

# Comprehensive analysis of key lncRNAs in HCV-positive hepatocellular carcinoma

Jingqi Liu<sup>1</sup>, Ligang Chen<sup>2</sup>, Jinshui Pan<sup>2</sup>, Meiya Chen<sup>2</sup>, Jingping Zhou<sup>2</sup>, Fei Zhou<sup>2</sup>, Peizhong Chen<sup>2</sup>, Yang Song<sup>2</sup>

<sup>1</sup>Department of Geriatrics, Zhongshan Hospital affiliated to Xiamen University, Xiamen, Fujian, China

<sup>2</sup>Department of Gastroenterology, Zhongshan Hospital affiliated to Xiamen University, Xiamen, Fujian, China

**Submitted:** 15 August 2017

**Accepted:** 15 January 2018

Arch Med Sci 2021; 17 (1): 142–151

DOI: <https://doi.org/10.5114/aoms.2020.100675>

Copyright © 2020 Termedia & Banach

**Corresponding author:**

Yang Song

Department of

Gastroenterology

Zhongshan Hospital affiliated

to Xiamen University

361004, Xiamen, China

Phone: +8618559666260

E-mail: [tcsongyang@sina.com](mailto:tcsongyang@sina.com)

## Abstract

**Introduction:** Hepatocellular carcinoma (HCC) is one of the most common cancers worldwide. Despite the therapeutic advances in HCC in the past few decades, the mortality rate of HCC is still high. Hepatitis C (HCV) infection is one of the major etiological risk factors of HCCs. However, the underlying mechanisms of HCV-induced hepatocarcinogenesis remain largely unclear.

**Material and methods:** Our study represented the comprehensive analysis of differentially expressed lncRNAs in HCV-positive HCC for the first time by analyzing the public dataset GSE17856. Co-expression network and gene ontology (GO) analysis revealed the functions of those differentially expressed lncRNAs.

**Results:** We identified 256 upregulated lncRNAs and 198 downregulated lncRNAs in HCV-positive HCC compared to the normal liver tissues. Co-expression network and GO analysis showed that these lncRNAs were involved in regulating metabolism, energy pathways, proliferation and the immune response. Seven lncRNAs (LOC341056, CCT6P1, PTTG3P, LOC643387, LOC100133920, C3P1 and C22orf45) were identified as key lncRNAs and co-expressed with more than 100 differentially expressed genes (DEGs) in HCV-related HCC. Kaplan-Meier analysis showed that higher expression levels of LOC643387, PTTG3P, LOC341056, CCT6P1 and lower expression levels of C3P1 and C22orf45 were associated with shorter survival time in the TCGA dataset.

**Conclusions:** We believe that this study can provide novel potential therapeutic and prognostic biomarkers for HCV-positive HCC.

**Key words:** Long non-coding RNA, HCV, hepatocellular carcinoma, co-expression network, biomarkers.

## Introduction

Hepatocellular carcinoma (HCC) is one of the most common cancers and the second leading cause of cancer death worldwide [1]. Despite the therapeutic advances in HCC in the past few decades, the mortality rate of HCC is still high [2]. Chronic hepatitis B (HBV) and hepatitis C (HCV) infection were the major etiological risk factors of HCCs [3]. HCV is a sense-strand RNA virus and is unable to integrate into the host genome [4] and HCV infection represents the main risk factor of primary HCC [5]. Although a previous study showed that several proteins (such as COX-2)

were associated with decreased overall and disease-free survival [6], and HCV encoded proteins (such as E1, E2, P7, NS2 and NS3) could also promote HCC formation by activating oncogenic molecular pathways *in vivo* [7–9], the underlying mechanisms of HCV-induced hepatocarcinogenesis remain largely unclear.

lncRNAs are a class of non-coding RNAs longer than 200 nucleotides with little or no protein-coding potential [10]. Previous studies had reported that lncRNA expression was frequently dysregulated and could contribute to the progress of various human cancer types, including prostate [11], breast [12], colon [13] and liver cancers [14]. For example, Fan *et al.* revealed that more than 1000 lncRNA transcripts were significantly differentially expressed in HBV-related HCC patients [15]. Mechanically, lncRNAs could regulate protein-coding genes' expression at epigenetic, transcriptional, and post-translational levels [16]. One of the most well-known mechanisms was acting as miRNAs sponges by interacting with miRNAs [17]. However, the potential roles of an enormous number of lncRNAs in HCC, especially in HCV-positive HCC, remains to be further elucidated.

In this study, we analyzed the public dataset GSE17856 to identify differentially expressed lncRNAs (DElncs) and mRNAs in HCV-related HCC patients. Next, a series of bioinformatics analyses (including GO, KEGG and co-expression analysis) were performed to explore potential roles of DElncs. The present study aimed to provide useful information to identify novel lncRNAs as biomarkers for HCV-related HCC.

## Material and methods

### Microarray data

Microarray data were obtained from the study by Tsuchiya *et al.* [18], which is referenced in the Gene Expression Omnibus (GEO) database ([www.ncbi.nlm.nih.gov/geo/](http://www.ncbi.nlm.nih.gov/geo/)) under accession number GSE17856. The platform is GPL6480 Agilent-014850 Whole Human Genome Microarray 4x44K G4112F Array. A total of 44 non-tumoral liver and 43 tumors sample were obtained from 47 subjects with HCV-associated HCC. lncRNAs having fold changes  $\geq 2$  and  $p$ -values  $< 0.05$  were selected as of significantly differential expression.

