Original Paper

RNA Sequencing and Bioinformatics Analysis Implicate the Regulatory Role of a Long Noncoding RNA-mRNA Network in Hepatic Stellate Cell Activation

Can-Jie Guo^a Xiao Xiao^a Li Sheng^a Lili Chen^b Wei Zhong^a Hai Li^a Jing Hua^a Xiong Ma^a

^aDivision of Gastroenterology and Hepatology, Key Laboratory of Gastroenterology and Hepatology, Ministry of Health, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai Institute of Digestive Disease, Shanghai, ^bGucun Town Community Health Service, Baoshan District, Shanghai, China

Key Words

Lncrna • Liver fibrosis • Hepatic stellate cell

Abstract

Background/Aims: To analyze the long noncoding (lncRNA)-mRNA expression network and potential roles in rat hepatic stellate cells (HSCs) during activation. **Methods:** lncRNA expression was analyzed in quiescent and culture-activated HSCs by RNA sequencing, and differentially expressed lncRNAs verified by quantitative reverse transcription polymerase chain reaction (qRT-PCR) were subjected to bioinformatics analysis. *In vivo* analyses of differential lncRNA-mRNA expression were performed on a rat model of liver fibrosis. **Results:** We identified upregulation of 12 lncRNAs and 155 mRNAs and downregulation of 12 lncRNAs and 374 mRNAs in activated HSCs. Additionally, we identified the differential expression of upregulated lncRNAs (NONRATT012636.2, NONRATT016788.2, and NONRATT021402.2) and downregulated lncRNAs (NONRATT007863.2, NONRATT019720.2, and NONRATT024061.2) in activated HSCs relative to levels observed in quiescent HSCs, and Gene Ontology and Kyoto Encyclopedia of Genes and Genomes pathway analyses showed that changes in lncRNAs associated with HSC activation revealed 11 significantly enriched pathways according to their predicted targets. Moreover, based on the predicted co-expression network, the relative dynamic levels of NONRATT013819.2 and lysyl oxidase (*Lox*) were compared during HSC activation both *in vitro* and *in vivo*. Our results confirmed the upregulation of lncRNA NONRATT013819.2 and *Lox* mRNA associated with the extracellular matrix (ECM)-related signaling pathway in HSCs and fibrotic livers. **Conclusion:** Our results detailing a dysregulated lncRNA-mRNA network might provide new treatment strategies for hepatic fibrosis based on findings indicating potentially critical roles for NONRATT013819.2 and *Lox* in ECM remodeling during HSC activation.

© 2017 The Author(s)
Published by S. Karger AG, Basel

C.J. Guo and X. Xiao X. contributed equally to this work

Xiong Ma

Shanghai Institute of Digestive Disease,
145 Middle Shandong Road, Shanghai 2000127 (China)
Tel./Fax 008621-63200874; E-Mail maxiongmd@163.com

Introduction

Fibrosis is characterized by an excessive deposition of collagen due to an imbalance between collagen synthesis and degradation [1]. Hepatic fibrogenesis results from the accumulation of extracellular matrix (ECM), including collagen, proteoglycan, and adhesive glycoproteins, which is mainly produced by hepatic stellate cells (HSCs). HSCs are types of mesenchymal cells located in the perisinusoidal space (space of Disse) beneath the endothelial barrier [2]. Upon activation by fibrogenic stimuli, HSC morphology and physiology change significantly during myofibroblastic transdifferentiation, which is regulated by multiple cellular signal-transduction pathways and cytokines [3]. Among these, *Lox* and the *Lox*-mediated ECM-related pathway are probably the most important [4]. As the key enzyme in liver fibrosis, *Lox* controls collagen and elastin maturation, promoting ECM cross-linking and participating in the initiation and/or progression of hepatic fibrogenesis [5].

Previous studies attempted to reveal the mechanisms underlying HSC activation mainly at the mRNA level [6]; however, a comprehensive understanding of fibrogenesis remains beyond our reach due to its complexity, especially involving the intricate regulation of gene expression. Fortunately, recent progress in RNA-Seq technology focusing on long noncoding RNAs (lncRNAs) and bioinformatics analysis added new depth to our knowledge of gene regulation during the process of HSC transformation [7].

To better understand the intricate regulatory network associated with HSC activation, we performed RNA sequencing and comparative bioinformatics analysis of quiescent and activated HSCs. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses were performed to identify the functions of target mRNAs exhibiting differential expression levels correlating with changes in lncRNA expression. Additionally, we established co-expression networks and potential mRNA-targeting relationships and also verified changes in NONRATT013819.2 and lysyl oxidase (*Lox*) expression during HSC activation both *in vitro* and *in vivo*. Moreover, we discussed the possible lncRNA-mRNA interactions associated with fibrogenesis and enhanced our knowledge regarding the mechanisms associated with liver disease.

Materials and Methods

Isolation, culture, and identification of rat HSCs

HSCs were isolated from three normal male Sprague-Dawley rats (weight: 300–400 g; Shanghai Laboratory Animal Center of Chinese Academy of Sciences, Shanghai, China) by *in situ* perfusion and density gradient centrifugation [8]. All materials for HSC isolation were obtained from commercial sources as previously described [9]. The rats received humane care according to the Guide for the Care and Use of Laboratory Animals of the Chinese Academy of Sciences.

