Long non-coding RNA expression profiles of hepatitis C virus-related dysplasia and hepatocellular carcinoma

Haohai Zhang[1,*], Chengpei Zhu[1,*], Yi Zhao[2,*], Ming Li[3], Liangcai Wu[1], Xiaobo Yang[1], Xueshuai Wan[1], Anqiang Wang[1], Michael Q. Zhang[1,2], Xinting Sang[1], Haitao Zhao[1]

[1]Department of Liver Surgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

[2]School of Medicine, MOE Key Laboratory of Bioinformatics and Bioinformatics Division, Center for Synthetic and System Biology, Tsinghua University, Beijing, China

[3]Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China

[*]These authors have contributed equally to this work

Correspondence to:
Haitao Zhao, **e-mail:** zhaoht@pumch.cn
Michael Q. Zhang, **e-mail:** michaelzhang@tsinghua.edu.cn
Xinting Sang, **e-mail:** SangXT@pumch.cn

ABSTRACT

Recently, long non-coding RNAs (lncRNAs) were found to be implicated in cancer progression. However, the contributions of lncRNAs to Hepatitis C virus-related hepatocellular carcinoma (HCC) remain largely unknown. Here, we characterized lncRNA expression in 73 tissue samples from several different developmental stages of HCV-related hepatocarcinogenesis by repurposing microarray data sets. We found that the expression of 7 lncRNAs in preneoplastic lesions and HCC was significantly different. Among these significantly differently expressed lncRNAs, the lncRNA LINC01419 transcripts were expressed at higher levels in early stage HCC compared to dysplasia and as compared with early stage HCC, lncRNA AK021443 level increase in advanced stage HCC while lncRNA AF070632 level decrease in advanced stage HCC. Using quantitative real-time reverse-transcription PCR, we validated that LINC01419 was significantly overexpressed in HBV-related and HCV-related HCC when compared with matched non-tumor liver tissues. Moreover, functional predictions suggested that LINC01419 and AK021443 regulate cell cycle genes, whereas AF070632 is associated with cofactor binding, oxidation-reduction and carboxylic acid catabolic process. These findings provide the first large-scale survey of lncRNAs associated with the development of hepatocarcinogenesis and may offer new diagnostic biomarkers and therapeutic targets for HCV-related HCC.

INTRODUCTION

Hepatocellular carcinoma (HCC) is one of the most frequently diagnosed cancers in the world and is the second leading cause of cancer-related deaths worldwide [1]. Despite improvements in HCC patient survival that have resulted from advances in surgical strategies and improved surveillance of patients with a high risk for HCC development, five-year survival rates remain at approximately 26% [2]. The major risk factors for the development of HCC include chronic liver disease as a result of infection with either the hepatitis B virus (HBV) or the hepatitis C virus (HCV) and exposure to carcinogens, such as aflatoxin B1 [3]. Chronic infection with HBV is related to HCC in developing countries, whereas chronic infection with HCV is the main risk factor in developed countries [4]. Hepatocellular carcinoma pathogenesis is a slow process that may take more than 30 years to develop after the initial diagnosis of chronic infection with HBV or HCV. During the long preneoplastic stage, genetic alteration progressively changes the hepatocellular phenotype, leading to cell death, cellular proliferation, dysplasia and

neoplasia [5]. An investigation of the molecular events across this slow process could identify crucial genes involved in HCC pathogenesis [6–8]. Such genes are potential biomarkers for the early detection of HCC [8, 9] or for predicting the prognosis of patients with HCC [10–12]. Unfortunately, most studies to date have explored only protein-coding genes or miRNA sequences. However, in addition to approximately 20,000 protein-coding genes and encoded small RNA molecules, the human transcriptome comprises a large number of long, nonprotein-coding RNAs (long noncoding RNAs; lncRNAs) [13].

lncRNAs are defined as nonprotein-coding RNAs that are greater than 200 nucleotides in length [14]. Multiple studies have suggested that several lncRNAs participate in a variety of biological process [15–18] and are associated with human diseases [19, 20]. Furthermore, lncRNAs have been reported to be involved in HCC through various mechanisms, including epigenetic silencing, mRNA splicing and translation, lncRNA-protein interaction and genetic variation [21, 22]. Several lncRNAs serve as potential biomarkers for the detection or prognosis of HCC. For example, the HULC gene, located on chromosome 6p24.3, is significantly overexpressed in HCC tissues [23]. The expression of HULC can be up-regulated by the transcription factor CREB (cAMP responsive element binding protein) and the Hepatitis B virus X protein (HBX) [24, 25]. In addition, HULC expression in HCC tissue and plasma of HCC patients is associated with Edmondson histological grades and HBV status. Moreover, HULC detection frequencies in HCC patients are higher than those in healthy people. These characteristics suggest that HULC may act as a diagnostic and/or prognostic biomarker for HCC [26]. Another potential lncRNA biomarker for HCC is lncRNA-HEIH, whose expression in HBV-related HCC tissue was reported as an independent risk factor for recurrence-free survival (RFS) [27]. Moreover, lncRNA-HEIH was overexpressed in HCC tissue compared with adjacent normal liver tissue and was higher in cirrhotic liver tissue than in healthy liver tissue [27]. Taken together, these previous findings suggest that lncRNAs may play an important role in the pathogenesis of HCC. However, most previous studies have only examined lncRNA expression in either HBV-induced HCC [26–28] or in mixed etiology (HBV and HCV) HCC [29]. The specific role of lncRNAs in the pathogenesis of HCV-related HCC remains unclear. Thus, the goal of our study was to explore lncRNA expression at the different stages of HCV-induced hepatocarcinogenesis.