### lncRNA classification pipeline

To evaluate the lncRNA expression pattern in microarray data, we applied a pipeline described by Zhang *et al.* [19]. Briefly, the probe set ID was mapped to the NetAffx Annotation Files (HG-U133 Plus 2.0, Annotations, CSV format, release 31, 08/23/10). The annotations included the probe set ID, gene symbol, gene title, ensemble gene ID, Ref-

seq transcript ID and other informative items for the specific probe sets. The probe ID-centric gene expression data were joined with the annotation files on the probe set ID. Secondly, the probe sets which were assigned with a Refseq transcript ID and/or Ensemble gene ID in the, NetAffx annotations were extracted. For the probe sets with Refseq IDs, we only retained those labeled as “NR\_” (NR indicates non-coding, RNA in the Refseq database). For the probe sets with Ensemble gene IDs, we only retained those annotated with “lncRNA”, “processed, transcripts”, “non-coding” or “misc\_RNA” in Ensemble annotations (accessible at the UCSC genome browser: <http://www.genome.ucsc.edu/>). Then, we filtered the probe sets obtained in step 2 by filtering out rRNAs, microRNAs and other short RNAs, including tRNAs, snRNAs and snoRNAs. Finally, 2448 annotated lncRNA transcripts with corresponding Affymetrix probe IDs were generated.

### Identification of DElncs

A  $t$ -test [20] in the limma [21] package, in R [22] was used to identify genes which were differentially expressed between normal and HCV-positive HCC samples. The threshold for the DEGs was set as corrected  $p$ -value  $< 0.05$  and  $|\log_2$  fold change, (FC)  $|\geq 1$ .

### GO and KEGG pathway analysis

To identify functions of DEGs in HCC, we performed GO function enrichment analysis in three functional ontologies: biological process (BP), cellular component (CC) and molecular function (MF). Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis was also performed to identify pathways enriched in HCC using the DAVID system (<https://david.ncifcrf.gov/>). The  $p$ -value was calculated by hypergeometric distribution and a pathway with  $p < 0.05$  was considered as significant.

### Co-expression network construction and analysis

In this study, the Pearson correlation coefficient of DEG-lncRNA pairs was calculated according to the expression value of them. The co-expressed DEG-lncRNA pairs with the absolute value of Pearson correlation coefficient  $\geq 0.5$  were selected and the co-expression network was established using Cytoscape software.

### Statistical analysis

The numerical data were presented as mean  $\pm$  standard deviation (SD) of at least three determinations. Statistical comparisons between groups of normalized data were performed using the  $t$ -test or

Mann-Whitney *U*-test according to the test condition. A value of  $p < 0.05$  was considered to indicate statistical significance with a 95% confidence level.

## Results

### Screening the DElncs in the HCV-positive hepatocellular carcinoma

In the present study, the public dataset GSE17856 was analyzed to identify differentially expressed lncRNAs. A total of 44 non-tumoral liver and 43 tumor samples were obtained from 47 subjects with HCV-associated HCC. LncRNAs having fold changes  $\geq 2$  and  $p$ -values  $< 0.05$  were selected as of significantly differential expression. In total, 454 transcripts were observed to be expressed differentially compared to the normal liver tissues, including 256 upregulated transcripts and 198 downregulated transcripts. Hierarchical clustering showed systematic variations in the expression of lncRNAs in the HCV-positive HCC samples (Figure 1).

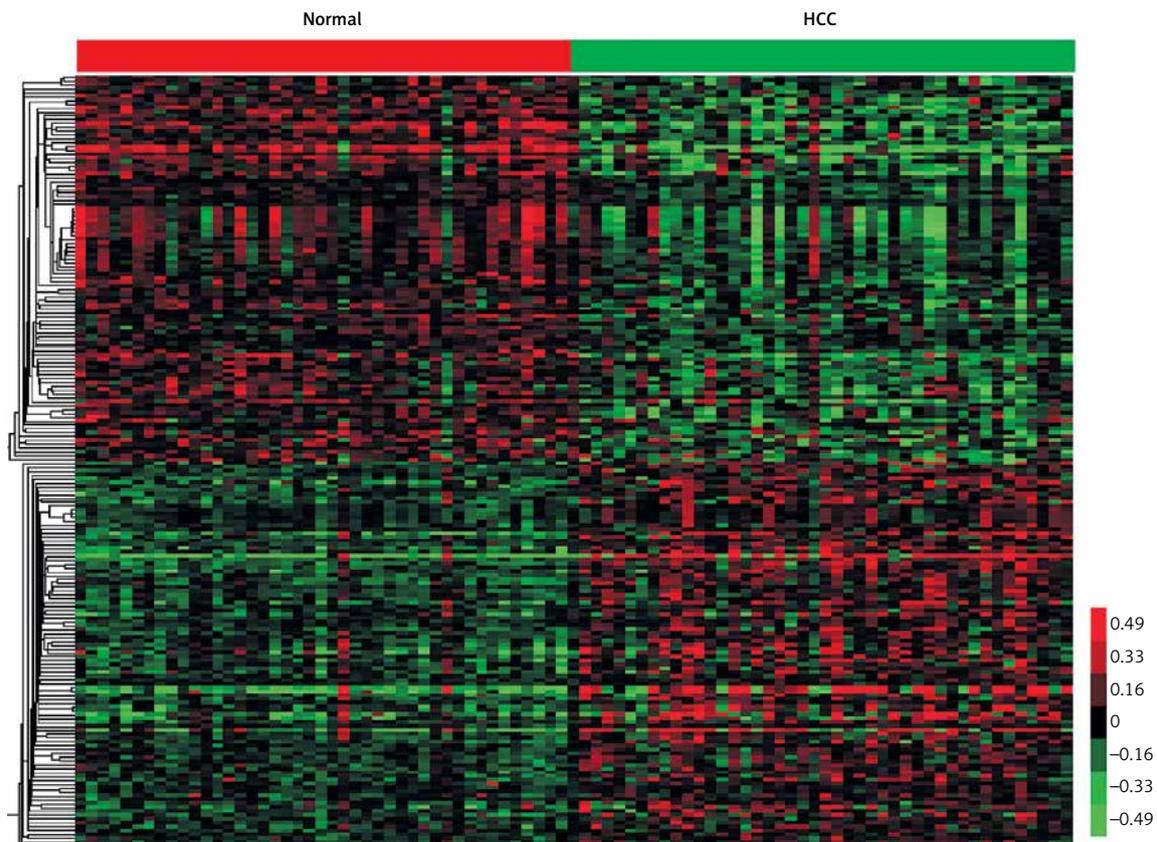
### Co-expression network analysis

Here, we constructed up- and down-regulated lncRNA-mRNA co-expressed networks to reveal