Isolated rat HSCs were cultured in Dulbecco's modified Eagle medium supplemented with 10% fetal bovine serum, 100 U/mL penicillin, 100 mg/mL streptomycin, and 2 mmol/L glutamine. Thereafter, quiescent and activated HSCs were harvested at days 2 and 14 for experimentation.

Cell purity was verified by immunohistochemical examination using monoclonal antibodies specific for desmin (1:100; Sigma-Aldrich, St. Louis, MO, USA). Immunohistochemical staining for desmin was performed as previously described [10]. Nuclei were labeled with Hoechst 33258 (Roche, Basel, Switzerland).

Establishment of an animal model of liver fibrosis and histological examination

Thirty-six Sprague-Dawley rats were divided into three groups (normal, control, and fibrosis model). Fibrosis-model rats were injected subcutaneously with 40% CCl₄ (3 mL/kg CCl₄:olive oil = 2:3) every 3 days for 10 weeks [11]. Control rats received only olive oil. Model rats were sacrificed at 4, 6, 8, or 10 weeks, respectively, and the degree of liver fibrosis was evaluated by microscopy.

Intrahepatic HSCs were isolated from CCl₄-induced fibrotic rat livers as previously described. Liver tissues were fixed with 4% formaldehyde in phosphate-buffered saline and embedded in paraffin, and

5- μ m-thick sections were prepared. All sections were stained with hematoxylin and eosin and underwent standard Van Gieson staining. Total RNA from rat HSCs was extracted, and fibrosis was graded according to the Ishak modified-staging system [12]. Histopathology was interpreted by two independent board-certified pathologists who were blind to the study.

RNA isolation, library preparation, and sequencing analysis

Total RNA was extracted from HSCs using TRIzol (Invitrogen, Carlsbad, CA, USA) and purified according to manufacturer instructions. Total RNA was quantified using a NanoDrop ND-2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA), and RNA integrity was determined using an Agilent 2100 system and an RNA 6000 Nano kit (Agilent Technologies, Santa Clara, CA, USA). Ribosomal RNA was removed using an EpicentreRibo-zero rRNA removal kit (Thermo Fisher Scientific) to enable accurate lncRNA analysis. RNA-Seq was conducted on an Illumina genome analyzer IIx from Shanghai Biotechnology Corporation (Shanghai, China) using next-generation sequencing analysis. Genome mapping was performed using TopHat software (v2.0.9; <https://ccb.jhu.edu/software/tophat/index.shtml>).

We determined differentially expressed genes and lncRNAs using edgeR analysis [13] for gene-expression profiling (<http://bioconductor.org/packages/release/bioc/html/edgeR.html>) and the Cufflinks algorithm [14] for identification of transcripts from RNA-Seq data. False-discovery rate (FDR) was used to determine the significance threshold of the *p*-value for multiple tests. A fold change ≥ 2 and an FDR ≤ 0.05 represented the thresholds signifying genes that were considered as being differentially expressed.

lncRNA expression, identification, and classification

Raw RNA-Seq samples were obtained from sequencing Poly-A+ fractions using paired-end Illumina reads (Illumina, San Diego, CA, USA). The mapped reads of each sample were assembled using the Cufflinks. Application function of the Cuffdiff algorithm (<http://cufflinks.cbcb.umd.edu/manual.html#cuffdiff>) to determine differences between samples for lncRNA analysis. Differences in lncRNA-screening criteria were determined by a *q*-value ≤ 0.05 and a fold-change in expression ≥ 2 .

Cuffcompare [15] was used to compare candidate sequences with known lncRNAs, and we used the following process to select new lncRNAs: 1) transcription length ≥ 200 bp and number of exons ≥ 2 ; 2) a predicted open reading frame length < 300 bp; 3) intersecting prediction results from the protein family database [16] (Pfam; <http://www.uniprot.org/database/DB-0073>), the coding potential calculator [17] (CPC; <http://cpc.cbi.pku.edu.cn/>), and the coding-noncoding index [18] (CNCI; <https://omictools.com/coding-non-coding-index-tool>) selected based on a CPC score < 0 , a CNCI score < 0 , and Pfam for determination of non-significant transcripts associated with potential lncRNAs; and 4) sequences matching those of known lncRNAs were removed.

According to their locations relative to the nearest protein-coding genes, the annotated lncRNAs were subdivided into intergenic lncRNAs (lincRNAs), intronic-antisense lncRNAs, exonic-sense lncRNAs, bidirectional lncRNAs, or exonic-antisense and intronic-sense lncRNAs [19].

Feature comparisons of transcripts in HSCs

To analyze the differences between lncRNAs and protein-coding genes, the distributions of lncRNA- and mRNA-transcript length and exon number were analyzed in HSCs. To compare transcript-expression levels and verify the consistency of predicted lncRNA with general characteristics, lncRNA- and mRNA-transcript levels were averaged, and a box diagram was created using $\log_{10}$ [fragments per kilobase of transcript per million (FPKM) + 1] values.