In this study, we examined lncRNA expression in 73 tissue samples, including healthy liver tissue, cirrhotic tissue, dysplastic nodules, and HCC samples. The cirrhotic tissue, dysplastic nodules, and HCC samples were obtained from patients with HCV infections. Our data show that several lncRNAs could serve as biomarkers for the detection or prognosis of HCV-related HCC. Furthermore, we predicted the functions of the lncRNAs by investigating coexpression networks of the differentially expressed lncRNAs and protein-coding genes.

RESULTS

## Transcriptomic analysis of HCV-induced dysplasia and hepatocellular carcinoma

Of the publicly available gene expression microarray data sets, we focused on repurposing microarray data sets comprising 75 samples that represent all stages of HCV-induced HCC. For these Affymetrix microarray platforms (HGU133plus2.0), we re-annotated the entire collection of probes using the ncFANs utility. As a result, we identified probes for 17,600 protein-coding genes and 2,863 lncRNAs in these published microarrays. Next, we reanalyzed the expression patterns of the protein-coding genes across 5 groups of patient samples: control, cirrhosis, dysplasia, early HCC and advanced HCC. Because HCC always occurs in cirrhotic livers, we compared the gene profiles associated with the preneoplastic lesion stage (cirrhosis and dysplasia stages) with those of HCC samples to identify the critical molecular events associated with HCC progression. Our results indicated that 585 protein-coding genes (Supplementary File 1) and 7 lncRNA genes (Table 1) were differentially expressed between preneoplastic lesions and HCC samples (p value < 0.05, fold change >2). Then, we performed a gene ontology (GO) analysis to enrich for gene sets based on differentially expressed protein-coding genes. We observed that genes that were up-regulated in HCC samples mainly participated in mitosis and the cell cycle (Fig. 1A, Supplementary Table 1), whereas genes that were down-regulated in HCC samples were associated with other functional capacities, such as response to steroid hormone stimulus, immune and inflammatory responses and normal liver function, which includes glucose and organic acid metabolism, organic acid biosynthesis and coagulation (Fig. 1B, Supplementary Table 2).

Deregulated lncRNAs in HCC

Next, we evaluated the changes in expression of the 7 lncRNAs during the five stages of HCC (healthy liver, cirrhosis, dysplasia, early HCC and advanced HCC) using one-way ANOVA (paired with an F test). We noted that LINC01419 was characterized by a significant increase in transcript expression from dysplasia to early HCC (Fig. 2A). The lncRNA AK021443 was also up-regulated in advanced HCC samples when compared with early HCC (Fig. 2B). Moreover, expression of LINC01419 and AK021443 was up-regulated in HCC tissues when compared with non-tumor liver tissue. AF070632 expression was down-regulated in HCC and was decreased in advanced HCC when compared with early HCC (Fig. 2C). These results suggest that LINC01419 may

Table 1: Summary of lncRNAsthat are differentially expressed between preneoplasticlesions and HCC

Gene symbol	Chromosomal location	Preneoplastic lesions vs. HCC	
		Regulation	*P*-value
LINC01419	Chr8:84315993-84321132(+)	Up	0.000633277280341197
BC014579	Chr10:5015558-5017273(+)	Up	0.0000959501811649745
AK021443	Chr10:96372044-96373661(+)	Up	6.56187823907914E-06
RP11-401P9.4	Chr16:50679720-50683160(+)	Up	0.00383700431326679
RP11-304 L19.5	Chr16:2204798-2205359(−)	Up	0.0000692618363817383
CTB-167B5.2	Chr7:87900207-87903065(−)	Down	1.43661312617711E-09
AF070632	Chr1:112313293-112315191(−)	Down	2.60357616541779E-06
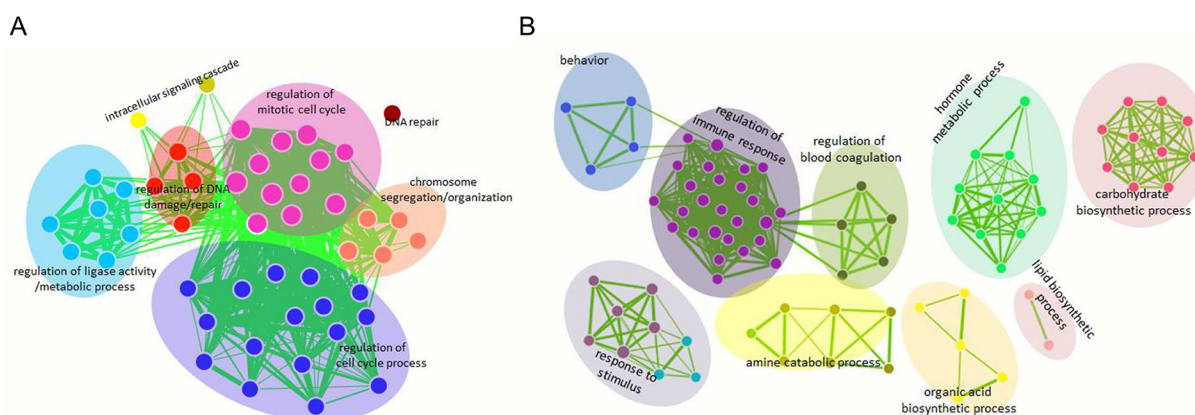
Figure 1: Functional enrichment maps of deregulated protein-coding genes in HCC. The gene sets for GO terms were visualized using the Cytoscape Enrichment Map plugin. Each node represents a GO term. Node size is indicative of the number of genes in a set. The thickness of each line is indicative of the number of genes shared between the connected gene sets. The enriched gene sets of up-regulated protein-coding genes in HCC are shown in **A.** and the down-regulated protein-coding genes in HCC are shown in **B.**