the potential roles of lncRNAs in HCV-positive HCC. The co-expressed lncRNA-mRNA pairs with the absolute value of the Pearson correlation coefficient  $\geq 0.5$  were selected for network construction. As shown in Figure 2, a total of 125 lncRNAs and 1469 DEGs were included in the up-regulated lncRNA mediated co-expression network. In this network, we found that LOC341056, CCT6P1, PTTG3P, LOC643387, CXADRP2, LOC100128098, LOC100133920 played key roles and co-expressed with more than 100 DEGs. Meanwhile, we observed that the up-regulated lncRNA mediated co-expression network contained 67 lncRNAs and 558 DEGs (Figure 2 A). Twelve lncRNAs (FLJ35390, MGC45922, LOC92249, MTMR9LP, LOC440461, LOC92659, LOC100130557, LINC00304, RRP7B, FAM201A, PYY2, PIN1P1) co-expressed with more than 100 DEGs in the down-regulated lncRNA mediated network (Figure 2 B).

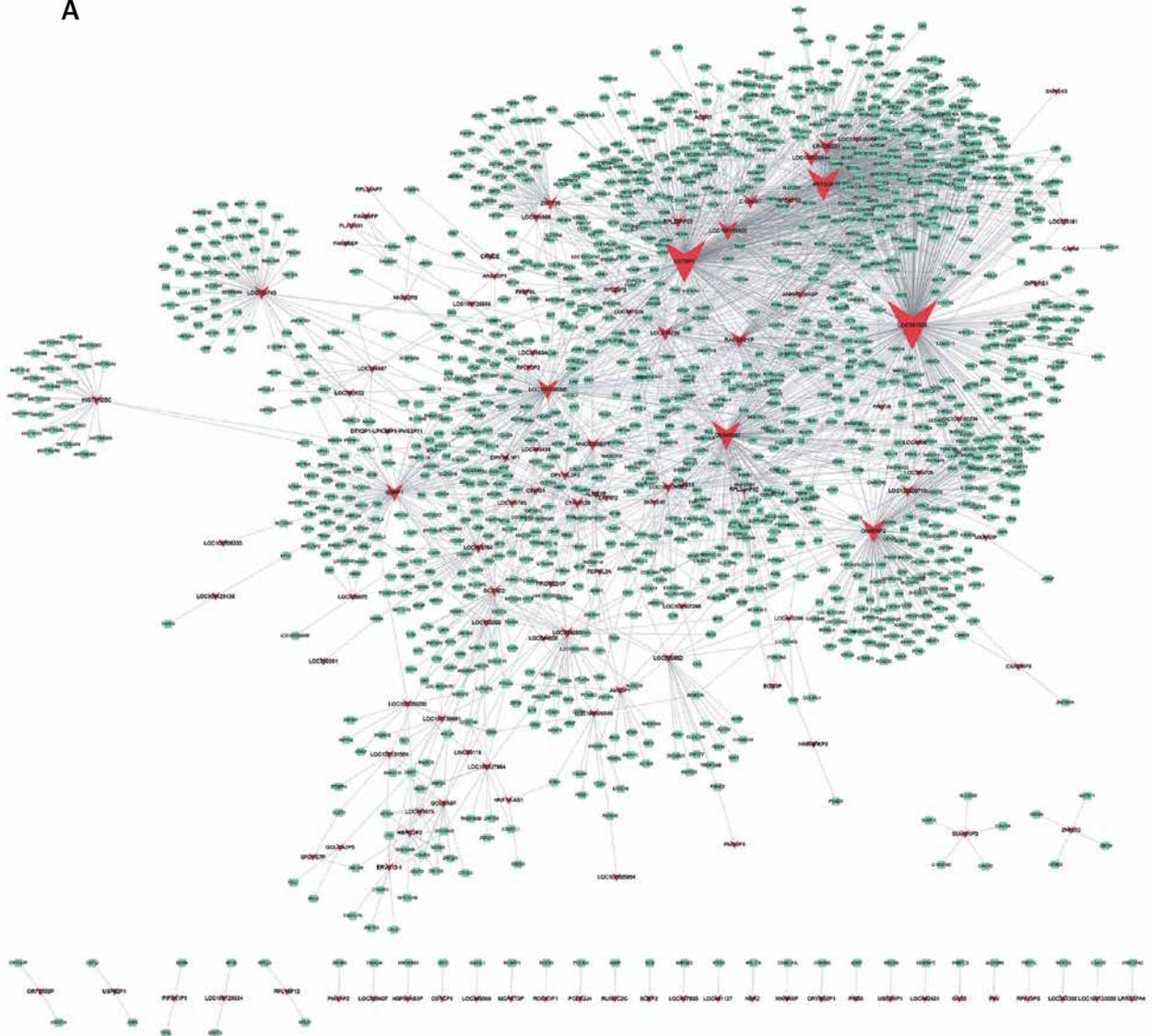
### Functional annotation of differentially expressed lncRNAs

Based on co-expression networks, we performed GO and KEGG analysis for differentially expressed lncRNAs using the set of co-expressed



**Figure 1.** Hierarchical clustering of systematic variations in the expression of lncRNAs in the HCV-positive HCC samples. We identified 256 upregulated lncRNAs and 198 downregulated lncRNAs in HCV-positive HCC compared to the normal liver tissues. Red indicates high relative expression and green indicates low relative expression. Normal represents nontumoral liver and HCC represents tumor samples which were obtained from subjects with HCV-associated HCC

A



**Figure 2.** Co-expression network of upregulated lncRNAs with DEGs (A). Triangle nodes, lncRNAs; square nodes, lncRNAs co-expressed genes

mRNAs. According to the GO analysis, we found that up-regulated lncRNAs were mainly involved in regulating metabolism, energy pathways, regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism, regulation of the cell cycle, amino acid and derivative metabolism, cell cycle, DNA repair, and DNA replication (Figure 3 A). Meanwhile, we observed that the downregulated lncRNAs were enriched in regulating metabolism, energy pathways, immune response, regulation of biological process, and cell-cell adhesion (Figure 3 B).