Target-gene prediction

To determine lncRNA function, we performed transcriptome analysis by lncRNA sequencing (lncRNA-Seq) to predict potential lncRNA targets. First, we predicted the cis- and trans-target genes of lncRNAs, with the trans-target genes predicted based on the rat mRNA database (<http://rgd.mcw.edu/>). Complementary or similar sequences were selected using BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>), and similar free energy values between sequences were calculated using RNAplex [20] (www.bioinf.uni-leipzig.de/Software/RNAplex/), with similarities above a specified threshold selected. For the cis-target genes, lncRNAs > 10 kb in length were selected.

GO analysis and KEGG pathway annotation

The functions of predicted lncRNA targets were studied via bioinformatics analysis. The DAVID program (<http://david.abcc.ncifcrf.gov/>) was used to perform GO analysis to identify the molecular functions of potential target genes. Additionally, KEGG (<http://www.genome.ad.jp/kegg/>) and BioCarta (<http://www.biocarta.com>) were used to analyze the potential functions of these genes in given pathways. Significant correlations between target genes and their associated functions and/or pathways were assessed based on a threshold of $p < 0.05$.

Analysis of lncRNA-mRNA regulatory networks

Co-expression-network analysis was performed to determine the correlated expression of genes in order to construct weighted co-expression networks [21]. To determine relationships between lncRNA expression and regulation of mRNA co-expression, Pearson's correlation coefficients were calculated [22], and the R value was used to compare the Pearson's correlation coefficients between lncRNAs and differentially expressed mRNAs, with relationships exhibiting coefficients ≥ 0.98 used to construct the network via Cytoscape (<http://www.cytoscape.org/>). Co-expression networks of possible lncRNA-mRNA interactions representing critical lncRNAs and target mRNAs were established accordingly.

Quantitative reverse transcription polymerase chain reaction (qRT-PCR) validation

RNAs were extracted according to manufacturer instructions, and RT reactions were performed according instructions associated with the ReverTra Ace kit (Toyobo, Osaka, Japan). The expression of selected upregulated lncRNAs (NONRATT012636.2, NONRATT016788.2, and NONRATT021402.2), downregulated lncRNAs (NONRATT007863.2, NONRATT019720.2, and NONRATT024061.2), and a lncRNA-mRNA pair (NONRATT013819.2 and *Lox*) were analyzed by qRT-PCR using a SYBR Green PCR kit (TaKaRa, Shiga, Japan). Glyceraldehyde 3-phosphate dehydrogenase (*Gapdh*) mRNA was used as an internal control. To determine relationships between lncRNA-mRNA pairs (NONRATT013819.2 and *Lox*), we also calculated correlation coefficients (using Pearson's correlation). For quantitative results, expression of each lncRNA was represented as a fold change using the $2^{-\Delta\Delta Ct}$ method, followed by statistical analysis. Primer sequences are listed in Table 1.

Statistical analyses

All data were expressed as the mean $\pm$ standard deviation. Statistical analysis was performed using Student's *t* test for comparison of two groups. A $p < 0.05$ was considered statistically significant, and the FDR was calculated to correct the *p*-value.

Results

In vitro activation of rat HSCs

HSCs ($\sim 5 \times 10^8$ /mL) were harvested from each rat, with the isolated cells exhibiting a round morphology and a quiescent phenotype at 48 h. Gradually, activation was achieved after a 7- to 14-day culture, with only cultures $>95\%$ viable (as determined by Trypan blue exclusion) used for experiments. The purity of the quiescent

Table 1. Primers used in qRT-PCR. F: forward primer; R: reverse primer

gene	Annealing Temperature (°C)	Primer sequence (5'→3')
NONRATT012636.2	60°C	F: 5'-ACCACGATGACCTTGAATAAAAACT -3' R: 5'-AAAAGTGCAGTAAAGGCACCTTGAA -3'
NONRATT013819.2	60°C	F: 5'-TCCTTCGCGGGATCTGAGT -3' R: 5'-CCCGGCTCGTCCCTTCT -3'
NONRATT016788.2	60°C	F: 5'-CCTCGCTACACCTTGATCT -3' R: 5'-AGGAGGGCTGCTTTTCATTT -3'
NONRATT021402.2	60°C	F: 5'-GGCAGAAGACAACACTTTGGA-3' R: 5'-CAACTATTCAACTGTCTCCATGC-3'
NONRATT007863.2	60°C	F: 5'-ACTGCTCCATCTGCACTCCT -3' R: 5'-AGTCACAAGGGTCGCCATTA -3'
NONRATT019720.2	60°C	F: 5'-TCCCAAAACACCACCAGCAG -3' R: 5'-AGTCTCGAAGGAAGCCGTGA -3'
NONRATT024061.2	60°C	F: 5'-TGACGGGATGTGCTTGTAAG-3' R: 5'-TAGTGACCGGCAAAAGGAAG-3'
<i>Lox</i>	60°C	F: 5'-GTTGAGTCCCGATGTTATGA -3' R: 5'-AGCTGGGGTTTACACTGACCT -3'
<i>Gapdh</i>	60°C	F: 5'-TGGCCTCCAAGGAGTAAGAAAC -3' R: 5'-GGCCTCTCTTCTGCTCTCAGTATC -3'

Fig. 1. Characterization of HSCs isolated from rat livers (200×). (A) Cells were cultured for 2 days (quiescent HSCs), (B) 14 days (activated HSCs). (C) 14 days (activated HSCs), (E) HSC-T6 rat hepatic stellate cell line. (D, F) Immunocytochemical staining of quiescent and activated HSCs for desmin.