be related to the initiation of HCC, whereas AK021443 and AF070632 may be associated with the progression of HCC.

Because early detection biomarkers are critical for effective HCC care, we chose to further validate LINC01419, whose expression increased sharply from dysplasia to early HCC. We analyzed LINC01419 expression using quantitative real-time polymerase chain reaction (qRT-PCR) in 15 pairs of HCV-related HCC samples, 55 pairs of HBV-related HCC samples and their corresponding adjacent non-tumor liver tissues (Fig. 2D). These results confirmed that LINC01419 was overexpressed in HCV-related HCC and HBV-related HCC. We noted that LINC01419 was poorly expressed in adjacent normal liver tissue. The Ct values were greater than 30 in all adjacent normal liver tissues (data not shown). In addition, using the ENCODE database, we identified a CpG methylation site 113bp upstream of the transcription start site for LINC01419 in all six cell lines (HepG2, GM12878, H1-hESC, K562, Hela-S3, and HUVEC; Fig. 2E). Because DNA methylation of a

promoter can repress the expression of the gene under the control of that promoter [30], this finding may partially explain the low level of LINC01419 expression. However, these findings should be confirmed experimentally.

Prediction of lncRNA function

To investigate the role of lncRNAs in HCC pathogenesis, we used the gene expression profiles of 75 patient samples to construct a coding-non-coding co-expression network. The resulting network includes 220 lncRNA genes and 3,924 coding genes with 105,508 coding-coding edges, 7,114 noncoding-coding edges and 324 noncoding-noncoding edges (online resource). Next, we predicted the functions of the lncRNAs using two different methods, a module-based analysis and a hub-based analysis. First, the co-expression network was parsed into forty module-based subnetworks using an MCL algorithm. Twenty of these modules had at least one enriched GO term. Next, we obtained 154 lncRNA-centered subnetworks with at least one enriched GO term
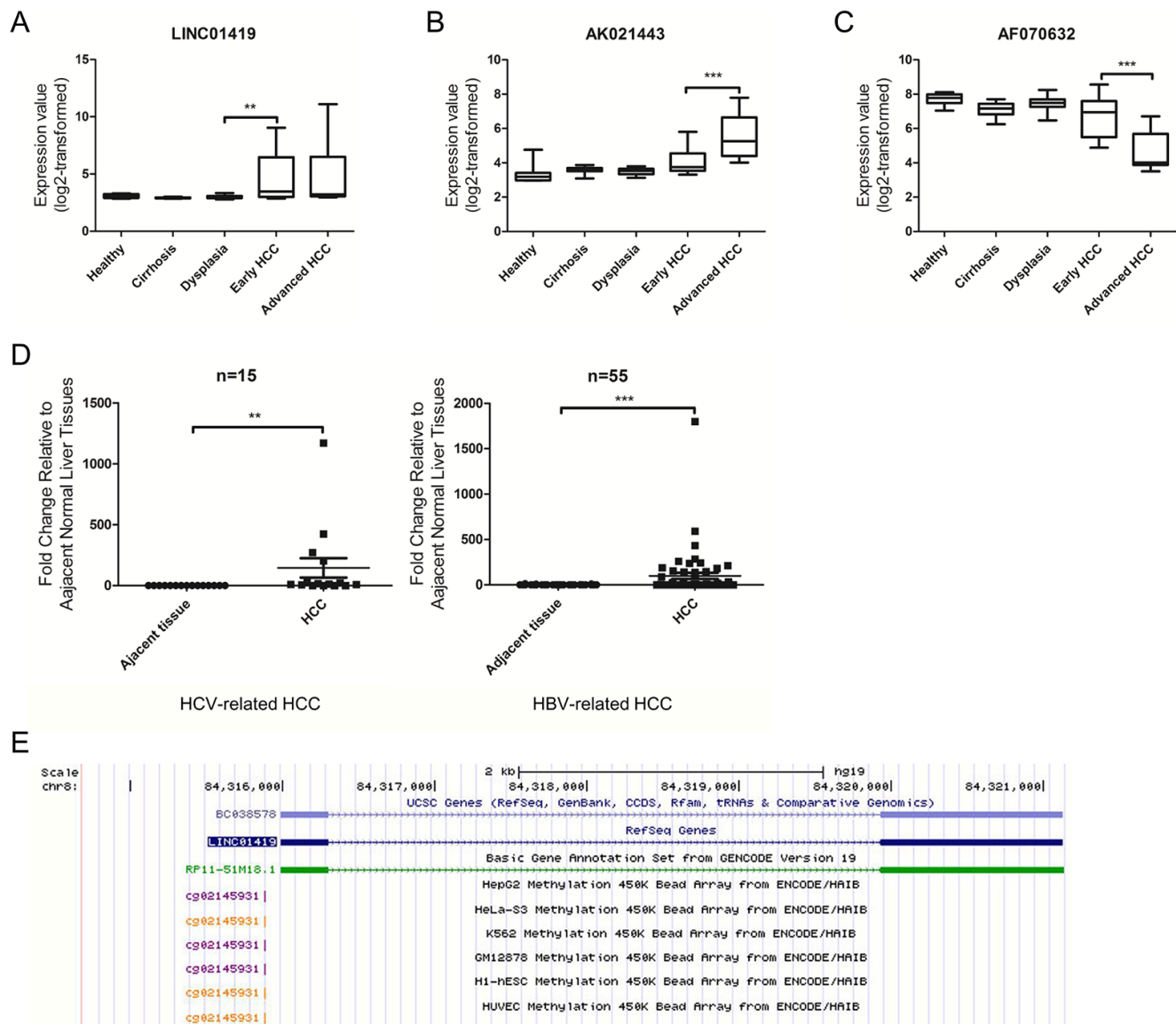
Figure 2: Expression of deregulated lncRNAs in HCC. Expression values of LINC01419 (A), AK021443 (B), and AF070632 (C) in the five stages of HCC based on microarray data from73 patient samples (10 healthy liver samples, 12 cirrhotic liver samples, 17 dysplastic liver samples, 16 early HCC samples, and 17 advanced HCC samples). For **A, B,** and **C.** the *p*-values were calculated using the Mann-Whitney test. **D.** LINC01419 expression in 55 pairs of HBV-related HCC and 15 pairs of HCV-related HCC. The expression level of LINC01419 was analyzed by qRT-PCR. The *p*-values were calculated using the Wilcoxon signed-rank test. **E.** Genomic context of LINC01419. A CpG methylation site was found upstream of the LINC01419 coding sequence.