#### Identification of key lncRNAs in regulating HCV-positive HCC development

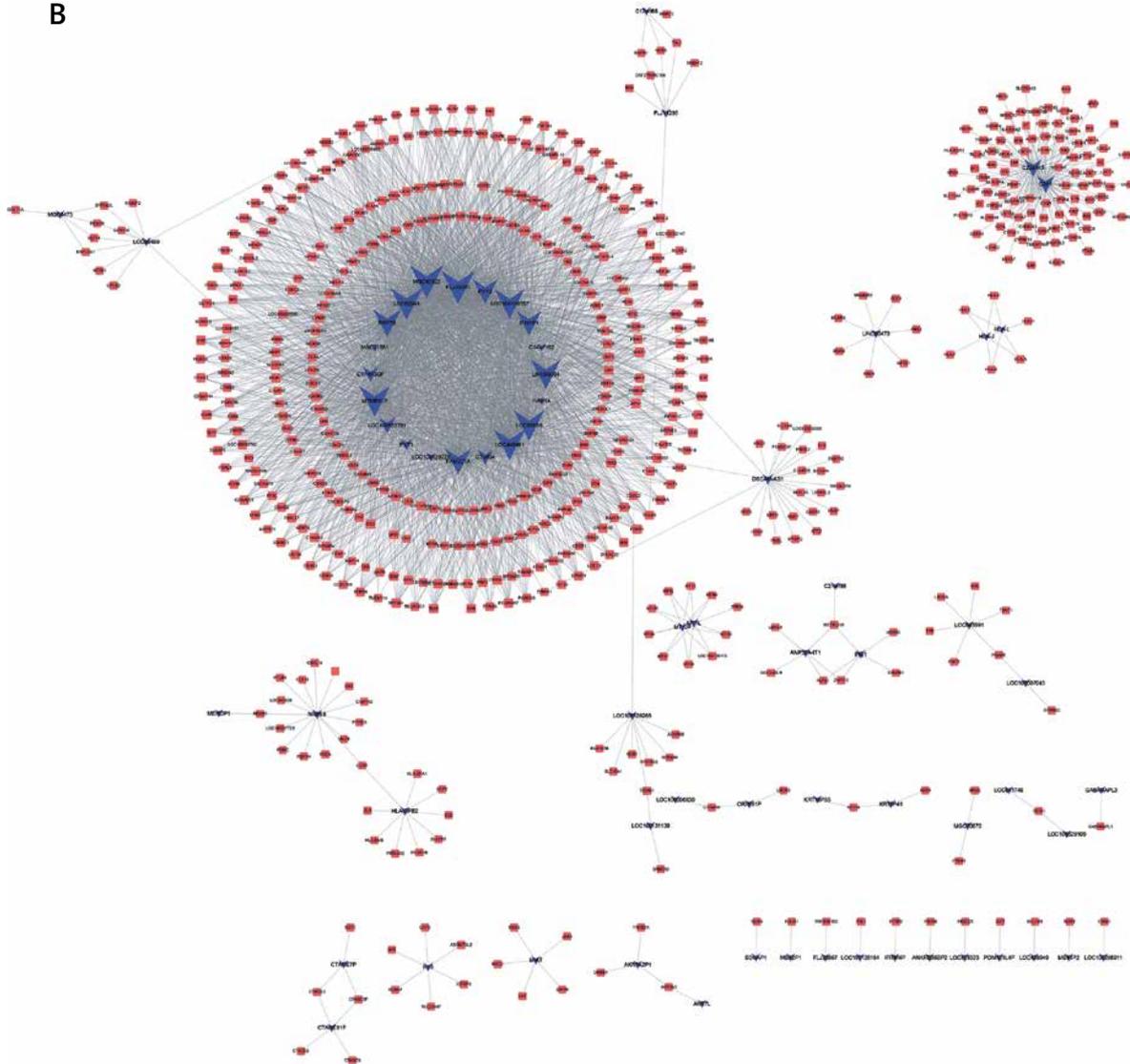
For further explore the key lncRNAs in regulating key biological processes in HCV-positive HCC development, we constructed lncRNA-mRNA-biological processes. We observed that LOC341056,

CCT6P1, PTTG3P, LOC643387, LOC100133920 were mainly involved in regulating metabolism and regulation of nucleobase and cell proliferation processes were most significant in the upregulated network (Figure 4 A). In the down-regulated network, C3P1 and C22orf45 played crucial roles in regulating immune response and energy pathways (Figure 4 B).

