Am J Cancer Res 2017;7(3):688-699
www.ajcr.us /ISSN:2156-6976/ajcr0047172

Original Article

Integrated mRNA and lncRNA expression profiling for exploring metastatic biomarkers of human intrahepatic cholangiocarcinoma

Lisheng Lv[1,2*], Miaoyan Wei[1*], Peiyi Lin[3], Zhisheng Chen[4], Peng Gong[5], Zhiwei Quan[1], Zhaohui Tang[1]

[1]Department of General Surgery, Xinhua Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China; [2]Department of General Surgery, Shengli Oilfield Central Hospital, Binzhou Medical University, Binzhou, Shandong, China; [3]Department of General Surgery, The Second Hospital, Xiamen University, Xiamen, China; [4]Department of General Surgery, Qingpu Hospital, Fudan University Shanghai Medical College, Shanghai, China; [5]Department of Hepatobiliary Surgery, First Affiliated Hospital of Dalian Medical University, Dalian Medical University, Dalian, China. [*]Equal contributors.

Received December 22, 2016; Accepted December 26, 2016; Epub March 1, 2017; Published March 15, 2017

Abstract: Long noncoding RNAs (lncRNAs) is crucial for various human cancers, but the function and mechanism of lncRNAs is largely unknown in human intrahepatic cholangiocarcinoma (ICC), the second most common liver cancer. In this study, we performed transcriptomic profiling of ICC and normal tissues, and found 2148 lncRNAs and 474 mRNAs were significantly upregulated, whereas 568 lncRNAs and 409 mRNAs were downregulated in ICC tissues. Enrichment analysis suggests these differentially expressed genes mainly focus on response to stimulus, development, and cell proliferation. Further, potential lncRNAs involved in five signaling pathways (ERBB, JAK/STAT, MAPK, VEGF and WNT) were constructed by highly co-expressed with mRNAs in these signaling pathways. The differentially expressed lncRNA-mRNA co-regulated signaling pathways in ICC were further confirmed by lncRNA target prediction. Finally, the differentially expressed lncRNAs were confirmed by quantitative real-time PCR in 32 paired ICC and adjacent tissues. The correlation analysis between the expression levels of lncRNAs and clinicopathologic characteristics showed that EMP1-008, ATF3-008, and RCOR3-013 were observed significantly downregulated in ICC with tumor metastasis. These findings suggested that lncRNA expression profiling in ICC is profoundly different from that in noncancerous tissues, and lncRNA may be used as a potential diagnostic and prognostic biomarker for ICC metastasis.

Keywords: lncRNA, microarray, biomarkers, intrahepatic cholangiocarcinoma, metastasis

Introduction

Intrahepatic cholangiocarcinoma (ICC), the second most common liver cancer following hepatocellular carcinoma (HCC), accounts for 5%-10% of all primary liver cancers [1, 2]. Recently, the morbidity and mortality of ICC patients have greatly increased worldwide [3, 4]. Since highly molecular heterogeneous, it has been reported as poor prognosis and limited therapeutic options for ICC patients. One third of ICC tumors are amenable to surgical resection, but most cases are diagnosed at advanced stages and chemotherapy as the only established standard of practice [2]. Currently, no molecular therapies are available for the treatment of this neoplasm. Poor understanding of ICC and lack of known oncogenic addiction loops has hindered the development of effective targeting therapies. Diagnostic methods for ICC need to be improved, especially for biomarkers. The carbohydrate antigen CA-199 and CA-125 as serum tumor markers are frequently used [5], however as the low sensitivity and specificity, they are not suitable for monitoring disease progression. As compared with noncancerous biliary epithelial cells, Obama et al. found the different gene expression pattern in ICC [6], whereas no molecular target was obtained for ICC diagnosis. Thus, exploring novel biomarkers especially epigenetic markers such as noncoding RNAs are urgently needed for early diagnosis of human ICC.

Long noncoding RNAs (lncRNAs) are a subgroup of noncoding RNA transcripts that are larger than 200 nt and have no protein-coding capacity [7]. Unlike microRNAs, the length of lncRNAs

allows them to fold into intricate structures, and function as RNA sequences by themselves through secondary and tertiary structural determinants. The intrinsic ability of lncRNAs to interact with DNA, RNA and proteins by acting as guides, tethers, decoys and scaffolds offers the most compelling explanation for regulating gene expression, including epigenetic transcriptional control as associated with chromatin remodeling complexes [8], splicing [9], translation [10] and protein stability. Generally, lncRNA transcripts exhibit low-level but tissue-specific expression and poorly conserved [11]. Recently, several studies have shown the importance of lncRNAs to normal physiology as well as to gene expression regulation, by modulating key cellular processes such as cell proliferation, senescence, migration and apoptosis [12]. The critical role of a large number of lncRNAs have been observed in a variety of biological processes, even in stem cells [13, 14]. Moreover, increasingly experimental evidences suggest the connections between lncRNA and microRNAs. A new function for lncRNAs has become apparent from their ability to regulate microRNA activity by acting as either competitive endogenous RNAs or sponges for microRNAs [15].

Furthermore, it has been found that aberrant lncRNA expression in various types of cancers [16-18] may be associated with metastasis and disease prognosis [19]. The expression of specific lncRNAs with oncogenic features is closely linked to the capability of promoting matrix invasion of cancer cells and tumor growth [20], suggesting lncRNAs have the potential to serve as diagnostic markers and therapeutic targets for cancers. Emerging studies have revealed that Malat1 is highly expressed in many malignant diseases including liver cancer, and contributes to enhancing cell migration or facilitating proliferation [21]. Furthermore, lncRNA metallothionein 1D (MT1DP) was found as a tumor suppressor. Overexpression of MT1DP resulted in reduced cell proliferation and colony formation in soft agar, and increased apoptosis in liver cancer cells [22]. Another study reported one lncRNA, highly up-regulated in liver cancer (HULC), promotes angiogenesis in liver cancer through miR-107/E2F1/SPHK1 signaling [23]. Recently, a novel lncRNA, down-regulated in liver cancer stem cells (lnc-DILC) was identified as mediating the crosstalk between TNF-α/NF-κB signaling and autocrine IL-6/STAT3 cascade, and connecting hepatic inflammation with LCSC (liver cancer stem cells) expansion, which suggested that lnc-DILC is not only a potential prognostic biomarker, but also a possible therapeutic target against LCSCs [24].

In ICC, carbamoyl-phosphate synthase 1 (CPS1) and its lncRNA CPS1 intronic transcript 1 (CPS1-IT1) were observed to be upregulated and served important roles in ICC by promoting the proliferation of ICC cells. Furthermore, CPS1 and CPS1-IT1 were associated with poor liver function and reduced survival rates, and may be potential prognostic indicators for patients with ICC [25]. Recently, it has been reported that the lncRNA expression profiling in ICC tissues is profoundly different from that in noncancerous tissues, suggesting lncRNA as a potential diagnostic and prognostic biomarker for ICC [26].