Fig. 2. Classification of HSC lncRNAs according to genomic position and overlap with protein-coding genes. The six main classes of lncRNAs are labeled by column (lncRNAs, intronic-antisense lncRNAs, exonic-sense lncRNAs, bidirectional lncRNA, exonic-antisense and intronic-sense lncRNAs). The proportions of the six lncRNA subtypes were calculated.

(>95%) and activated (>95%) HSCs was demonstrated by immunohistochemical examination to determine desmin levels (Fig. 1).

Differentially expressed lncRNAs and mRNAs during HSC activation

After discarding adaptor and low-quality sequences, we obtained raw data consisting of >84,970,970 clean reads (93.3–94.1%) from the Illumina HiSeq platform. TopHat was subsequently used for genome mapping of the clean reads, and Cufflinks was used to reconstruct the transcripts. After basic filtering, we obtained 16,816 tentative lncRNAs containing 3559 lncRNAs, 281 intronic-antisense lncRNAs, 9235 exonic-sense lncRNAs, 1503 bidirectional lncRNA, 281 exonic-antisense lncRNAs, and 2055 intronic-sense lncRNAs (Fig. 2). Following sequence verification, 1454 novel lncRNAs were identified for further analysis.

We then determined differentially expressed lncRNAs between quiescent and activated rat HSCs, identifying 24 lncRNAs meeting the selection threshold (fold change ≥ 2 and $p \leq 0.05$). Among these, 12 lncRNAs were identified as upregulated by >2-fold in activated HSCs as compared with their levels in quiescent HSCs (Fig. 3). Additionally, we found that among 24 different lncRNAs, only two (ENSRNOT00000081905 and TCONS_00042359) did not predict the cis-target gene, with no trans-target genes predicted (Table 2).

To investigate possible lncRNA-mRNA interactions, we performed RNA-Seq analysis of the mRNA data, revealing 529 differentially expressed mRNAs between quiescent and activated rat HSCs. Among these, 155 mRNAs were upregulated in activated rat HSCs, whereas 374 mRNAs were downregulated ($p < 0.01$) (Fig. 3).

Table 2. The list of lncRNAs and its target in quiescent and activated rat HSCs, upregulated lncRNAs. A) Up-regulated lncRNAs. The list of lncRNAs identified in quiescent and activated rat hscs with their mean expression values determined following global normalization and statistical analysis using student t-test as described in the materials and methods. Fold increase in activated HSCs compared to quiescent HSC experiments is shown (Increased expression 2-fold). The P values for activated vs quiescent HSCs for each gene are <0.05. B) Downregulated lncRNAs. Fold decrease in activated HSCs compared to quiescent HSC experiments is shown (Decreased expression 0.5-fold). The P values for activated vs quiescent HSCs for each gene are <0.05

A) lncRNA_id	locus	length	Day14- FPKM	Day2- FPKM	Fold change	p-value	Cis Target
ENSRNOT00000081905	1:88632187-88635003	170	55.2006	1.00E-05	5520060	3.59E-13	Unknown
ENSRNOT00000084803	20:4535717-45357393	214	74.1778	1.05986	69.9883	2.55E-17	Gtf3c6
NONRATT001884.2	1:276659541-276730514	243	21.7661	0.00563	3866.017	0.001518	Tcf7l2
NONRATT007938.2	12:37186274-37210389	229	94.1115	33.4507	2.813439	2.31E-05	Zfp664
NONRATT012636.2	17:80882665-80891212	838	165.454	40.0786	4.128238	1.37E-16	Vim
NONRATT013819.2	18:47500329-47634127	512	445.615	190.366	2.340833	2.26E-22	Lox
NONRATT016196.2	2:66600150-66608324	254	25.1569	0.663993	37.8873	0.000124	Cln5
NONRATT016788.2	2:211344016-211344442	426	131.773	30.8377	4.273114	2.07E-13	SCARNA2
NONRATT017029.2	20:3135790-3166569	258	104.351	39.3096	2.654593	2.25E-05	Rps2-ps1
NONRATT021402.2	4:78905379-78924188	1813	52.5509	0.001371	38327.83	2.35E-12	Tra2a
NONRATT026984.2	7:107895041-107963142	244	34.4094	1.38961	24.76191	4.24E-06	St3gal1
TCONS_00014889	1:101120344-101137147	378	23.1952	1.00E-05	2319520	0.000433	SNORD35
B) lncRNA_id	locus	length	Day14- FPKM	Day2- FPKM	Fold change	p-value	Cis Target
ENSRNOT00000079074	1:88632187-88635003	215	63.4774	179.142	0.354341	2.32E-10	Gtf3c6
NONRATT001882.2	1:276659541-276730514	237	4.3077	38.1529	0.112906	0.000123	Tcf7l2
NONRATT007863.2	12:23674901-23682381	718	184.21	498.522	0.369512	8.21E-29	Upk3b
NONRATT014648.2	19:38237964-38321528	502	24.8743	64.3156	0.386754	0.025128	Smpd3
NONRATT019720.2	3:125503637-125546195	2530	35.4308	85.428	0.414745	0.006307	Lrrn4
NONRATT019765.2	3:143219685-143223615	210	0.000392	43.0062	9.11E-06	3.59E-09	Cst3
NONRATT024061.2	6:24160225-24184216	2691	14.3205	46.7537	0.306297	0.031803	Lbh
NONRATT031134.2	X:104494096-104494370	274	56.9167	116.942	0.486709	0.007273	Pcdh19
TCONS_00018480	1:225906525-225907599	715	44.1917	112.687	0.392163	8.87E-05	Scgb2a1
TCONS_00042359	10:95545667-95563842	250	30.5304	87.3907	0.349355	0.000208	Unknown
TCONS_00218427	7:141428062-141528397	271	1.22477	22.334	0.054839	0.01089	Lima1
TCONS_00229990	8:36625732-36710949	271	1.2499	23.5761	0.053016	0.00653	U1