#### Key lncRNAs were associated with hepatocellular carcinoma prognosis

Furthermore, we evaluated the possible prognostic value of LOC341056, CCT6P1, PTTG3P, LOC643387, LOC100133920, C3P1 and C22orf45 in HCC using TCGA RNA-seq data. As shown in Figures 5 A–D, according to Kaplan-Meier analysis, we found that higher expression levels of LOC643387, PTTG3P, LOC341056, and CCT6P1

**B**



**Figure 2.** Cont. Co-expression network of downregulated lncRNAs with DEGs (**B**). Triangle nodes, lncRNAs; square nodes, lncRNAs co-expressed genes

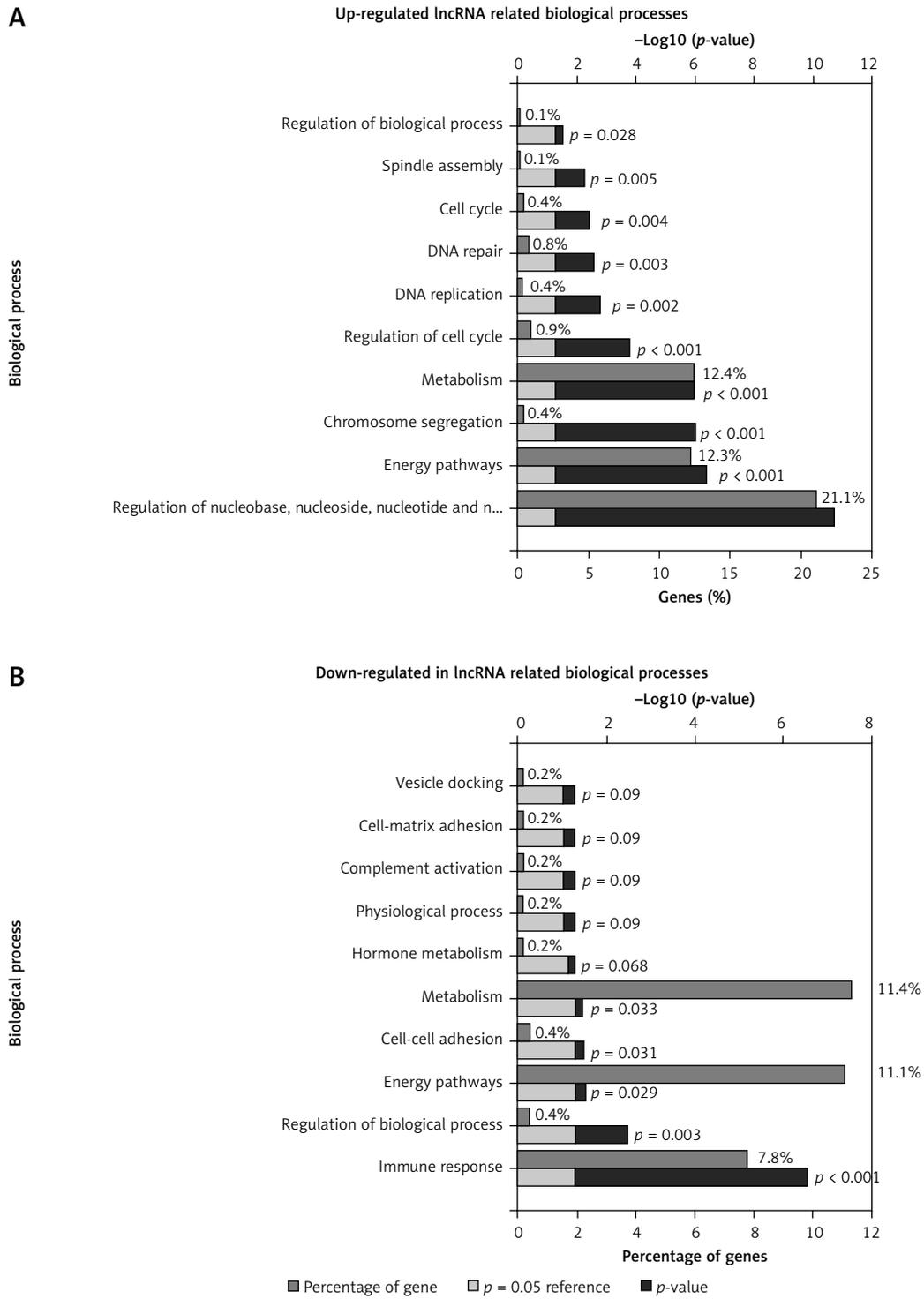
were associated with shorter survival time in the TCGA dataset. However, we found that C3P1-high (Figure 5 F) and C22orf45-high (Figure 5 E) patients showed higher survival rates compared to C3P1- and C22orf45-low patients.