In this study, we performed profiling of lncRNA and mRNA through microarray analysis to determine their association with molecular characteristics of ICC especially for metastasis. Results showed that 2716 of lncRNAs and 883 of mRNAs were differentially expressed in ICC tissues as compared with paired noncancerous tissues. Through correlation analysis and lncRNA target prediction, the co-regulated signaling network of lncRNAs and mRNA were constructed. The correlation between clinical characteristics and expression level of lncRNAs was clearly confirmed in more ICC samples. These observations suggest that a larger number of lncRNAs may contribute to the development and progression of ICC especially for metastasis, integrated analysis of lncRNA-mRNA profiles might provide new biomarkers and targets for ICC diagnosis and treatment.

Methods

Tissue samples of ICC patients

There were thirty-two of ICC patients and four of normal patients were included in the present study in Xinhua Hospital Affiliated to Shanghai Jiao Tong University (Shanghai, China). We collected and immediately frozen the tissues in liquid nitrogen. The paired noncancerous tissues were > 2 cm distant from the tumor.

This study was approved by the ethics committee of Xinhua Hospital Affiliated to Shanghai

Table 1. The information of 4 ICC sample and 4 normal samples used in microarray

ID	Gender	Age	Diagnosis	CA199 (U/ml)	AFP (IU/ml)	CEA (ng/ml)	CA125 (U/ml)	CA153 (U/ml)	CA724 (U/ml)	CA242 (U/ml)
T475	M	43	ICC	76.9	3.2	1.6	10.2	19.41	2.4	29.9
T553	M	62	ICC	40.2	1.5	3	24.1	11.35	1.4	18.2
T816	M	48	ICC	66.3	3.3	8	14.5	13.15	3.1	17
T012	F	51	ICC	355.6	26.9	1.3	31.6	26.73	22.6	88.01
C913	F	49	Normal	2039	1.37	0.69	50.51	9.85	1.04	13.19
C001	M	41	Normal	92.94	0	0	0	0	0	0
C339	F	58	Normal	25.6	1.25	2.27	2.37	8.15	184.6	107.82
C910	M	42	Normal	9.73	3.89	0	0	0	0	0

Table 2. The information and primers for 4 lncRNAs used in qRT-PCR

ncRNA_Accession	NONHSAT 027037	NONHSAT 009381	NONHSAT 009284	NONHSAT 009859
Chromosome	Chr.12	Chr.1	Chr.1	Chr.1
Start	13349659	212786777	211477352	226044404
End	13368329	212794119	211486367	226047241
Strand	+	+	+	-
ENSEMBL	ENST 00000542289	ENST 00000492118	ENST 00000528066	ENST 00000487971
Gene Symbol	EMP1-008	ATF3-008	RCOR3-013	TMEM63A-005
Sig.	Down	Down	Up	Up
Forward primer	GCTGGGACCCTTCAGAACTC	GAGACTTGCCAGAGGGCTTC	GCGCTTTGGAAAATGTCTGTCT	TTCTGACCCACCTAAGCCAC
Reverse primer	GGCTTCCAGGCTTAGCGTAT	TGGCCCCAATTTGGAAGAGT	CTGCTCCTCTGTGGTCCAAC	TGATCGGGTTATTCAGCGCA

Jiao Tong University, and all participants were informed of the requirements and provided written consent. The study procedures were carried out in accordance with the approved programs.

LncRNA and mRNA profiling

The lncRNA and mRNA expression profiling was carried out with 4 of ICC tissues and 4 of normal tissues using lncRNA and mRNA microarray (Agilent Human lncRNA 4*180 K, Design ID: 062918), which included 78,243 probes for lncRNAs and 32,776 probes for coding transcripts. The lncRNA probes on microarray were designed according to the known lncRNAs from BroadlncRNA, RefSeq, ENSEMBL, lncRNAdb, NONCODE (V4) and frnadb (V3.4). Total RNA was quantified by the NanoDrop ND-2000 (Thermo Scientific) and the RNA integrity was assessed using Agilent Bioanalyzer 2100 (Agilent Technologies). The sample labeling, microarray hybridization and washing were performed based on the manufacturer's standard protocols. Then, the arrays were scanned by the Agilent Scanner G2505C (Agilent Technologies). Feature Extraction software (version10.7.1.1, Agilent Technologies) was used to analyze array images to get raw data. Genespring were employed to finish the basic analysis with the raw data. To begin with, the raw data was normalized with the quantile algorithm. The probes that have 75% flags in "P" at both conditions were chosen for further data analysis. Differentially expressed genes or lncRNAs were then identified as a fold change ≥ 2.0 and a *P* value < 0.05.

QRT-PCR analyses

Total RNA was extracted from 32 pairs of snap-frozen ICC tissues and matched noncancerous tissues using TRIzol reagent (Invitrogen, California) according to the manufacturer's protocol. The first strand cDNA synthesis was synthesized by M-MLV Reverse Transcriptase (Takara Biotechnology Co., Ltd., Dalian, China) with 1 μg total RNA according to the manufacturer's instructions. The qRT-PCR included 40 cycles of amplification was performed with SYBR® Premix Ex Taq™ II (Takara Biotechnology Co., Ltd., Dalian, China) on Mx3000P QPCR System (Agilent Technologies, Inc., USA). Expression of target genes ($2^{-\Delta\Delta Ct}$) was normalized