by parsing the co-expression network into different hub-based subnetworks.

The functions of the 3 deregulated lncRNAs associated with the initiation or progression of HCC were predicted using the module-based method or the hub-based method. Our results suggest that LINC01419 and AK021443 were classified into the same module with the other 36 lncRNAs and 637 protein-coding genes. Because the co-expressed module could be regarded as a functional unit, LINC01419 and AK021443 may have the similar function with those 637 protein coding genes. Thus, by performing Go analysis on those protein-coding genes, we conclude that the functions of LINC01419 and AK021443 were mostly involved with cell cycle progression (Fig. 3A, Supplementary Table 3). The function of AF070632 was predicted using hub-based analysis. We found that it was surrounded by 64 differentially expressed protein-coding genes (Fig. 3B). Consistence with enriched function of these 64 protein-coding genes, AF070632 is mainly related to oxidation-reduction, cofactor binding and carboxylic acid catabolic process (Supplementary Table 4).
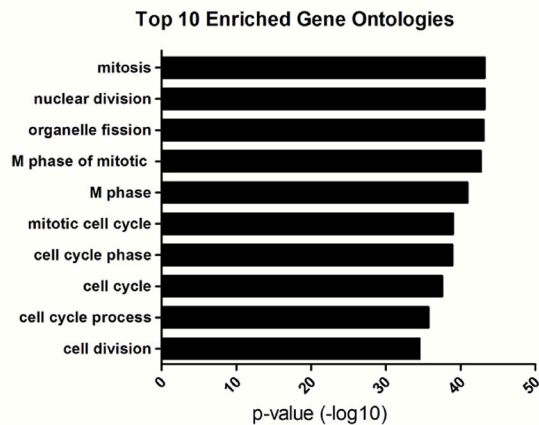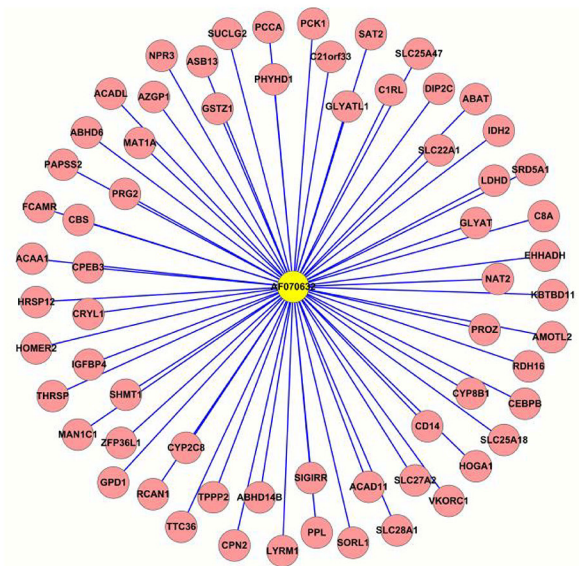
Figure 3: LncRNA functional prediction. A. Functional enrichment result of the module containing LINC01419 and AK021443. **B.** The AF070632-centered subnetwork. Genes colored in red are protein-coding genes.