Distribution of transcript length and exon number between HSC lncRNAs and protein-coding genes

Previous studies showed that both plant and animal lncRNAs are shorter and harbor fewer exons than protein-coding genes. To determine whether HSC lncRNAs share these features, we compared the transcript length, exon number, and expression level of 16,816 lncRNAs with those of protein-coding genes present in HSCs. Figure 4 shows that >80% of lncRNAs ranged in size from 700 to 900 nucleotides, whereas the length of all mRNAs exceeded 1000 bp. Interestingly, most (90%) of the genes encoding HSC lncRNAs only contained one to three exons, whereas the number of exons in protein-coding genes ranged from three to 20. Compared to mRNA-expression levels, lncRNA expression was lower than that of protein coding genes (Fig. 4). These results indicated that the majority of HSC lncRNAs were shorter than HSC protein-coding genes, contained fewer exons, and exhibited lower expression levels.

GO analysis and KEGG pathway annotations

Bioinformatics analyses were used to analyze lncRNA function based on their target genes. GO analysis focusing on biological processes, molecular functions, and cellular components resulted in classification of the top 30 of 783 GO terms (Fig. 5). The most significant GO functions associated with potential lncRNA targets were related to the extracellular exosome, extracellular vesicle, and extracellular region. KEGG pathway analysis

Fig. 3. Differentiation of activated HSCs (14-day cultures) from quiescent HSCs (2-day cultures) by RNA-Seq. Volcano plot shows the upregulated and downregulated (A) lncRNAs and (C) mRNAs in HSCs, respectively. The horizontal axis represents the fold change between quiescent and activated HSCs. The vertical axis represents the p-value of the t test describing the differences between samples.

The box plot displays the distribution of the (B) lncRNA and (D) mRNA datasets. The distributions of log10 (FPKM) among three samples were highly similar.

Fig. 4. Feature comparisons among lncRNA and mRNA transcripts. (A) Length, (B) exon number, and (C) expression level (as represented by FPKM) of lncRNAs relative to protein-coding region and mRNAs.

Fig4. the feature comparison of transcript

(Fig. 6) identified associations with 11 pathways ($p < 0.05$) ranked by target-gene count, with some highly enriched pathways related to ECM remodeling involving mucin-type *O*-glycan biosynthesis and glycosaminoglycan biosynthesis shown to participate in HSC activation. Moreover, *Lox*, a critical member of ECM-related pathways, is a significant target of the lncRNA NONRATT013819.2, both of which were revealed as differentially expressed in this study.

Co-expression relationships between lncRNAs and mRNAs

Networks describing the co-expression of genes and lncRNAs can provide a deeper understanding of the relationship between lncRNAs and regulation of mRNA levels. According to the similarities in gene-/lncRNA-expression profiles, we calculated Pearson's correlation coefficients between lncRNA and gene expression and used expression relationships ≥ 0.98 to construct a co-expression network. Our network consisted of 632 matched lncRNA-mRNA pairs representing 24 differentially expressed lncRNAs and 122 differentially expressed mRNAs (Fig. 7), with 352 and 280 pairs representing positive and negative correlations, respectively. Our results showed that lncRNA NONRATT013819.2 exhibited the highest correlation with co-expressed mRNAs, indicating a potentially important role in regulating differential gene expression between activated and quiescent HSCs.

Fig. 5. GO enrichment histogram for predicted lncRNA cis-target genes. GO enrichment results presented as a histogram, reflecting the differences in target-gene-product functions based on their distribution by the GO terms biological process, cellular component, and molecular function. We selected the 30 most significant GO terms showing high degrees of enrichment for subsequent analysis. The vertical axis represents the GO category; the horizontal axis represents the enrichment of the respective GO terms.

Fig. 6. Pathway analysis based on lncRNA-target genes. The vertical axis represents the KEGG pathway; the horizontal axis represents pathway enrichment. Dot size indicates the number of differentially expressed genes in the pathway.

Validation of genes present in the co-expression network by qRT-PCR

Twenty four lncRNAs were randomly chosen for qRT-PCR validation of significantly differential expression patterns exhibited between activated and quiescent HSCs ($p < 0.05$). Additionally, genes exhibiting higher FPKM values were also selected for qRT-PCR analysis. As shown in Figure 8, qRT-PCR analysis confirmed that NONRATT007863.2, NONRATT019720.2, and NONRATT024061.2 expression was downregulated, and NONRATT012636.2, NONRATT016788.2, and NONRATT021402.2 expression was upregulated in activated HSCs as compared with their levels in quiescent HSCs, which agreed with RNA-Seq results.