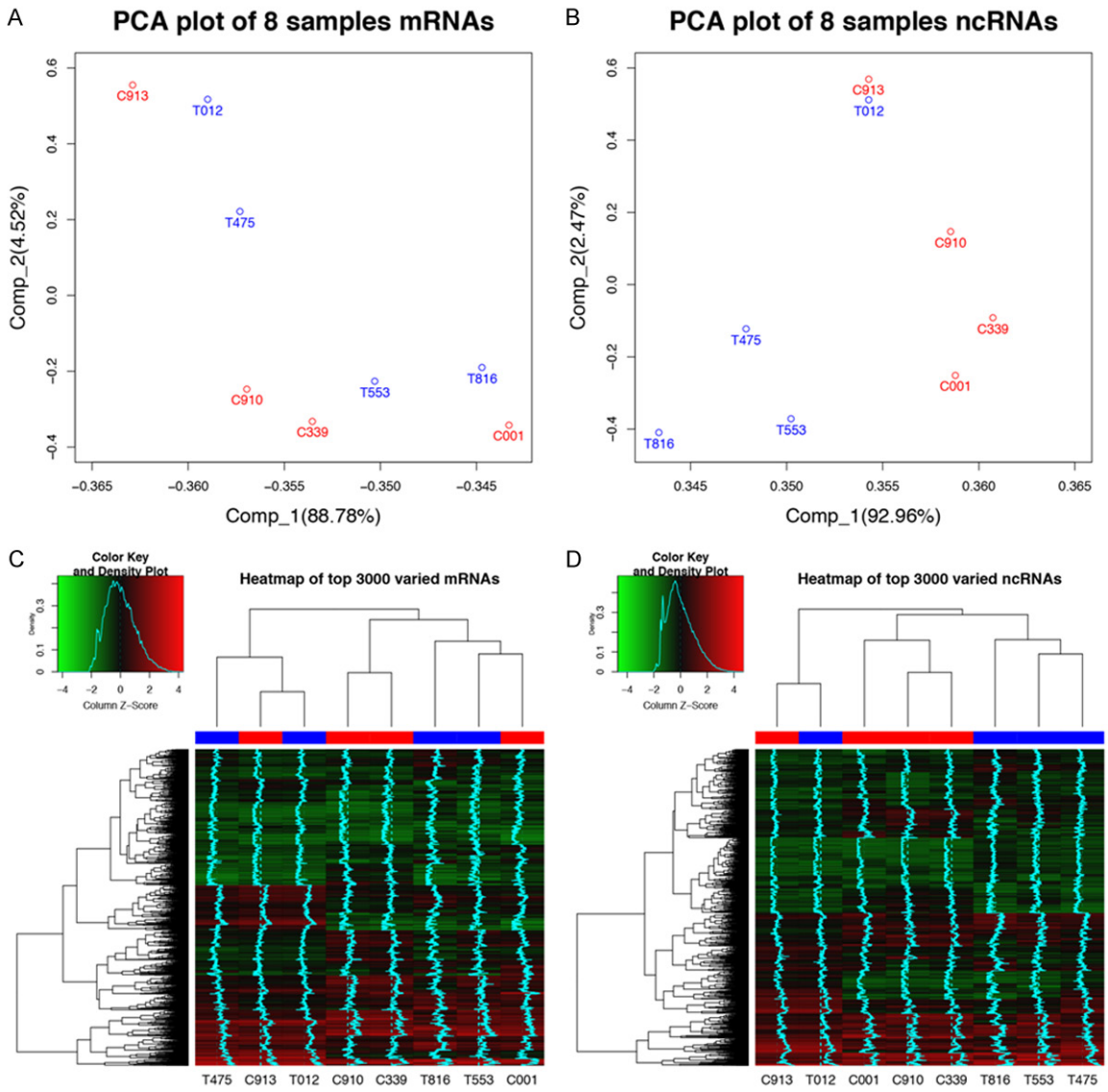

Figure 1. Relationship of the transcriptomes for 4 ICC sample (T) and 4 normal samples (C). A. The PCA plot explained by first principle component and second principle component by using profile of all mRNAs expression after filtering low signal. B. The PCA plot explained by first principle component and second principle component by using profile of all lncRNAs expression after filtering low signal. C. The heatmap plot by using top 3000 varied mRNAs. D. The heatmap plot by using top 3000 varied lncRNAs.

against GAPDH [27]. The primer sequences used are listed in **Table 2**. QRT-PCR was performed in 32 pairs of samples (tumor and adjacent tissues).

Enrichment analyses

The enrichment analyses of GO and KEGG pathway for differentially expressed gene (DEG) sets were completed by DAVID web servers [28] and Cytoscape app [29]. We used the BiNGO app to exhibit GO enrichment accompanied with GO hierarchy [30], and presented results of KEGG pathway analysis as a rich factor plot. To extend into the whole transcriptome, the GSEA analysis was used [31].

Regulatory network construction for lncRNA-mRNAs

To investigate the relationship between lncRNAs and mRNAs, we constructed the lncRNA-mRNA co-regulated network using differentially expressed lncRNAs and mRNAs. Pearson cor-

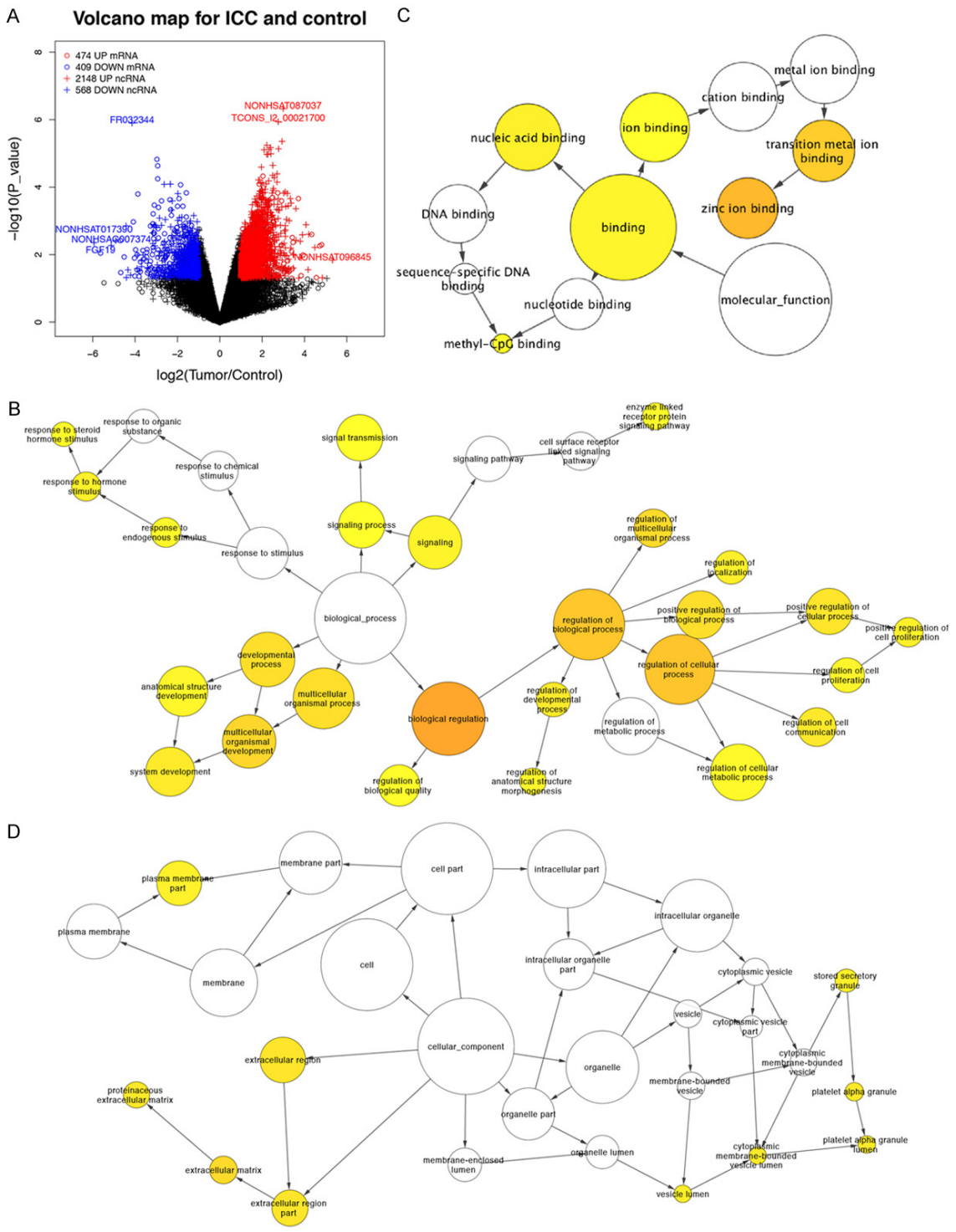

Figure 2. The DEGs between ICC samples and normal samples and gene enrichment analysis of DEG mRNAs. A. The volcano plot for DEGs between ICC sample and normal sample, x-axes index the log foldchange and y-axes index the -log (*P*_value), circle node stand for mRNA and cross node stand for lncRNA. B. The BiNGO plot (cytoscape app) in GO_BP for DEG mRNAs. C. The BiNGO plot (cytoscape app) in GO_MF for DEG mRNAs. D. The BiNGO plot (cytoscape app) in GO_CC for DEG mRNAs.