**Discussion**

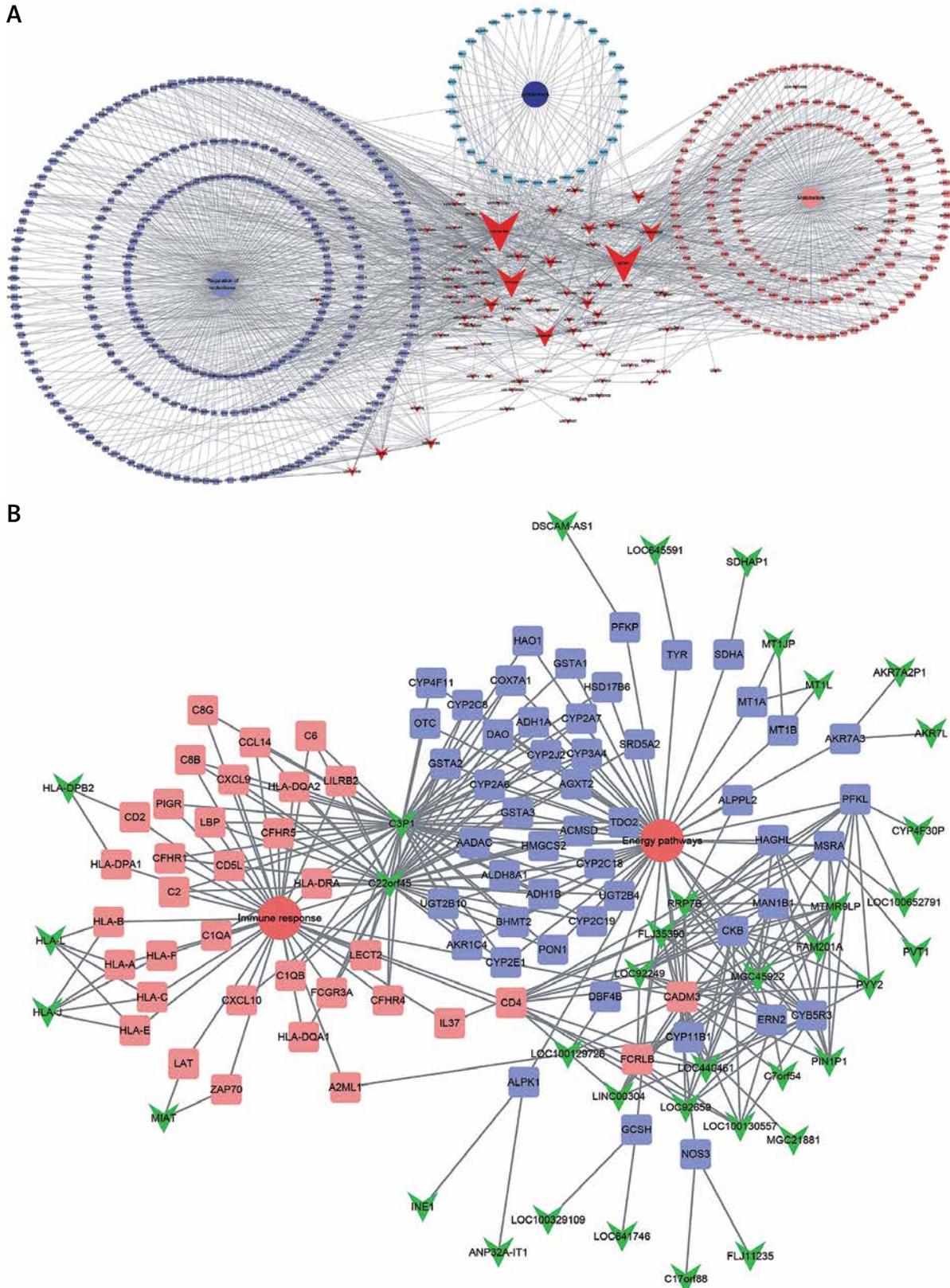
Hepatocellular carcinoma is one of the most common cancers worldwide [1]. Hepatitis C infection was one of the major etiological risk factors of HCCs [8]. However, the underlying mechanisms of HCV-induced hepatocarcinogenesis remain largely unclear. lncRNAs were a class of non-coding RNAs of more than 200 nucleotides in length [10]. A previous study demonstrated the key roles of lncRNAs in various human cancer types, including liver cancers [14]. In this study, we analyzed the public dataset GSE17856 and found that 454

lncRNAs (including 256 upregulated lncRNAs and 198 downregulated lncRNAs) were differentially expressed compared to the normal liver tissues.

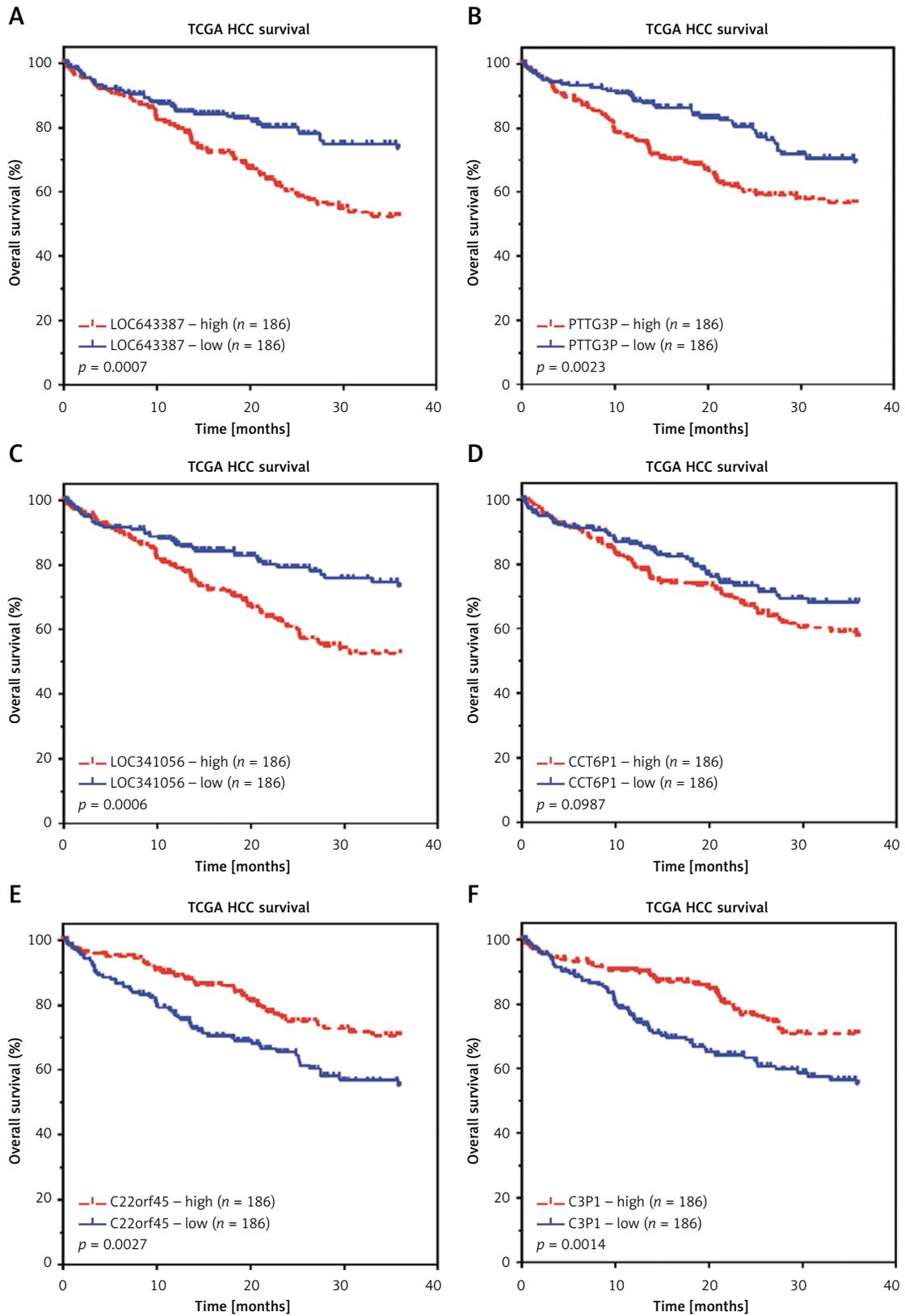
Recently, emerging studies have shown that lncRNAs were also associated with the progression of HCC. For example, authors [23–25] identified differently expressed lncRNAs in HCV-related HCC using RNA sequencing and expression profile analysis. Zhu *et al.* reported that LINC00052 upregulated EPB41L3 to inhibit migration and invasion of hepatocellular carcinoma by binding miR-452-5p [26]. lncRNA TSLNC8 was reported as a tumor suppressor by inactivating the IL-6/STAT3 signaling pathway [27]. However, very few reports have focused on the molecular function of lncRNAs in HCV-positive HCC. In this study, we performed bioinformatics analysis to explore the potential roles of lncRNAs in regulating HCV-pos-