Correlation analysis of NONRATT013819.2-Lox co-expression during HSC activation in vitro and in vivo

Because of similarities in NONRATT013819.2-Lox co-expression in the network, we analyzed their chromosomal location and promoter-driven expression. NONRATT013819.2

Fig. 7. Construction of a lncRNA-mRNA co-expression network. The co-expression network comprised 632 connections between 24 lncRNAs and 122 mRNAs. The red nodes represent lncRNAs, and the green nodes represent mRNAs. The correlation between lncRNAs and mRNAs ≥ 0.98 , $p < 0.0001$.

is an antisense lncRNA located on rat chromosome 18:47500329–47634127 and adjacent to the *Lox* gene located on rat chromosome 18:47500330–47577819 (<http://www.bioinfo.org/noncode/>) (Fig. 9). We observed upregulated expression of NONRATT013819.2 and *Lox* according to qRT-PCR analysis during HSC activation (on days 2, 7, and 14 and in the HSC-T6 rat HSC line) and relative to levels observed in quiescent HSCs, as well as significantly positive correlations in their co-expression patterns (Figs. 1 and 9).

Histological examination revealed that the degree of liver fibrosis progressed in rats that received CCL₄ injections as compared with that observed in controls. We also observed that fatty tissue degeneration, necrosis, and fibrogenesis became more evident over time following CCL₄ administration. Furthermore, intrahepatic HSCs isolated from rat livers and subjected to qRT-PCR analysis confirmed the significantly upregulated co-expression of NONRATT013819.2 and *Lox* associated with HSC activation status and fibrogenesis progression as compared with that observed in HSCs isolated from control rats (Fig. 10).

Discussion

HSCs represent the primary mesenchymal cells of the liver and play a major role in ECM maintenance and as inducers of hepatic fibrogenesis upon exposure of the liver to stress [23]. In such cases, HSCs exhibit altered morphology and physiology through the process of activation, which is accompanied by a series of pathological changes based on large-scale gene overexpression and/or repression [24-27]. However, limited understanding of gene

Fig. 8. Validation of RNA-Seq data by qRT-PCR. Each RNA sample was validated in triplicate, and the relative expression level of each gene was normalized to that of GAPDH. * $p < 0.05$, statistically significant differences between quiescent and activated HSCs.

Figure 8. qRT-PCR validation of some differentially expressed lncRNAs and Lox mRNA in the co-expression network

regulation associated with these processes rendered these HSC-specific mechanisms vague and controversial.

lncRNA are small noncoding RNAs >200 nucleotides in length and intimately involved in various critical cell processes, such as development, proliferation, differentiation, and pluripotency [28]. Recent studies reported low expression levels of hepatocellular carcinoma lncRNA maternally expressed gene 3 in CCl₄-induced mouse liver fibrosis and significantly lower expression in the human-shaped cell line LX-2 stimulated by transforming growth factor-β1 [29]. However, the expression and biological function of lncRNAs in HSCs remain largely unknown. Therefore, we performed RNA-Seq analysis of rat quiescent and activated HSCs, with results revealing significant differential expression of 24 lncRNAs and 529 mRNAs. Tests of NONRATT012636.2, NONRATT016788.2, NONRATT021402.2, NONRATT007863.2, NONRATT019720.2, NONRATT024061.2, NONRATT013819.2, and Lox further validated the reliability of RNA-Seq results.

To gain insight into the functions affected by lncRNA activity, GO and KEGG analyses were performed to elucidate the possible biological functions and mechanisms of lncRNAs in HSCs, resulting in the identification of 783 GO terms and 11 pathways. GO terms involving cellular components (extracellular vesicles, extracellular exosomes, and intracellular membrane-bound organelles) dominated the significant GO categories. Moreover, we noted that the significantly enriched terms derived from KEGG analysis agreed with the GO results. Among these pathways and predicted target genes, ECM-related pathways (mucin-type O-Glycan biosynthesis and glycosaminoglycan biosynthesis) and ECM protein cross-linking by Lox appear to play important roles in HSC activation. Integrated bioinformatics analysis of differentially expressed lncRNAs verified that NONRATT013819.2 was the only lncRNA targeting Lox, with its location adjacent to the Lox gene on the same rat chromosome. Furthermore, construction of a lncRNA-mRNA co-expression network displayed critical lncRNAs and mRNAs potentially involved in HSC activation. Additionally, co-expression analyses identified a subset of lncRNAs tightly linked to collagen genes and multiple proteins that regulate ECM remodeling during hepatic fibrogenesis. In this network, positively correlated co-expression of the lncRNA NONRATT013819.2 and Lox mRNA exhibited the highest level of enrichment in ECM-related pathways. Lox, also known as protein-lysine 6-oxidase, is rarely detected in normal or non-proliferative senescent HSCs, but is expressed

Fig. 9. Correlation analysis of NONRATT013819.2-*Lox* co-expression. (A) NONRATT013819.2 is located on rat chromosome 18:47500329–47634127 adjacent to the *Lox* gene located on chr18:47500330–47577819. (B) Positive correlation between NONRATT013819.2 and *Lox* co-expression in HSCs. correlation = 0.98, $p < 0.0001$.