relation coefficients (PCCs) were calculated to explore the correlations between the expression levels of lncRNAs and mRNAs with $abs_{(PCC)}$ > 0.95 were considered meaningful. LncTAR, a bioinformatics method, was used to filter relationship between lncRNA and targets. Finally,

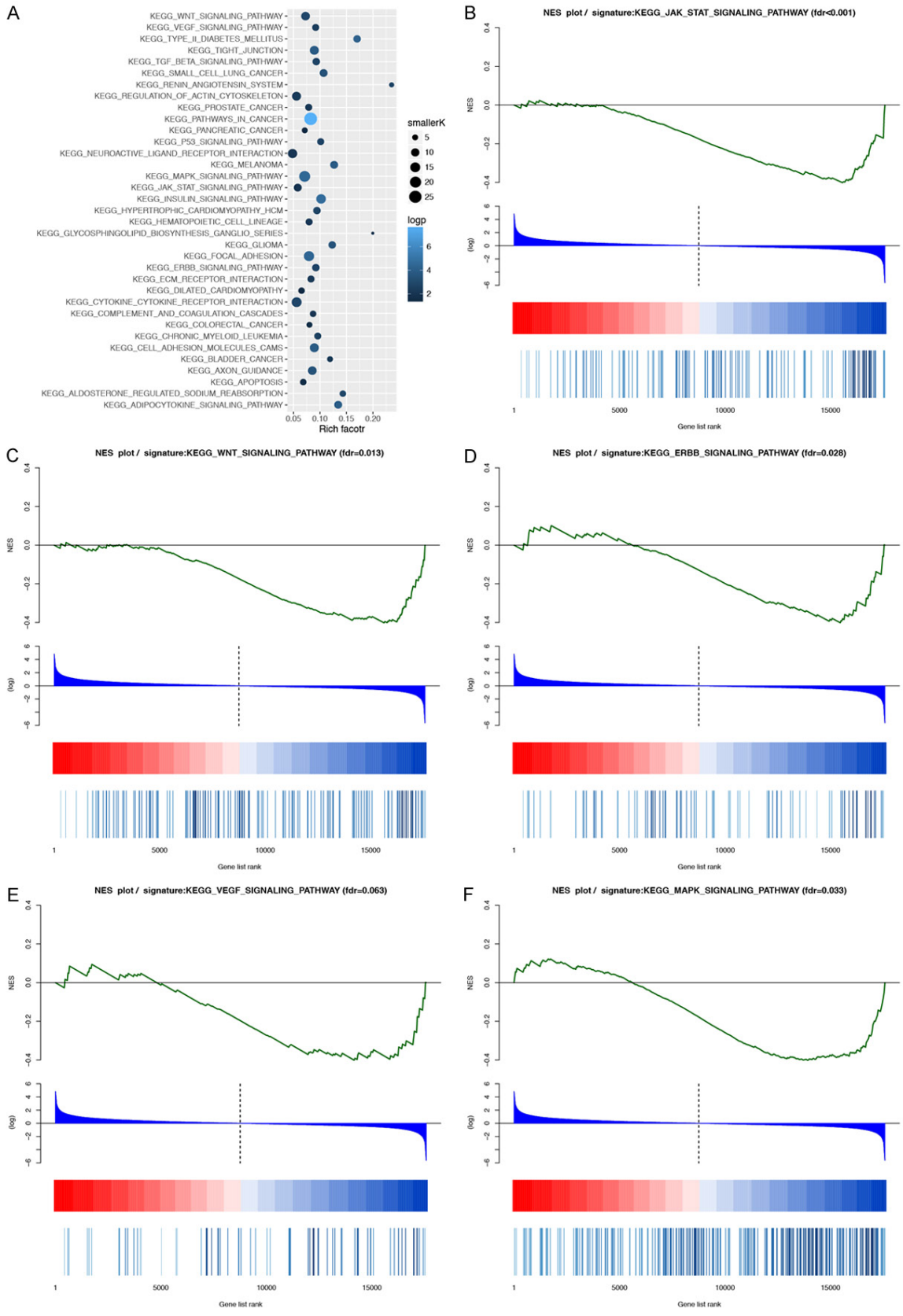

Figure 3. KEGG pathway enrichment analysis of DEG mRNAs and GSEA plot for signaling pathway. A. The rich factor plot of KEGG pathway enrichment analysis result. (The degree of color stands for the *P* value; size of node stands

for the gene count in this item). B. GSEA plot for KEGG JAK/STAT signaling pathway in $\log_2$ fold change for whole transcriptome. C. GSEA plot for KEGG WNT signaling pathway in $\log_2$ fold change for whole transcriptome. D. GSEA plot for KEGG ERBB signaling pathway in $\log_2$ fold change for whole transcriptome. E. GSEA plot for KEGG VEGF signaling pathway in $\log_2$ fold change for whole transcriptome. F. GSEA plot for KEGG MAPK signaling pathway in $\log_2$ fold change for whole transcriptome.

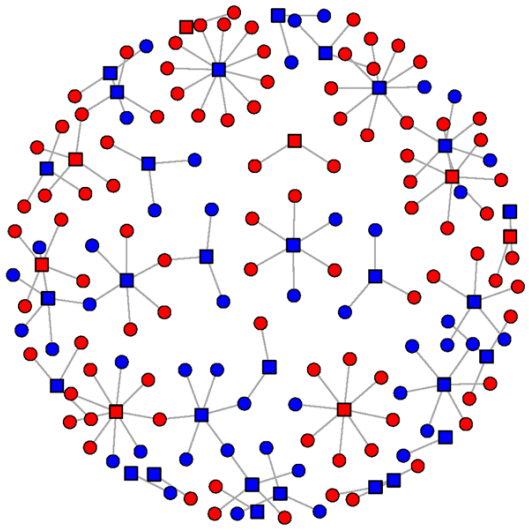

Figure 4. ICC special lncRNA-mRNA regulator network in signaling pathway. In the network, a red node represents up-regulated genes, and a blue node represents down-regulated genes. The square node represents mRNA, while a circle node represents lncRNA.

the co-regulated network was constructed using the package igraph in R, while red and blue nodes represent up-regulated and down-regulated genes, respectively. The square node represents mRNA, while a circle node represents lncRNA.