**Figure 3.** GO analysis of up-regulated lncRNAs in biological process (A) and down-regulated lncRNAs in biological process (B)



**Figure 4.** Identification of key lncRNAs in regulating HCV-positive HCC development. We constructed upregulated lncRNA-mRNA-biological processes. LOC341056, CCT6P1, PTTG3P, LOC643387, and LOC100133920 were involved in regulating metabolism, regulation of nucleobase and cell proliferation processes (A). Down-regulated lncRNAs (C3P1 and C22orf45) were involved in regulating immune response and energy pathways (B). Triangle nodes, lncRNAs; square nodes, lncRNAs co-expressed genes; circle, biological processes



**Figure 5.** Key lncRNAs were associated with hepatocellular carcinoma prognosis. According to Kaplan-Meier analysis, we found that higher expression levels of LOC643387 (A), PTTG3P (B), LOC341056 (C), and CCT6P1 (D) were associated with shorter survival time in TCGA dataset. We found that C22orf45- (E) and C3P1- (F) high patients showed higher survival rates compared to C3P1- and C22orf45-low patients

Significance was defined as  $p < 0.05$  (\* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ).

itive HCC development. We first constructed dysregulated lncRNA-mRNA co-expressed networks. We found 125 up-regulated lncRNAs co-expressed with 1469 DEGs and 67 down-regulated lncRNAs co-expressed with 558 DEGs. Bioinformatics analysis showed that dysregulated lncRNAs in HCV-positive HCC were mainly involved in regulating metabolism (such as energy pathways, nucleic acid metabolism and amino acid and derivative metabolism) and cell proliferation related biological processes (including cell cycle, DNA repair, and DNA replication). Of note, emerging studies have indicated the important roles in HCC progression [28, 29]. However, very few reports have focused on exploring the roles of lncRNAs in HCC metabolism regulation. Our study for the first time provides useful information for exploring the effect of lncRNA on HCC metabolism.

The roles of most lncRNAs in HCC have remained unclear. In the present study, we identified lncRNAs (LOC341056, CCT6P1, PTTG3P, LOC643387, LOC100133920, C3P1 and C22orf45) as key lncRNAs. Furthermore, we constructed key lncRNA mediated mRNA-bp networks to reveal the potential roles of these key lncRNAs. Of these lncRNAs, PTTG3P has been reported to promote gastric tumor cell proliferation and invasion. However, the other lncRNAs were never reported in human diseases. According to bioinformatics analysis, we found that LOC341056, CCT6P1, PTTG3P, LOC643387, LOC100133920 were mainly involved in regulating metabolism and cell proliferation related processes, and C3P1 and C22orf45 were associated with immune response and energy pathways. Moreover, we performed Kaplan-Meier analysis to evaluate the prognostic value of these key lncRNAs. Higher expression levels of LOC643387, PTTG3P, LOC341056, CCT6P1 and lower C3P1 and C22orf45 expression were associated with shorter survival time in the TCGA dataset. These results suggested that these lncRNAs in HCV-positive HCC development could serve as biomarkers for hepatocellular carcinoma prognosis.

In conclusion, our study represents a comprehensive analysis of differentially expressed lncRNAs in HCV-positive HCC for the first time by analyzing the public dataset GSE17856. We identified 256 upregulated lncRNAs and 198 downregulated lncRNAs in HCV-positive HCC compared to the normal liver tissues. Co-expression network and GO analysis showed that these lncRNAs were involved in regulating metabolism, energy pathways, proliferation and immune response. Seven lncRNAs (LOC341056, CCT6P1, PTTG3P, LOC643387, LOC100133920, C3P1 and C22orf45) were identified as key lncRNAs and co-expressed with more than 100 DEGs in HCV-positive HCC. Kaplan-Meier analysis showed that higher expression levels of LOC643387, PTTG3P, LOC341056,

CCT6P1 and lower expression levels of C3P1 and C22orf45 were associated with shorter survival time in the TCGA dataset. We believe that this study can provide novel potential therapeutic and prognostic targets for HCV-positive HCC.

## Acknowledgments

This work was supported by Huimin Science and Technology Foundation Project of Xiamen City of China (No. 3502Z20164015) and Natural Science Foundation of Fujian Province of China (No. 2016J01629).

## Conflict of interest

The authors declare no conflict of interest.

