Arch Biol Sci. 2022;74(1):25-39 https://doi.org/10.2298/ABS211207001L

# Integration of miRNA-lncRNA-mRNA profiles in liver tissue from *EpCAM* knockout mice

Zili Lei[1], Yuting Lei[1], Guibin Chen[1], Shaomin Liu[1,2], Wanwan Liu[1], Li Huang[1,2], Lanxiang Yang[1,2], Huijuan Wu[1,2] and Yanhong Yang[3,*]

[1]*Guangdong Metabolic Diseases Research Center of Integrated Chinese and Western Medicine; Guangdong TCM Key Laboratory for Metabolic Diseases; Key Laboratory of Glucolipid Metabolic Disorder, Ministry of Education of China; Institute of Chinese Medicine*, *Guangdong Pharmaceutical University*, Guangzhou 510006, P.R. China

[2]*School of Traditional Chinese Medicine, Guangdong Pharmaceutical University, Guangzhou Higher Education Mega Center*, Guangzhou 510006, P.R. China

[3]*First Affiliated Hospital (School of Clinical Medicine), Guangdong Pharmaceutical University*, Guangzhou 510080, P.R. China

***Corresponding author**: 1764941457@qq.com

**Received**: December 7, 2021; **Revised**: December 28, 2021; **Accepted**: December 30, 2021; **Published online**: January 11, 2022

**Abstract:** The epithelial cell adhesion molecule (EpCAM) is highly expressed in the liver during development and diseases. However, its role in the development and pathology of liver remains to be explored. The liver tissues of EpCAM$^{-/-}$ and wildtype (WT) mice at P0 stage were used for RNA sequencing. The differently expressed miRNAs, lncRNAs and mRNAs were selected and confirmed by qPCR. The expression of metabolism-related gene SET domain bifurcated 2 (*Setdb2*) was significantly increased in the liver of EpCAM$^{-/-}$ mice; the triglyceride (TG) and total cholesterol (TC) levels in the liver were also markedly decreased in EpCAM$^{-/-}$ mice. The microRNA (miRNA)-long noncoding RNA (lncRNA)-mRNA regulatory networks indicated that *EpCAM* may play important roles in glucose and lipid metabolism of the liver during development and in disease. The comprehensive miRNA, lncRNA and mRNA expression profiles in the developing liver of EpCAM$^{-/-}$ mice established here might help to elucidate functions and mechanisms of *EpCAM* during development and in diseases of the liver.

**Keywords:** epithelial cell adhesion molecule (EpCAM); metabolism-related gene SET domain bifurcated 2 (Setdb2); miRNA-lncRNA-mRNA regulatory network; mouse liver

## INTRODUCTION

EpCAM has been found to play an important role in maintaining the intestinal barrier by recruiting claudins to cell-cell junctions, with *EpCAM* knockout mice exhibiting defects in the intestinal barrier and dying soon after birth as a result of intestinal erosion [1]. Apart from its role in cell-cell junctions, *EpCAM* is also involved in several other physiological and biological processes, including cell signaling, proliferation, differentiation, development, formation and maintenance of organ morphology, which we summarized in detail in our previous report [2]. *EpCAM* is involved in several signaling pathways, including the nPKC-dependent pathway, the Wnt signaling pathway and the embryonic stem cell-expressed Ras (ERas)/protein kinase B (AKT) pathway [2]. Due to its important roles in maintaining the intestinal barrier, *EpCAM* mutations could lead to severe intestinal disease, such as congenital tufting enteropathy (CTE) [2-4], and *EpCAM* knockdown *in vivo* could increase the severity of inflammatory bowel disease (IBD) in mice [5].

*EpCAM* also plays an important role in the biological or pathological processes of the liver. It has been reported that *EpCAM* was expressed in the mouse liver bud on embryonic day (E) 9.5 and that hepatoblasts expressing both *EpCAM* and the δ-like 1 homolog could form the hepatic cords at the early stage of hepatogenesis; however, the expression of *EpCAM* in hepatoblasts was markedly decreased, along with liver development, and its expression almost disappeared

at E14.5 [6]. It was reported that the expression of a number of adhesion proteins, including epithelial cadherin (E-cadherin), neural cadherin (N-cadherin), *EpCAM*, intracellular cell adhesion molecule (ICAM), collagen 1α1 and α-actinin was low around prenatal day 14, but increased during the postnatal period in rats [7]. Abnormal *EpCAM* expression was associated with hepatic cancer. It has been reported that *EpCAM* is normally expressed in hepatic progenitors but re-expressed in hepatic cancer stem cells (HCSCS), and that it could activate canonical Wnt signaling in hepatitis B virus (HBV)-replicating cells, potentially leading to liver cancer [8]. A systematic review and meta-analysis revealed that *EpCAM* overexpression was associated with poor differentiation and a high α fetoprotein level in hepatocellular carcinoma (HCC); in addition, *EpCAM* overexpression was a predictor of an overall short survival period and disease-free survival [9]. However, the underlying mechanisms and signaling pathways of *EpCAM* in the physiological and pathological processes of the liver remain unclear.