Results

Expression profiling of lncRNAs and mRNA

The expression patterns including mRNA and lncRNAs between 4 ICC tumors (T) and 4 normal tissues (C) were obtained by Agilent Human lncRNA 4*180 K microarray. The raw data contains 78,243 probes for lncRNAs and 32,776 probes for coding transcripts, respectively. After filtering the low signal level probes (less than 75% flags of "P" in both groups), only 27,532 and 51,695 probes were remain to stand for expression level of 17,575 mRNAs and 51,674 lncRNAs. The most varied probe with largest standard deviation among different probes for the same gene was collected to stand for expression level of given gene.

The transcriptomic patterns of 8 samples as presented with PCA plots of lncRNA and mRNA respectively (**Figure 1A** and **1B**), showed that the expression level of both mRNA and lncRNA can classify the ICC and normal samples except for C913 and T012. Interestingly, the lncRNA PCA plot can better grouped the samples of ICC and normal, suggesting that lncRNA may have more potential for predicting ICC. Then, the top 3000 varied genes from mRNA and lncRNA were chose to plot heatmaps, respectively. Results showed that heatmaps by lncRNAs can clearly cluster samples into two groups expect for sample C913 and T012, while the heatmaps by mRNA failed (**Figure 1C** and **1D**). These findings suggested that more underlying information of lncRNAs than mRNA to be associated with the pathogenesis of ICC. Sample T012 is a female ICC patient of 49 years old, while C913 is female hepatolithiasis patient. Both of them have high level of CA199 (355.6 U/ml and 2039 U/ml), indicated a high potential to intestinal cancer (**Table 1**). These might partially explain the similarity in the expression profiling of C913 and T012 whereas not similar to other ICC samples.

Differentially expressed genes (DEG)

The t test for two independent samples were used to detect the differentially expressed genes (fold change ≥ 2; P value < 0.05), and the volcano plot showed the pattern of differentially expressed lncRNAs and mRNA (**Figure 2A**). As compared with normal tissues, in ICC there were 2148 and 568 lncRNAs significantly upregulated and downregulated respectively, while 474 and 409 mRNAs significantly upregulated and downregulated respectively (Table S1). The most significantly down-regulated coding and non-coding genes were FGF19 with $\log_2$ (fold change of T/C) = -5.63 and NONHSAT017390 (Processed transcript of MUC5) with $\log_2$ (fold change of T/C) = -5.93, while the most up-regulated coding and non-coding genes were VTCN1 with $\log_2$ (fold change of T/C) = 4.82 and NONHSAT096845 (Retained intron of ALB) with $\log_2$ (fold change of T/C) = 5.32.

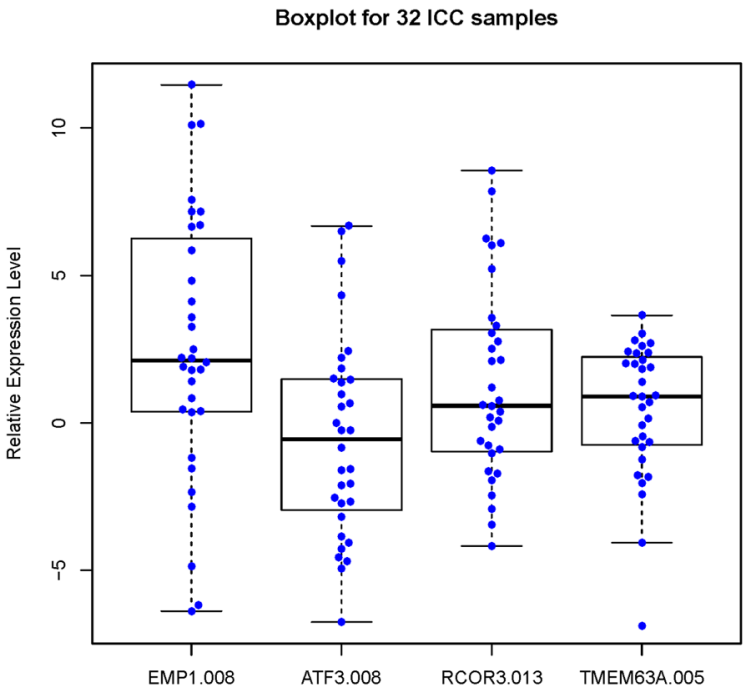

Figure 5. Boxplot for 4 lncRNAs in 32 ICC samples.

Enrichment analysis of DEG mRNAs

Gene ontology analysis was performed to assess gene and gene product enrichment in terms of biological processes, cellular components, and molecular functions by BINGO app in cytoscape [30]. The upregulated and down-regulated genes were involved in a variety of processes, including response to stimulus, system development, signaling and cell proliferation (**Figure 2B**), which tends to be located at plasma membrane part, proteinaceous extracellular matrix and platelet alpha granule lumen (**Figure 2C**). These DEG mRNAs tend to functions as zinc ion binding, growth factor activity, sequence-specific DNA binding, oxygen transporter activity and protein kinase regulator activity (**Figure 2D**).

KEGG pathway enrichment analysis completed by DAVID web server [28] indicated that the DEGs tend to be involved in various cancers (melanoma, small cell lung cancer, pancreatic, colorectal, prostate and bladder), several signaling pathways (insulin, MAPK, adipocytokine, WNT, TGF-β, ERBB, P53, VEGF and JAK/STAT), and adhesion pathway (**Figure 3A**). To expand DEGs into whole transcriptomes, the GSEA analysis was used to test the tendency of all genes in a given pathway for expression profiles [32]. The GSEA plot showed that ICC samples tend to express less in ERBB, JAK/STAT, MAPK, VEGF and WNT signaling pathway, by compared with normal tissues (**Figure 3B-F**).

ICC special lncRNA-mRNA regulator network in signaling pathway

We used 44 DEG mRNAs in five signaling pathway (ERBB, JAK/STAT, MAPK, VEGF and WNT) as a bait to find the regulatory lncRNAs. Firstly, only DEG lncRNAs were considered. Pearson correlation coefficients (PCCs) were calculated to find these co-expression lncRNAs with targets mRNAs by defined of abs (PCC) > 0.95, which found 710 lncRNAs and 993 lncRNA-mRNAs pairs in co-expression networks. We collected the sequence information of mRNAs and lncRNAs from databases of BroadlncRNA, RefSeq, ENSEMBL, lncRNAdb, NONCODE (V4) and frnadb (V3.4). The bioinformatics method of LncTAR [33] were used to filter relationship between lncRNA and targets with parameter of ndG < -10. The final special regulator-signaling network mediated by lncRNAs in ICC contains 37 mRNAs, 135 lncRNAs and 142 lncRNAs-mRNAs pairs, which was shown by igraph, a package of R (**Figure 4**). All regulating pairs of lncRNA and mRNAs were stored in Table S2.