DISCUSSION

In the present study, we reanalyzed a large microarray dataset to investigate the pathogenesis of HCV-related HCC. There are three main advantages of using this approach. First, with the exception often healthy liver tissue samples, the samples used in this microarray dataset were derived from patients with hepatitis C virus infection, which is the leading cause of HCC in the United States and Europe [4]. This study focused on lncRNA expression in HCV-induced HCC and identified lncRNAs that could be used as biomarkers for HCV-related HCC. Second, unlike some other studies comparing HCC tissues to adjacent normal tissues [28, 31], the samples in this microarray dataset were comprised of tissues representing each stage of HCV-related HCC [7], which provides better insight into the molecular basis underlying hepatocarcinogenesis. Finally, although RNA sequencing (RNA-seq) could be used to comprehensively identify lncRNA expression, the number of publically available RNA-seq databases and the application of RNA-seq for large-scale studies are limited by its cost. In contrast, there are many publically available microarray data sets for various types of cancer and clinical experiments. Because publically available microarray data contain probes that potentially match lncRNAs, we were able to repurpose these datasets to analyze lncRNA expression after reannotating the microarray probes. This approach has been used in other studies [32]. Moreover, the increased detection sensitivity of microarrays can facilitate the investigation of low-abundance lncRNAs.

Increasing evidence has suggested that lncRNAs serve critical functions [14]. Several studies have reported that lncRNAs are differentially expressed in different stages of HBV-related HCC, which suggests that lncRNAs may participate in the hepatocarcinogenic process. For example, lncRNA-HEIH is overexpressed in HCC tissue when compared with adjacent normal tissue, suggesting that this lncRNA could promote tumor growth. Notably, lncRNA-HEIH expression was also up-regulated in cirrhotic liver tissue when compared with healthy liver tissue [27]. Furthermore, lncRNA-ATB, another lncRNA, is overexpressed in HCC tissues when compared with adjacent normal tissues. lncRNA-ATB expression is negatively correlated with prognosis in HCC patients. Interestingly, the expression of lncRNA-ATB was up-regulated in portal vein tumor thrombi when compared with HCC tissues [33]. In the current study, we describe the lncRNA signatures of the multiple stages of HCV-related HCC and identify a set of lncRNAs that could play a key role in hepatocarcinogenesis. We also identify 7 lncRNAs that were deregulated between preneoplastic lesions and HCC. Further analysis showed that 3 of the 7 lncRNAs were differentially expressed between dysplasia, early HCC, and advanced HCC. Our results suggest that these 3 lncRNAs are associated with the initiation or progression of HCC and may serve as diagnostic or prognostic biomarkers for HCC. Among the 3 deregulated lncRNAs, the up-regulated expression of LINC01419 in HCC was confirmed by quantitative real-time polymerase chain reaction (qRT-PCR). Interestingly, LINC01419 is poorly expressed in adjacent normal liver

tissue, suggesting that LINC01419could serve as a specific biomarker for the detection of HCC.

The cell cycle controls cell division, which is critical to cancer. Cyclin-dependent kinases (CDKs), which are key regulators of the cell cycle, have been investigated as potential therapeutic targets for cancer [34]. Cell cycle deregulation is also common in HCC [35]. For example, a recent study showed that cell cycle-related kinase, a member of the latest cyclin-dependent kinase, could facilitate the development of HCC by activating β-catenin/T cell factor signaling [36]. TTK is a critical mitotic checkpoint protein and participates in the p53 signaling pathway. The expression of TTK in HCC tissues is associated with the Edmonson tumor grade, recurrence-free survival and overall survival [37]. Moreover, SKP2 and CKS1 in HCC tissues could promote G1-S transition and active DNA synthesis by degrading cell cycle inhibitors [38, 39]. However, the regulatory role of lncRNAs in the cell cycle is poorly understood. In this study, using gene co-expression network analysis, we found that LINC01419 and AK021443 were clustered into one functional module for cell cycle regulation. In addition, these two lncRNAs were over expressed in HCC tissues. Thus, our findings suggest that LINC01419 and AK021443 could regulate cell cycle activity and thereby promote tumor growth.

Taken together, our results provide a comprehensive understanding of lncRNAs in HCV-related HCC and a basis for identifying specific biomarkers or therapeutic targets for HCC.

MATERIALS AND METHODS

Microarray data information and computational process

The microarray data used in this study were generated using the Affymetrix platform HG-U133A Plus 2.0. We downloaded these publicly available data sets from the Gene Expression Omnibus (GEO) database (accession number: GSE6764). The microarray data assessed the gene expression profiles of 75 samples consisting of 10 normal liver tissue samples, 13 cirrhotic liver tissue samples, 17 dysplastic nodules, and 35 HCCs. The preneoplastic lesions (cirrhosis and dysplasia) and HCC samples were obtained from 38 patients infected with HCV. Patients with other etiologies, such as HBV infection, were excluded from our analysis.35 HCC samples were categorized using the four following criteria: (i) according to tumor size, where well-differentiated tumors <2 cm in diameter with no vascular invasion and satellites were categorized as very early stage HCC (8 cases); (ii) according to tumor differentiation, where microscopic tumors measuring <2 cm in diameter, well-to moderately differentiated tumors measuring 2–5 cm in diameter without vascular invasion or satellites and tumors with 2–3 well-differentiated nodules measuring

<3 cm in diameter were categorized as early stage HCC (10 cases); (iii) according to tumor number, where poorly differentiated tumors measuring >2 cm in diameter with microvascular invasion and satellites or tumors measuring >5 cm in diameter were categorized as advanced stage HCC (7 cases); (iv) according to vascular invasion and satellites, where tumors with macrovascular invasion or diffuse liver involvement were categorized as very advanced stage HCC (10 cases) [7].