Fig. 10. Hematoxylin and eosin (HE) and Van Gieson (VG) staining of normal and CCl₄-treated liver tissue (100×). (A) Liver-tissue staining. (B) LncRNA NONRATT013819.2 and *Lox* mRNA levels in intrahepatic HSCs following CCl₄ treatment and as assessed by qRT-PCR. * $p < 0.05$, statistically significant differences between quiescent and activated HSCs.

in culture-activated HSCs and intra-hepatic HSCs following CCl₄ injury [30, 31]. Recent studies attributed ECM remodeling, spontaneous reversion of fibrogenesis, and induced phenotypic reversion of HSCs to *Lox* overexpression [7, 32]. These findings suggested that the dysregulated lncRNA-mRNA network might be closely associated with ECM remodeling and HSC activation.

Our experimental findings were consistent with these reports. Subsequent examination of NONRATT013819.2 and *Lox* co-expression both *in vitro* and *in vivo* confirmed their upregulation during HSC activation relative to levels observed in quiescent HSCs and suggested their association with HSC activation and progression during fibrogenesis. Moreover, the most significant finding from data mining was the high degree of association between lncRNA NONRATT013819.2/*Lox* mRNA and the ECM, the production of which by HSC myofibroblast-like cells is the main cause of fibrogenesis and progression to liver failure in chronic liver diseases [33–35]. These results indicated that NONRATT013819.2-*Lox* co-expression and possible interaction might promote HSC transformation and proliferation by activating signaling pathways associated with ECM remodeling. However, the function of lncRNA in liver fibrosis and related mechanisms remains to be confirmed, given that a

single lncRNA might potentially regulate the expression of multiple protein-coding genes at diverse levels [36]. Therefore, additional evidence regarding the cross-talk between NONRATT013819.2 and *Lox* remains to be experimentally validated in our future work.

Conclusion

In conclusion, we investigated the differential expression of lncRNAs and mRNAs during HSC transformation, with results of RNA-Seq analysis indicating that lncRNAs were involved in pathological processes associated with the progression of fibrogenesis. Furthermore, bioinformatics analysis and *in vitro* and *in vivo* experiments identified correlations between lncRNA NONRATT013819.2 and *Lox* mRNA expression and potential interaction related to ECM-remodeling pathways. These findings provide valuable insight into HSC activation and also facilitate access to novel therapeutic strategies for the treatment of liver fibrogenesis.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (81570544 and 81100296, 81670546).

Disclosure Statement

All authors declare no conflicts of interest.

References

- 1 Yin X, Yi H, Wang L, Wu W, Wu X, Yu L: RSPOs facilitated HSC activation and promoted hepatic fibrogenesis. *Oncotarget* 2016;7:63767-63778.
- 2 Guo Z, Li D, Peng H, Kang J, Jiang X, Xie X, Sun D, Jiang H: Specific Hepatic Stellate Cell-penetrating Peptide Targeted Delivery of a KLA Peptide Reduces Collagen Accumulation by Inducing Apoptosis. *J Drug Target* 2017;1-26.
- 3 Ganai AA, Husain M: Genistein attenuates D-GalN induced liver fibrosis/chronic liver damage in rats by blocking the TGF-beta/Smad signaling pathways. *Chem Biol Interact* 2017;261:80-85.
- 4 Zhang Y, Ghazwani M, Li J, Sun M, Stolz DB, He F, Fan J, Xie W, Li S: MiR-29b inhibits collagen maturation in hepatic stellate cells through down-regulating the expression of HSP47 and lysyl oxidase. *Biochem Biophys Res Commun* 2014;446:940-944.
- 5 Wang TH, Hsia SM, Shieh TM: Lysyl Oxidase and the Tumor Microenvironment. *Int J Mol Sci* 2016;18.
- 6 He Y, Jin L, Wang J, Yan Z, Chen T, Zhao Y: Mechanisms of fibrosis in acute liver failure. *Mechanisms of fibrosis in acute liver failure. Liver Int* 2015;35:1877-1885.
- 7 Zhou C, York SR, Chen JY, Pondick JV, Motola DL, Chung RT, Mullen AC: Long noncoding RNAs expressed in human hepatic stellate cells form networks with extracellular matrix proteins. *Genome Med* 2016;8:31.
- 8 Friedman SL, Roll FJ: Isolation and culture of hepatic lipocytes, Kupffer cells, and sinusoidal endothelial cells by density gradient centrifugation with Stractan. *Anal Biochem* 1987;161:207-218.
- 9 Guo CJ, Pan Q, Li DG, Sun H, Liu BW: miR-15b and miR-16 are implicated in activation of the rat hepatic stellate cell: An essential role for apoptosis. *J Hepatol* 2009;50:766-778.
- 10 Pawlak A, Gil RJ, Walczak E, Fidzianska A: Evaluation of desmin activity using immunohistochemical and immunofluorescent staining of myocardial biopsies in patients with chronic heart failure. Comparison of the two methods. *Kardiologia Pol* 2007;65:229-235;discussion 236.
- 11 Guo CJ, Pan Q, Cheng T, Jiang B, Chen GY, Li DG: Changes in microRNAs associated with hepatic stellate cell activation status identify signaling pathways. *FEBS J* 2009;276:5163-5176.
- 12 Ishak K, Baptista A, Bianchi L, Callea F, De Groote J, Gudat F, Denk H, Desmet V, Korb G, MacSween RN: Histological grading and staging of chronic hepatitis. *J Hepatol* 1995;22:696-699.
- 13 Robinson MD, McCarthy DJ, Smyth GK: edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010;26:139-140.