Quantitative real-time PCR (QRT-PCR) validation

We next focused on extending the microarray results into 32 pairs of ICC samples and adjacent tissues to validate the expression pattern in ICC. The criteria of candidate lncRNAs included that 1) with annotation in Encode database (V4), 2) with high expression level (more than 5, easy to be detected), 3) with fold change of expression level more than 4, and 4) with highly significant difference (P < 0.01). Two up-regulated lncRNAs (NONHSAT009284: RCOR3-013 and NONHSAT009859: TMEM63A-005) and two down-regulated lncRNAs (NONHSAT027-

Table 3. The association between relative expression level of 4 lncRNAs and clinicopathologic characteristics

Variable	N	EMP1-008			ATF3-008			RCOR3-013			TMEM63A-005		
		Mean	SD	95% CI *P*. value	Mean	SD	95% CI *P*. value	Mean	SD	95% CI *P*. value	Mean	SD	95% CI *P*. value
Age (y)													
< 60	19	2.63	4.51	(-3.14, 3.60)	-0.09	3.54	(-1.43, 3.58)	1.34	3.57	(-2.33, 2.58)	0.47	1.99	(-1.653, 1.777)
≥ 60	13	2.4	4.68	0.89	-1.17	3.19	0.388	1.22	2.98	0.917	0.41	2.77	0.941
Gender													
F	15	1.82	4.85	(-4.62, 1.94)	-0.02	3.72	(-1.51, 3.43)	0.71	3.4	(-3.47, 1.30)	0.35	1.77	(-1.869, 1.504)
M	17	3.16	4.23	0.412	-0.98	3.12	0.431	1.8	3.21	0.359	0.53	2.73	0.826
Adbominal_pain													
0	18	2.61	4.33	(-3.16, 3.51)	-0.82	3.03	(-3.17, 1.83)	0.83	2.99	(-3.46, 1.34)	0	2.5	(-2.672, 0.638)
1	14	2.44	4.89	0.917	-0.15	3.89	0.586	1.89	3.66	0.374	1.02	1.93	0.219
Cirrhosis													
0	26	2.71	4.69	(-3.28, 5.17)	-0.87	3.55	(-4.94, 1.29)	1.02	3.5	(-4.49, 1.60)	0.81	1.91	(-0.079, 3.984)
1	6	1.77	3.9	0.652	0.95	2.23	0.241	2.47	1.95	0.34	-1.14	3.27	0.059
Hepatitis													
0	18	2.52	4.48	(-3.37, 3.30)	-0.94	3.36	(-3.42, 1.55)	1	3.11	(-3.09, 1.75)	0.49	2.14	(-1.597, 1.798)
1	14	2.55	4.72	0.983	0	3.49	0.448	1.67	3.59	0.575	0.39	2.56	0.905
CA199 (U/ml)													
0 (< 40)	14	2.75	4.31	(-2.96, 3.71)	0.16	3.07	(-1.25, 3.69)	1.36	3.39	(-2.32, 2.55)	-0.23	2.31	(-2.848, 0.426)
1 (≥ 40)	18	2.37	4.78	0.819	-1.06	3.62	0.321	1.24	3.31	0.923	0.98	2.2	0.141
AFP (IU/ml)													
0 (< 20)	26	2.74	4.6	(-3.13, 5.31)	-0.32	3.66	(-2.06, 4.27)	1.58	3.39	(-1.53, 4.55)	0.78	1.69	(-0.31, 3.80)
1 (≥ 20)	6	1.65	4.37	0.603	-1.42	1.78	0.482	0.07	2.73	0.319	-0.97	3.93	0.094
CEA (ng/ml)													
0 (< 5)	27	2.11	4.67	(-7.16, 1.73)	-0.75	3.25	(-4.79, 1.99)	1.18	3.4	(-4.03, 2.60)	0.32	2.39	(-3.11, 1.49)
1 (≥ 5)	5	4.82	2.86	0.222	0.65	4.33	0.407	1.89	2.9	0.664	1.13	1.71	0.48
CA125 (U/ml)													
0 (< 35)	22	2.9	4.04	(-2.38, 4.70)	-0.73	2.98	(-3.31, 2.04)	1.19	2.97	(-2.92, 2.29)	0.36	2.42	(-2.086, 1.543)
1 (≥ 35)	10	1.74	5.56	0.509	-0.09	4.32	0.633	1.51	4.08	0.807	0.64	2.1	0.762
CA153 (U/ml)													
0 (< 25)	26	2.59	4.2	(-3.92, 4.55)	-0.2	3.17	(-1.37, 4.87)	1.01	2.91	(-4.54, 1.54)	0.68	1.88	(-0.86, 3.35)
1 (≥ 25)	6	2.28	6.13	0.88	-1.95	4.27	0.261	2.51	4.77	0.322	-0.56	3.67	0.236
CA724 (U/ml)													
0 (< 10)	27	2.43	4.5	(-5.22, 3.88)	-0.73	3.34	(-4.66, 2.14)	1.12	3.38	(-4.38, 2.22)	0.71	1.89	(-0.59, 3.88)
1 (≥ 10)	5	3.1	5.03	0.765	0.54	3.85	0.454	2.2	2.92	0.509	-0.94	3.84	0.144
CA242 (U/ml)													
0 (< 20)	19	3	3.98	(-2.18, 4.50)	0.07	3.3	(-1.00, 3.95)	1.5	3.12	(-1.94, 2.96)	0.4	2.27	(-1.840, 1.589)
1 (≥ 20)	13	1.85	5.28	0.485	-1.41	3.47	0.233	0.99	3.63	0.673	0.52	2.42	0.882
Tumor_size (cm)													
0 (< 5)	16	3.32	4.37	(-1.69, 4.83)	0.1	2.99	(-1.19, 3.70)	1.09	2.69	(-2.81, 2.01)	0.9	1.62	(-0.743, 2.557)
1 (> 5)	16	1.75	4.65	0.332	-1.16	3.74	0.304	1.49	3.88	0.739	0	2.8	0.271
Tumor_metastasis													
0	21	3.95	4.15	(0.99, 7.24)	0.79	3.13	(1.63, 6.03)	2.19	3.14	(0.28, 4.97)	0.43	2.39	(-1.823, 1.724)
1	11	-0.17	4.05	0.012	-3.04	2.35	0.001	-0.43	2.98	0.03	0.48	2.22	0.955
Tumor_number													
Single	21	2.48	4.44	(-3.65, 3.32)	-0.6	3.24	(-2.83, 2.41)	1.51	3.38	(-1.90, 3.16)	0.24	1.99	(-2.357, 1.161)
Mulitple	11	2.64	4.86	0.925	-0.39	3.84	0.869	0.88	3.23	0.613	0.84	2.85	0.493
HBsAg													
0	24	2.59	4.28	(-3.62, 4.02)	-0.78	3.28	(-3.84, 1.86)	1.26	3.25	(-2.92, 2.66)	0.51	2.39	(-1.709, 2.178)
1	8	2.38	5.47	0.914	0.21	3.84	0.484	1.39	3.63	0.926	0.27	2.11	0.807
CK18													
0	18	1.67	4.18	(-5.23, 1.27)	-1.66	3.12	(-4.90, -0.26)	0.97	3.17	(-3.16, 1.68)	0.12	2.73	(-2.422, 0.928)
1	14	3.65	4.82	0.223	0.92	3.27	0.03	1.71	3.51	0.535	0.87	1.57	0.37
CK19													