Patients

Sixty-five frozen HCC tissues and matched adjacent normal liver tissues were randomly obtained from sixty-five patients who underwent hepatectomy at Peking Union Medical College Hospital (Beijing, China). This patient cohort consisted of fifty-five patients who were infected with HBV and fifteen patients who were infected with HCV. We used these samples to perform quantitative real-time polymerase chain reaction (qRT-PCR) analysis.

Repurposing microarray data to analyze the expression of protein-coding genes and lncRNA genes

Non-coding RNA Function Annotation Server (ncFANs) can be used to quantify the levels of messenger RNAs and long non-coding RNAs depending on the re-annotation of Affymetrix microarray probes [40]. We submitted microarray data (in CEL format) to ncFANs and obtained both protein-coding and lncRNA gene expression values (log2-transformed) for 73 samples using R (2 samples were eliminated for quality control reasons). The Robust Multichip Average (RMA) analysis, Student'st-test, one-way ANOVA (paired with an F test) and Benjamini Hochberg (BH) FDR correction were performed using the Limma statistical package. Genes with a fold-change >2 and BH FDR-adjusted p-values < 0.05 were considered differentially expressed.

PCR analysis

Total RNA was isolated using TRIzol reagent (Invitrogen). First-strand cDNA was synthesized using a High Capacity cDNA Reverse Transcription kit (Applied Biosystems) according to instructions suggested by the manufacturer. Real-time PCR was performed using Power SYBR® Green Master mix (Applied Biosystems) and a StepOne™ Real-Time PCR System (Applied Biosystems, Foster City, USA). The relative expression levels of the lncRNAs were calculated using the comparative Ct method. GAPDH gene expression was included as an internal control. The primers for LINC01419were (forward) 5′-GAAACTCCGAACACATCTG-3′ and (reverse) 5′-TTCTCCTGCTGGTTGATT-3′; for GAPDH, (forward) 5′- TCAAGGCTGAGAACGGGAAG-3′ and (reverse) 5′- GTGAAGACGCCAGTGGACT-3′.

Functional enrichment of protein-coding genes

The DAVID Bioinformatics Tool [41] was used to identify functional enrichment of target protein-coding genes. This tool can be used to perform gene ontology analysis to identify biological processes related to protein-coding genes. For this analysis, we used a *p*-value significance threshold of < 0.01. Finally, we used the Enrichment Map plugin [42] for Cytoscape [43] to visualize the biological process organization.

Functional analysis of lncRNAs

Gene expression data for 73 tissue samples was submitted to ncFANs to generate a coding-non-coding co-expression network. To build the co-expression network, we chose genes (both protein-coding genes and lncRNA genes) with expression differences that ranked in the top 75th percentile. A Pearson's correlation coefficient or a Spearman Rank correlation coefficient was used to estimate the relationship of each gene pair, including coding-coding, coding-lncRNA and lncRNA-lncRNA gene pairs. The *p*-values of the correlation coefficients for all gene pairs were calculated using Fisher's asymptotic test and then were adjusted using Bonferronimultiple testing correction. We chose co-expression gene pairs with *p*-values < 0.01 to construct the coding-non-coding co-expression network. Based on the coding-non-coding co-expression network, we used module-based and hub-based methods embedded in ncFANs to predict the function of the lncRNAs. In the module based method, modules of co-expressed genes in the coding-non-coding co-expression network were identified using the Markov cluster algorithm (MCL). The protein coding genes in the same module as the lncRNAs were further analyzed using gene ontology analysis to predict the functions of the lncRNAs in this module. In the hub-based method, the co-expression network was parsed to identify subnetworks comprising a central lncRNA and its neighboring protein-coding genes. The functional enrichment of these connected protein-coding genes was considered the function of the corresponding lncRNA. Only function enrichments with *p*-values < 0.01 were retained.

CONFLICTS OF INTEREST

The authors have nothing to disclose.

REFERENCES

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. CA: a cancer journal for clinicians. 2011; 61:69–90.

2. Simard EP, Ward EM, Siegel R, Jemal A. Cancers with increasing incidence trends in the United States: 1999 through 2008. CA: a cancer journal for clinicians. 2012; 62:118–128.

3. Bruix J, Boix L, Sala M and Llovet JM. Focus on hepatocellular carcinoma. Cancer Cell. 2004; 5:215–219.

4. Llovet JM, Burroughs A, Bruix J. Hepatocellular carcinoma. The Lancet. 2003; 362:1907–1917.

5. Thorgeirsson SS, Grisham JW. Molecular pathogenesis of human hepatocellular carcinoma. Nat Genet. 2002; 31:339–346.

6. Nam SW, Park JY, Ramasamy A, Shevade S, Islam A, Long PM, Park CK, Park SE, Kim SY, Lee SH, Park WS, Yoo NJ, Liu ET, Miller LD, Lee JY. Molecular changes from dysplastic nodule to hepatocellular carcinoma through gene expression profiling. Hepatology. 2005; 42:809–818.

