatocellular carcinoma after curative resection. *Lancet* 2003; 361: 923-9.
14. Zhou XD, Tang ZY, Yang BH, et al. Experience of 1000 patients who underwent hepatectomy for small hepatocellular carcinoma. *Cancer* 2001; 91: 1479-86.
15. Xing X, Liang D, Huang Y, et al. The application of proteomics in different aspects of hepatocellular carcinoma research. *J Proteomics* 2016; 145: 70-80.
16. Zhao WM, Fang G. MgcRacGAP controls the assembly of the contractile ring and the initiation of cytokinesis. *Proc Nat Acad Sci USA* 2005; 102: 13158-63.
17. Wang SM, Ooi LL, Hui KM. Upregulation of Rac GTPase-activating protein 1 is significantly associated with the early recurrence of human hepatocellular carcinoma. *Clin Cancer Res* 2011; 17: 6040-51.
18. Kawashima T, Hirose K, Satoh T, et al. MgcRacGAP is involved in the control of growth and differentiation of hematopoietic cells. *Blood* 2000; 96: 2116-24.
19. Saigusa S, Tanaka K, Mohri Y, et al. Clinical significance of RacGAP1 expression at the invasive front of gastric cancer. *Gastric Cancer* 2015; 18: 84-92.
20. Liu M, Hu Q, Tu M, et al. MCM6 promotes metastasis of hepatocellular carcinoma via MEK/ERK pathway and serves as a novel serum biomarker for early recurrence. *J Exp Clin Cancer Res* 2018; 37: 10.
21. Liu Z, Li J, Chen J, et al. MCM family in HCC: MCM6 indicates adverse tumor features and poor outcomes and promotes S/G2 cell cycle progression. *BMC Cancer* 2018; 18: 200.
22. Zheng T, Chen M, Han S, et al. Plasma minichromosome maintenance complex component 6 is a novel biomarker for hepatocellular carcinoma patients. *Hepatol Res* 2014; 44: 1347-56.
23. Kwon JH, Kim J, Park JY, et al. Overexpression of high-mobility group box 2 is associated with tumor aggressiveness and prognosis of hepatocellular carcinoma. *Clin Cancer Res* 2010; 16: 5511-21.
24. Liu W, Liang B, Liu H, et al. Overexpression of nonSMC condensin I complex subunit G serves as a promising prognostic marker and therapeutic target for hepatocellular carcinoma. *Int J Mol Med* 2017; 40: 731-8.
25. Lee JY, Lake RJ, Kirk J, Bohr VA, Fan HY, Hohng S. NAP1L1 accelerates activation and decreases pausing to enhance nucleosome remodeling by CSB. *Nucleic Acids Res* 2017; 45: 4696-707.
26. Zhou X, Yang F, Yang Y, et al. HBV facilitated hepatocellular carcinoma cells proliferation by up-regulating angiogenin expression through IL-6. *Cell Physiol Biochem* 2018; 46: 461-70.
27. Hu K, Wang J, Yao Z, et al. Expression of cytoskeleton regulatory protein Mena in human hepatocellular carcinoma and its prognostic significance. *Med Oncol* 2014; 31: 939.
28. Hu K, Huang P, Luo H, et al. Mammalian-enabled (MENA) protein enhances oncogenic potential and cancer stem cell-like phenotype in hepatocellular carcinoma cells. *FEBS Open Bio* 2017; 7: 1144-53.
29. Lin L, Yang XM, Li J, Zhang YL, Qin W, Zhang ZG. Microfilament regulatory protein MENA increases activity of RhoA and promotes metastasis of hepatocellular carcinoma. *Exp Cell Res* 2014; 327: 113-22.
30. Quintavalle C, Hindupur SK, Quagliata L, et al. Phosphoprotein enriched in diabetes (PED/PEA15) promotes migration in hepatocellular carcinoma and confers resistance to sorafenib. *Cell Death Dis* 2017; 8: e3138.