0	25	2.29	4.62	(-5.11, 2.86)	-0.91	3.28	(-4.68, 1.20)	1.2	3.28	(-3.34, 2.49)	0.18	2.49	(-3.204, 0.768)
1	7	3.41	4.31	0.568	0.83	3.7	0.237	1.63	3.56	0.767	1.4	1.06	0.22
CD10													
0	24	2.4	4.39	(-4.33, 3.30)	-0.37	3.49	(-2.24, 3.49)	1.53	3.13	(-1.82, 3.71)	0.79	1.85	(-0.495, 3.257)
1	8	2.92	5.15	0.784	-1	3.26	0.66	0.59	3.86	0.492	-0.59	3.23	0.143
OPN													
0	17	1.91	4.89	(-4.62, 1.94)	-0.64	3.38	(-2.72, 2.27)	1.21	3.13	(-2.59, 2.24)	0.19	2.15	(-2.236, 1.114)
1	15	3.24	4.08	0.412	-0.41	3.52	0.854	1.39	3.57	0.883	0.75	2.49	0.499

037: EMP1-008 and NONHSAT009381: ATF3-008) were chose to be confirmed using qRT-PCR with primers shown in **Table 2**. We observed the same expression patterns as microarray analysis between ICC tumors and adjacent tissues for RCOR3-013 (95% CI (0.106, 2.479), P = 0.034) with statistical significance, ATF3-008 (95% CI (-1.753, 0.694), P = 0.385) and TMEM63A-005 (95% CI (-0.379, 1.276), P = 0.278) without statistical significance. EMP1-008 was unexpectedly up-regulated in ICC tumors (95% CI (0.908, 4.160), P = 0.003), which maybe due to high false positive of microarray. The expression patterns of these four lncRNAs were shown in **Figure 5**.

Correlation analysis of lncRNAs with clinicopathologic characteristics

We further analyzed the associations between the expression levels of lncRNA and various specific clinicopathologic characteristics of ICC patients. The -ΔΔCT values were used to stand for the relative expression level of lncRNAs (**Table 3**). Interestingly, ICC samples with tumor metastasis tend to expressed in lower levels of EMP1-008 (95% CI (0.99, 7.24), P = 0.012), ATF3-008 (95% CI (1.63, 6.03), P = 0.001) and RCOR3-013 (95% CI (0.28, 4.97), P = 0.030), while no significant association with size and number of tumors and markers was found. Among the several protein levels of immunohistochemical experiments (HBsAg, CK18, CK19, CD10 and OPN), only the high expression level of ATF3-008 were associated with CK18 positive samples (95% CI (-4.90, -0.26), P = 0.03). However, there were no significant correlation for the four lncRNAs and gender, age groups and biomarker of tumors (CA199, AFP, CEA, CA125, CA153, CA724, CA242), suggesting that these four lncRNAs provide independent information of these tumor biomarkers.

Discussion

As the second most common liver cancer, the morbidity and mortality of ICC patients have greatly increased worldwide [3, 4]. Exploring novel biomarkers especially epigenetic markers such as noncoding RNAs are urgently needed for early diagnosis of human ICC. Accumulating data demonstrates lncRNAs with pivotal roles in gene expression regulation and tumor progression. However, little is known about the expression patterns and functions of lncRNAs in ICC. Here, we found that lncRNA expression profile of ICC is profoundly different from noncancerous tissues, and lncRNA may be used as a potential diagnostic and prognostic biomarker for ICC metastasis.

LncRNAs, as a kind of noncoding RNA more than 200 nt [7], can interact with DNA, RNA and proteins by acting as guides, tethers, decoys and scaffolds offers the most compelling explanation for regulating gene expression [8-10]. In combination with bioinformatics and integrated analyses of differentially expressed lncRNAs and mRNAs between ICC and normal samples, we firstly constructed lncRNA-mRNA co-regulated signaling pathways in human ICC. They were predicted to be involved in a variety of processes, including response to stimulus, system development, signaling, and cell proliferation. KEGG pathway enrichment analysis indicated that the DEGs tend to be involved in ICC associated and pivotal signaling pathways such as ERBB, JAK/STAT, MAPK, VEGF and WNT signaling [34]. However, the exact co-regulatory function and mechanism of lncRNA-mRNA remains to be investigated.

Recently, it has been reported that lncRNAs can modulate key cellular processes such as cell proliferation, senescence, migration and apoptosis [12]. Furthermore, it has been found that aberrant lncRNA expression may be associated with cancer metastasis and disease prognosis [19]. In ICC, only one study found that CPS1 and its lncRNA CPS1-IT1 could promote the proliferation of ICC cells [25], suggesting lncRNA as a potential diagnostic and prognostic biomarker for ICC. Here, metastasis

tumor tended to show low expression levels of EMP1-008, ATF3-008 and RCOR3-013, providing diagnostic and prognostic potential for ICC with metastasis. Wang and his colleagues identified 290 lncRNA-mRNA pairs involving in ICC through co-expression and cis-regulation within 10 k windows. This study discovered 142 news lncRNAs-mRNAs pairs in ICC from co-expression and trans-regulation by LncTar.

In conclusion, through integrated analysis of lncRNA and mRNA profiling with molecular characteristics of ICC, we found that 2716 of lncRNAs and 883 of mRNAs were differentially expressed in ICC tissues as compared with paired noncancerous tissues. Furthermore, with correlation analysis and lncRNA target prediction, we constructed the co-regulated signaling network of lncRNAs and mRNA. These findings suggest that lncRNAs may contribute to the development and progression of ICC especially for metastasis, providing new and potential diagnostic biomarkers and treatment targets for human ICC.

Disclosure of conflict of interest

None.

Address correspondence to: Zhaohui Tang, Department of General Surgery, Xinhua Hospital, School of Medicine, Shanghai Jiao Tong University, 1665 Kongjiang Road, Shanghai 200092, China. Tel: +86-21-61114264; Fax: +86-21-61126716; E-mail: tangzhaohui001@gmail.com

References

[1] Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J and Jemal A. Global cancer statistics, 2012. CA Cancer J Clin 2015; 65: 87-108.

[2] Moeini A, Sia D, Bardeesy N, Mazzaferro V and Llovet JM. Molecular pathogenesis and targeted therapies for intrahepatic cholangiocarcinoma. Clin Cancer Res 2016; 22: 291-300.