- 14 Ghosh S, Chan CK: Analysis of RNA-Seq Data Using TopHat and Cufflinks. *Methods Mol Biol* 2016;1374:339-361.
- 15 Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L: Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* 2010;28:511-515.
- 16 Sun L, Zhang Z, Bailey TL, Perkins AC, Tallack MR, Xu Z, Liu H: Prediction of novel long non-coding RNAs based on RNA-Seq data of mouse Klf1 knockout study. *BMC Bioinformatics* 2012;13:331.
- 17 Kong L, Zhang Y, Ye ZQ, Liu XQ, Zhao SQ, Wei L, Gao G: CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res* 2007;35:W345-349.
- 18 Sun L, Luo H, Bu D, Zhao G, Yu K, Zhang C, Liu Y, Chen R, Zhao Y: Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. *Nucleic Acids Res* 2013;41:e166.
- 19 Knauss JL, Sun T: Regulatory mechanisms of long noncoding RNAs in vertebrate central nervous system development and function. *Neuroscience* 2013;235:200-214.
- 20 Tafer H, Hofacker IL: RNAplex: a fast tool for RNA-RNA interaction search. *Bioinformatics* 2008;24:2657-2663.
- 21 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 Res* 2003;13:2498-2504.
- 22 Janosky JE: Pearson correlation coefficients vs reliability coefficients. *J Am Diet Assoc* 1991;91:912-913.
- 23 Baig MS, Yaqoob U, Cao S, Saqib U, Shah VH: Non-canonical role of matrix metalloprotease (MMP) in activation and migration of hepatic stellate cells (HSCs). *Life Sci* 2016;155:155-160.
- 24 Chen L, Cui X, Li P, Feng C, Wang L, Wang H, Zhou X, Yang B, Lv F, Li T: Suppression of MicroRNA-219-5p Activates Keratinocyte Growth Factor to Mitigate Severity of Experimental Cirrhosis. *Cell Physiol Biochem* 2016;40:253-262.
- 25 Yu F, Yang J, Huang K, Pan X, Chen B, Dong P, Zheng J: The Epigenetically-Regulated microRNA-378a Targets TGF-beta2 in TGF-beta1-Treated Hepatic Stellate Cells. *Cell Physiol Biochem* 2016;40:183-194.
- 26 Yu F, Fan X, Chen B, Dong P, Zheng J: Activation of Hepatic Stellate Cells is Inhibited by microRNA-378a-3p via Wnt10a. *Cell Physiol Biochem* 2016;39:2409-2420.
- 27 Li G, Li J, Li C, Qi H, Dong P, Zheng J, Yu F: MicroRNA-125a-5p Contributes to Hepatic Stellate Cell Activation through Targeting FIH1. *Cell Physiol Biochem* 2016;38:1544-1552.
- 28 Wei W, Liu Y, Lu Y, Yang B, Tang L: LncRNA XIST Promotes Pancreatic Cancer Proliferation Through miR-133a/EGFR. *J Cell Biochem* 2017;9999:1-10.
- 29 He Y, Wu YT, Huang C, Meng XM, Ma TT, Wu BM, Xu FY, Zhang L, Lv XW, Li J: Inhibitory effects of long noncoding RNA MEG3 on hepatic stellate cells activation and liver fibrogenesis. *Biochim Biophys Acta* 2014;1842:2204-2215.
- 30 El Taghdouini A, Najimi M, Sancho-Bru P, Sokal E, van Grunsven LA: In vitro reversion of activated primary human hepatic stellate cells. *Fibrogenesis Tissue Repair* 2015;8:14.
- 31 Liu SB, Ikenaga N, Peng ZW, Sverdlov DY, Greenstein A, Smith V, Schuppan D, Popov Y: Lysyl oxidase activity contributes to collagen stabilization during liver fibrosis progression and limits spontaneous fibrosis reversal in mice. *FASEB J* 2016;30:1599-1609.
- 32 Li SY, Yan JQ, Song Z, Liu YF, Song MJ, Qin JW, Yang ZM, Liang XH: Molecular characterization of lysyl oxidase-mediated extracellular matrix remodeling during mouse decidualization. *FEBS Lett* 2017;591:1394-1407.
- 33 Santos-Laso A, Munoz-Garrido P, Felipe-Agirre M, Bujanda L, Banales JM, Perugorria MJ: New advances in the molecular mechanisms driving biliary fibrosis and emerging molecular targets. *Curr Drug Targets* 2015.
- 34 Sziksz E, Pap D, Lippai R, Beres NJ, Fekete A, Szabo AJ, Vannay A: Fibrosis Related Inflammatory Mediators: Role of the IL-10 Cytokine Family. *Mediators Inflamm* 2015;2015:764641.
- 35 Kumar V, Mahato RI: Delivery and targeting of miRNAs for treating liver fibrosis. *Pharm Res* 2015;32:341-361.
- 36 Costa MC, Leitao AL, Enguita FJ: Noncoding Transcriptional Landscape in Human Aging. *Curr Top Microbiol Immunol* 2016;394:177-202.