7. Wurmbach E, Chen YB, Khitrov G, Zhang W, Roayaie S, Schwartz M, Fiel I, Thung S, Mazzaferro V, Bruix J, Bottinger E, Friedman S, Waxman S, Llovet JM. Genome-wide molecular profiles of HCV-induced dysplasia and hepatocellular carcinoma. Hepatology. 2007; 45:938–947.

8. Chuma M, Sakamoto M, Yamazaki K, Ohta T, Ohki M, Asaka M, Hirohashi S. Expression profiling in multistage hepatocarcinogenesis: identification of HSP70 as a molecular marker of early hepatocellular carcinoma. Hepatology. 2003; 37:198–207.

9. Capurro M, Wanless IR, Sherman M, Deboer G, Shi W, Miyoshi E, Filmus J. Glypican-3: a novel serum and histochemical marker for hepatocellular carcinoma. Gastroenterology. 2003; 125:89–97.

10. Lee JS, Chu IS, Heo J, Calvisi DF, Sun Z, Roskams T, Durnez A, Demetris AJ, Thorgeirsson SS. Classification and prediction of survival in hepatocellular carcinoma by gene expression profiling. Hepatology. 2004; 40:667–676.

11. Ye QH, Qin LX, Forgues M, He P, Kim JW, Peng AC, Simon R, Li Y, Robles AI, Chen Y, Ma ZC, Wu ZQ, Ye SL, Liu YK, Tang ZY, Wang XW. Predicting hepatitis B virus-positive metastatic hepatocellular carcinomas using gene expression profiling and supervised machine learning. Nature medicine. 2003; 9:416–423.

12. Iizuka N, Oka M, Yamada-Okabe H, Nishida M, Maeda Y, Mori N, Takao T, Tamesa T, Tangoku A, Tabuchi H,

Hamada K, Nakayama H, Ishitsuka H, Miyamoto T, Hirabayashi A, Uchimura S, et al. Oligonucleotide microarray for prediction of early intrahepatic recurrence of hepatocellular carcinoma after curative resection. The Lancet. 2003; 361:923–929.

13. Wilusz JE, Sunwoo H, Spector DL. Long noncoding RNAs: functional surprises from the RNA world. Genes Dev. 2009; 23:1494–1504.

14. Nagano T, Fraser P. No-nonsense functions for long non-coding RNAs. Cell. 2011; 145:178–181.

15. Tian D, Sun S, Lee JT. The long noncoding RNA, Jpx, is a molecular switch for X chromosome inactivation. Cell. 2010; 143:390–403.

16. Hung T, Wang Y, Lin MF, Koegel AK, Kotake Y, Grant GD, Horlings HM, Shah N, Umbricht C, Wang P, Wang Y, Kong B, Langerod A, Borresen-Dale AL, Kim SK, van de Vijver M, et al. Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. Nat Genet. 2011; 43:621–629.

17. Orom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, Bussotti G, Lai F, Zytnicki M, Notredame C, Huang Q, Guigo R, Shiekhattar R. Long noncoding RNAs with enhancer-like function in human cells. Cell. 2010; 143:46–58.

18. Huarte M, Guttman M, Feldser D, Garber M, Koziol MJ, Kenzelmann-Broz D, Khalil AM, Zuk O, Amit I, Rabani M, Attardi LD, Regev A, Lander ES, Jacks T, Rinn JL. A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. Cell. 2010; 142:409–419.

19. Szymanski M, Barciszewska MZ, Erdmann VA, Barciszewski J. A new frontier for molecular medicine: non-coding RNAs. Biochim Biophys Acta. 2005; 1756:65–75.

20. Costa FF. Non-coding RNAs: new players in eukaryotic biology. Gene. 2005; 357:83–94.

21. Li CH, Chen Y. Targeting long non-coding RNAs in cancers: progress and prospects. The international journal of biochemistry & cell biology. 2013; 45:1895–1910.

22. Huang JL, Zheng L, Hu YW, Wang Q. Characteristics of long non-coding RNA and its relation to hepatocellular carcinoma. Carcinogenesis. 2014; 35:507–514.

23. Panzitt K, Tschernatsch MM, Guelly C, Moustafa T, Stradner M, Strohmaier HM, Buck CR, Denk H, Schroeder R, Trauner M, Zatloukal K. Characterization of HULC, a novel gene with striking up-regulation in hepatocellular carcinoma, as noncoding RNA. Gastroenterology. 2007; 132:330–342.

24. Wang J, Liu X, Wu H, Ni P, Gu Z, Qiao Y, Chen N, Sun F, Fan Q. CREB up-regulates long non-coding RNA, HULC expression through interaction with microRNA-372 in liver cancer. Nucleic Acids Res. 2010; 38:5366–5383.

25. Du Y, Kong G, You X, Zhang S, Zhang T, Gao Y, Ye L, Zhang X. Elevation of highly up-regulated in liver cancer (HULC) by hepatitis B virus X protein promotes hepatoma

cell proliferation via down-regulating p18. J Biol Chem. 2012; 287:26302–26311.

26. Xie H, Ma H, Zhou D. Plasma HULC as a promising novel biomarker for the detection of hepatocellular carcinoma. BioMed research international. 2013; 2013:136106.