[3] Zhang H, Yang T, Wu M and Shen F. Intrahepatic cholangiocarcinoma: epidemiology, risk factors, diagnosis and surgical management. Cancer Lett 2016; 379: 198-205.

[4] de Martel C, Plummer M and Franceschi S. Cholangiocarcinoma: descriptive epidemiology and risk factors. Gastroenterol Clin Biol 2010; 34: 173-180.

[5] Khan SA, Davidson BR, Goldin RD, Heaton N, Karani J, Pereira SP, Rosenberg WM, Tait P, Taylor-Robinson SD, Thillainayagam AV, Thomas HC, Wasan H; British Society of Gastroenterology. Guidelines for the diagnosis and treatment of cholangiocarcinoma: an update. Gut 2012; 61: 1657-1669.

[6] Obama K, Ura K, Li M, Katagiri T, Tsunoda T, Nomura A, Satoh S, Nakamura Y and Furukawa Y. Genome-wide analysis of gene expression in human intrahepatic cholangiocarcinoma. Hepatology 2005; 41: 1339-1348.

[7] Ponting CP, Oliver PL and Reik W. Evolution and functions of long noncoding RNAs. Cell 2009; 136: 629-641.

[8] Brockdorff N. Noncoding RNA and Polycomb recruitment. RNA 2013; 19: 429-442.

[9] Guo F, Li Y, Liu Y, Wang J, Li Y and Li G. Inhibition of metastasis-associated lung adenocarcinoma transcript 1 in CaSki human cervical cancer cells suppresses cell proliferation and invasion. Acta Biochim Biophys Sin (Shanghai) 2010; 42: 224-229.

[10] Gong C and Maquat LE. lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. Nature 2011; 470: 284-288.

[11] Diederichs S. The four dimensions of noncoding RNA conservation. Trends Genet 2014; 30: 121-123.

[12] Feng S, Yao J, Chen Y, Geng P, Zhang H, Ma X, Zhao J and Yu X. Expression and functional role of reprogramming-related long noncoding RNA (lincRNA-ROR) in glioma. J Mol Neurosci 2015; 56: 623-630.

[13] Ouyang J, Hu J and Chen JL. lncRNAs regulate the innate immune response to viral infection. Wiley Interdiscip Rev RNA 2016; 7: 129-143.

[14] Hu S and Shan G. LncRNAs in stem cells. Stem Cells Int 2016; 2016: 2681925.

[15] Johnsson P, Ackley A, Vidarsdottir L, Lui WO, Corcoran M, Grander D and Morris KV. A pseudogene long-noncoding-RNA network regulates PTEN transcription and translation in human cells. Nat Struct Mol Biol 2013; 20: 440-446.

[16] Liz J and Esteller M. lncRNAs and microRNAs with a role in cancer development. Biochim Biophys Acta 2016; 1859: 169-176.

[17] Tang WW, Wu Q, Li SQ, Tong YS, Liu ZH, Yang TX, Xu Y and Cao XF. Implication of lncRNAs in pathogenesis of esophageal cancer. Onco Targets Ther 2015; 8: 3219-3226.

[18] Huarte M. The emerging role of lncRNAs in cancer. Nat Med 2015; 21: 1253-1261.

[19] Han L, Zhang EB, Yin DD, Kong R, Xu TP, Chen WM, Xia R, Shu YQ and De W. Low expression of long noncoding RNA PANDAR predicts a poor prognosis of non-small cell lung cancer and affects cell apoptosis by regulating Bcl-2. Cell Death Dis 2015; 6: e1665.

[20] Qiu MT, Hu JW, Yin R and Xu L. Long noncoding RNA: an emerging paradigm of cancer research. Tumour Biol 2013; 34: 613-620.

[21] Wang J, Wang H, Zhang Y, Zhen N, Zhang L, Qiao Y, Weng W, Liu X, Ma L, Xiao W, Yu W, Chu Q, Pan Q and Sun F. Mutual inhibition between YAP and SRSF1 maintains long non-coding RNA, Malat1-induced tumourigenesis in liver cancer. Cell Signal 2014; 26: 1048-1059.

[22] Yu W, Qiao Y, Tang X, Ma L, Wang Y, Zhang X, Weng W, Pan Q, Yu Y, Sun F and Wang J. Tumor suppressor long non-coding RNA, MT1DP is negatively regulated by YAP and Runx2 to inhibit FoxA1 in liver cancer cells. Cell Signal 2014; 26: 2961-2968.

[23] Lu Z, Xiao Z, Liu F, Cui M, Li W, Yang Z, Li J, Ye L and Zhang X. Long non-coding RNA HULC promotes tumor angiogenesis in liver cancer by up-regulating sphingosine kinase 1 (SPHK1). Oncotarget 2016; 7: 241-254.

[24] Wang X, Sun W, Shen W, Xia M, Chen C, Xiang D, Ning B, Cui X, Li H, Li X, Ding J and Wang H. Long non-coding RNA DILC represses self-renewal of liver cancer stem cells via inhibiting autocrine IL-6/STAT3 axis. J Hepatol 2016; 64: 1283-94.

[25] Ma SL, Li AJ, Hu ZY, Shang FS and Wu MC. Co-expression of the carbamoylphosphate synthase 1 gene and its long noncoding RNA correlates with poor prognosis of patients with intrahepatic cholangiocarcinoma. Mol Med Rep 2015; 12: 7915-7926.

[26] Wang J, Xie H, Ling Q, Lu D, Lv Z, Zhuang R, Liu Z, Wei X, Zhou L, Xu X and Zheng S. Coding-noncoding gene expression in intrahepatic cholangiocarcinoma. Transl Res 2016; 168: 107-121.

[27] Livak KJ and Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 2001; 25: 402-408.

[28] Huang DW, Sherman BT, Tan Q, Kir J, Liu D, Bryant D, Guo Y, Stephens R, Baseler MW, Lane HC and Lempicki RA. DAVID bioinformatics resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. Nucleic Acids Res 2007; 35: W169-175.

[29] Lotia S, Montojo J, Dong Y, Bader GD and Pico AR. Cytoscape app store. Bioinformatics 2013; 29: 1350-1351.

[30] Maere S, Heymans K and Kuiper M. BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. Bioinformatics 2005; 21: 3448-3449.

[31] Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES and Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005; 102: 15545-15550.

[32] Subramanian A, Kuehn H, Gould J, Tamayo P and Mesirov JP. GSEA-P: a desktop application for gene set enrichment analysis. Bioinformatics 2007; 23: 3251-3253.

[33] Li J, Ma W, Zeng P, Wang J, Geng B, Yang J and Cui Q. LncTar: a tool for predicting the RNA targets of long noncoding RNAs. Brief Bioinform 2015; 16: 806-812.

[34] Chong DQ and Zhu AX. The landscape of targeted therapies for cholangiocarcinoma: current status and emerging targets. Oncotarget 2016; 7: 46750-46767.