## References

1. Bruix J, Gores GJ, Mazzaferro V. Hepatocellular carcinoma: clinical frontiers and perspectives. *Gut* 2014; 63: 844-55.
2. Giannelli G, Villa E, Lahn M. Transforming growth factor-beta as a therapeutic target in hepatocellular carcinoma. *Cancer Res* 2014; 74: 1890-4.
3. Gupta P, Cairns MJ, Saksena NK. Regulation of gene expression by microRNA in HCV infection and HCV-mediated hepatocellular carcinoma. *Virology* 2014; 11: 64.
4. Lee MH, Yang HJ, Yuan Y, L'Italien G, Chen CJ. Epidemiology and natural history of hepatitis C virus infection. *World J Gastroenterol* 2014; 20: 9270-80.
5. Kasprzak A, Rogacki K, Adamek A, et al. Tissue expression of beta-catenin and E- and N-cadherins in chronic hepatitis C and hepatocellular carcinoma. *Arch Med Sci* 2017; 13: 1269-80.
6. Chen G, Li X, Yang J, et al. Prognostic significance of cyclooxygenase-2 expression in patients with hepatocellular carcinoma: a meta-analysis. *Arch Med Sci* 2016; 12: 1110-7.
7. Gurtsevitch VE. Human oncogenic viruses: hepatitis B and hepatitis C viruses and their role in hepatocarcinogenesis. *Biochemistry (Mosc)* 2008; 73: 504-13.
8. Bouchard MJ, Navas-Martin S. Hepatitis B and C virus hepatocarcinogenesis: lessons learned and future challenges. *Cancer Lett* 2011; 305: 123-43.
9. Ringelhan M, Heikenwalder M, Protzer U. Direct effects of hepatitis B virus-encoded proteins and chronic infection in liver cancer development. *Dig Dis* 2013; 31: 138-51.
10. Huarte M. The emerging role of lncRNAs in cancer. *Nat Med* 2015; 21: 1253-61.
11. Martens-Uzunova ES, Bottcher R, Croce CM, Jenster G, Visakorpi T, Calin GA. Long noncoding RNA in prostate, bladder, and kidney cancer. *Eur Urol* 2014; 65: 1140-51.
12. De Leeneer K, Claes K. Non coding RNA molecules as potential biomarkers in breast cancer. *Adv Exp Med Biol* 2015; 867: 263-75.
13. Deng H, Wang JM, Li M, et al. Long non-coding RNAs: New biomarkers for prognosis and diagnosis of colon cancer. *Tumour Biol* 2017; 39: 1393383332.
14. He Y, Meng XM, Huang C, et al. Long noncoding RNAs: Novel insights into hepatocellular carcinoma. *Cancer Lett* 2014; 344: 20-7.
15. Fan H, Zhang Q, Zhao X, Lv P, Liu M, Tang H. Transcriptomic profiling of long non-coding RNAs in hepatitis

- B virus-related hepatocellular carcinoma. *Oncotarget* 2017; 8: 65421-34.
16. Li X, Wu Z, Fu X, Han W. lncRNAs: insights into their function and mechanics in underlying disorders. *Mutat Res Rev Mutat Res* 2014; 762: 1-21.
  17. Liang WC, Fu WM, Wong CW, et al. The lncRNA H19 promotes epithelial to mesenchymal transition by functioning as miRNA sponges in colorectal cancer. *Oncotarget* 2015; 6: 22513-25.
  18. Tsuchiya M, Parker JS, Kono H, Matsuda M, Fujii H, Rusyn I. Gene expression in nontumoral liver tissue and recurrence-free survival in hepatitis C virus-positive hepatocellular carcinoma. *Mol Cancer* 2010; 9: 74.
  19. Zhang X, Sun S, Pu JKS, et al. Long non-coding RNA expression profiles predict clinical phenotypes in glioma. *Neurobiol Dis* 2012; 48: 1-8.
  20. Cui X, Churchill GA. Statistical tests for differential expression in cDNA microarray experiments. *Genome Biol* 2003; 4: 210.
  21. Diboun I, Wernisch L, Orengo CA, Koltzenburg M. Microarray analysis after RNA amplification can detect pronounced differences in gene expression using limma. *BMC Genomics* 2006; 7: 252.
  22. Gondro C, Porto-Neto LR, Lee SH. R for genome-wide association studies. *Methods Mol Biol* 2013; 1019: 1-17.
  23. Lv J, Fan HX, Zhao XP, et al. Long non-coding RNA Uni-gene56159 promotes epithelial-mesenchymal transition by acting as a ceRNA of miR-140-5p in hepatocellular carcinoma cells. *Cancer Lett* 2016; 382: 166-75.
  24. Pan YF, Qin T, Feng L, Yu ZJ. Expression profile of altered long non-coding RNAs in patients with HBV-associated hepatocellular carcinoma. *J Huazhong Univ Sci Technol Med Sci* 2013; 33: 96-101.
  25. Gong X, Wei W, Chen L, Xia Z, Yu C. Comprehensive analysis of long non-coding RNA expression profiles in hepatitis B virus-related hepatocellular carcinoma. *Oncotarget* 2016; 7: 42422-30.
  26. Zhu L, Yang N, Chen J, et al. LINC00052 upregulates EP-B41L3 to inhibit migration and invasion of hepatocellular carcinoma by binding miR-452-5p. *Oncotarget* 2017; 8: 63724-37.
  27. Zhang J, Li Z, Liu L, et al. Long noncoding RNA TSLNC8 is a tumor suppressor that inactivates the IL-6/STAT3 signaling pathway. *Hepatology* 2018; 67: 171-87.
  28. Ockner RK, Kaikous RM, Bass NM. Fatty-acid metabolism and the pathogenesis of hepatocellular carcinoma: review and hypothesis. *Hepatology* 1993; 18: 669-76.
  29. Li J, Huang Q, Long X, et al. CD147 reprograms fatty acid metabolism in hepatocellular carcinoma cells through Akt/mTOR/SREBP1c and P38/PPARalpha pathways. *J Hepatol* 2015; 63: 1378-89.