27. Yang F, Zhang L, Huo XS, Yuan JH, Xu D, Yuan SX, Zhu N, Zhou WP, Yang GS, Wang YZ, Shang JL, Gao CF, Zhang FR, Wang F, Sun SH. Long noncoding RNA high expression in hepatocellular carcinoma facilitates tumor growth through enhancer of zeste homolog 2 in humans. Hepatology. 2011; 54:1679–1689.

28. Yuan SX, Yang F, Yang Y, Tao QF, Zhang J, Huang G, Yang Y, Wang RY, Yang S, Huo XS, Zhang L, Wang F, Sun SH, Zhou WP. Long noncoding RNA associated with microvascular invasion in hepatocellular carcinoma promotes angiogenesis and serves as a predictor for hepatocellular carcinoma patients' poor recurrence-free survival after hepatectomy. Hepatology. 2012; 56:2231–2241.

29. Ishibashi M, Kogo R, Shibata K, Sawada G, Takahashi Y, Kurashige J, Akiyoshi S, Sasaki S, Iwaya T, Sudo T, Sugimachi K, Mimori K, Wakabayashi G, Mori M. Clinical significance of the expression of long non-coding RNA HOTAIR in primary hepatocellular carcinoma. Oncol Rep. 2013; 29:946–950.

30. Roadmap Epigenomics C, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, Kheradpour P, Zhang Z, Wang J, Ziller MJ, Amin V, Whitaker JW, Schultz MD, Ward LD, Sarkar A, et al. Integrative analysis of 111 reference human epigenomes. Nature. 2015; 518:317–330.

31. Oona Delpuech, Jean-Baptiste Trabut, Francoise Carnot, Jean Feuillard, Christian Brechot and Dina Kremsdorf. Identification, using cDNA macroarray analysis, of distinct gene expression profiles associated with pathological and virological features of hepatocellular carcinoma. Oncogene. 2002; 21:2926–2937.

32. Du Z, Fei T, Verhaak RG, Su Z, Zhang Y, Brown M, Chen Y, Liu XS. Integrative genomic analyses reveal clinically relevant long noncoding RNAs in human cancer. Nature structural & molecular biology. 2013; 20:908–913.

33. Yuan JH, Yang F, Wang F, Ma JZ, Guo YJ, Tao QF, Liu F, Pan W, Wang TT, Zhou CC, Wang SB, Wang YZ, Yang Y, Yang N, Zhou WP, Yang GS, et al. A long noncoding RNA activated by TGF-beta promotes the invasion-metastasis cascade in hepatocellular carcinoma. Cancer Cell. 2014; 25:666–681.

34. Asghar U, Witkiewicz AK, Turner NC, Knudsen ES. The history and future of targeting cyclin-dependent kinases in cancer therapy. Nature reviews Drug discovery. 2015; 14:130–146.

35. Frau M, Biasi F, Feo F, Pascale RM. Prognostic markers and putative therapeutic targets for hepatocellular carcinoma. Molecular aspects of medicine. 2010; 31:179–193.

36. Yu Z, Gao Y-Q Feng, H Lee Y-Y, Li MS, Tian Y, Go MYY, Yu D-Y Cheung Y-S, Lai PBS, Yu J,

Wong VWS, Sung JJY, Chan HLY, Cheng ASL. Cell cycle-related kinase mediates viral-host signalling to promote hepatitis B virus-associated hepatocarcinogenesis. Gut. 2014; 63:1793–1804.

37. Miao R, Luo H, Zhou H, Li G, Bu D, Yang X, Zhao X, Zhang H, Liu S, Zhong Y, Zou Z, Zhao Y, Yu K, He L, Sang X, Zhong S, et al. Identification of prognostic biomarkers in hepatitis B virus-related hepatocellular carcinoma and stratification by integrative multi-omics analysis. J Hepatol. 2014; 61:840–849.

38. Calvisi DF, Pinna F, Ladu S, Muroni MR, Frau M, Demartis I, Tomasi ML, Sini M, Simile MM, Seddaiu MA, Feo F, Pascale RM. The degradation of cell cycle regulators by SKP2/CKS1 ubiquitin ligase is genetically controlled in rodent liver cancer and contributes to determine the susceptibility to the disease. International Journal of Cancer. 2010; 126:1275–1281.

39. Calvisi DF, Ladu S, Pinna F, Frau M, Tomasi ML, Sini M, Simile MM, Bonelli P, Muroni MR, Seddaiu MA, Lim DS, Feo F, Pascale RM. SKP2 and CKS1 promote degradation of cell cycle regulators and are associated with hepatocellular carcinoma prognosis. Gastroenterology. 2009; 137:1816–1826:e1811–1810.

40. Liao Q, Xiao H, Bu D, Xie C, Miao R, Luo H, Zhao G, Yu K, Zhao H, Skogerbo G, Chen R, Wu Z, Liu C, Zhao Y. ncFANs: a web server for functional annotation of long non-coding RNAs. Nucleic Acids Res. 2011; 39:W118–124.

41. Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Research. 2009; 37:1–13.

42. Merico D, Isserlin R, Stueker O, Emili A, Bader GD. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. PLoS One. 2010; 5:e13984.

43. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. Genome Research. 2003; 13:2498–2504.