High-throughput sequencing (HTS) techniques make it possible to identify numerous different biological components simultaneously; they provide useful tools for understanding the underlying molecular mechanisms of several biological processes or diseases [10]. Non-coding RNAs (ncRNAs) include housekeeping and regulatory RNAs; housekeeping RNAs include ribosomal RNAs (rRNAs), small nuclear RNAs (snRNAs) and transfer RNAs (tRNAs), which are part of the ribosomal machinery; regulatory RNAs include short ncRNAs that are <200 nucleotides in length, such as microRNAs (miRNAs) and piwi-interacting RNAs (piRNAs), as well as long noncoding RNAs (lncRNAs) that are >200 nucleotides in length [11]. ncRNAs constitute about 60% of the transcriptional output in human cells and are involved in several developmental and pathological processes [12]. The dysregulation of ncRNAs may lead to the development of certain types of disease, including cancer, as well as neurological, cardiovascular, developmental and other diseases [13]. With the development of HTS technology, an increasing number of new ncRNAs are being identified, and it is difficult to determine the separate functions of different types of ncRNAs. For example, miRNAs could target several mRNAs, and the mRNA of one gene could also be targeted by several miRNAs. In addition, miRNAs could interact with circular RNAs (circRNAs) or lncRNAs to regulate their stability, and lncRNAs and circRNAs could also regulate the abundance of miRNAs [12]. Furthermore, it has been reported that lncRNAs could act as a competing endogenous RNA (ceRNA) for miRNAs and play important roles in tumor development [14]. ceRNAs are endogenous RNA transcripts that share miRNA response elements, which regulate the expression of mRNAs by competing for miRNA regulators at the posttranscriptional level, linking ncRNAs such as miRNAs, lncRNAs, pseudogenes and circRNAs [15]. The complexity of the interactions among ncRNAs makes them essential in the regulation of biological processes, and the breakdown of the interactions might lead to different diseases.

In order to systematically study the functions of *EpCAM* in the biological and physiological processes of the liver, the effects of *EpCAM* on the expression profiles of mRNAs, lncRNAs and miRNAs in the liver tissue of *EpCAM* gene knockout mice at P0 stage were investigated by HTS. The expression profile of circRNAs in the liver of *EpCAM*$^{-/-}$ mice was analyzed in our previous study [16] in which several novel circRNAs were identified and a circRNA-miRNA-mRNA regulatory network was established. In the present study, the expression profiles of mRNAs, lncRNAs and miRNAs in the liver of *EpCAM*$^{-/-}$ mice at P0 were established, and the interactions of microRNA-lncRNA-mRNA regulatory networks were analyzed, providing important information about the functions and mechanisms of EpCAM in the developing liver. Directions for the future development of novel therapeutic targets for liver disease are presented.

## MATERIALS AND METHODS

### Ethics statement

All animal experiments were approved by the Experimental Animal Ethics Committee of the Guangdong Pharmaceutical University (Guangzhou, China).

### Experimental mice

The total number, age and sex of all mice used in the present study are summarized in Supplementary Table S1. The generation of *EpCAM*$^{-/-}$ mice has

been described in detail in our previous study [16]. Wildtype (WT) and *EpCAM*$^{-/-}$ mice were killed on the first day after birth by cervical dislocation; the livers were weighed, frozen in liquid nitrogen, and stored at -80˚C for preservation. The livers of every 3 mice (≥200 mg) were collected as 1 sample. The control WT and *EpCAM*$^{-/-}$ groups had 3 samples each [16]. The liver index was calculated by dividing the liver weight by body weight multiplied by 100.

**Library construction and miRNA sequencing analysis**

miRNA isolation, library construction, and miRNA sequencing were performed by the Gene Denovo Biotechnology Company (Guangzhou, China). Total RNA was extracted using a TRIzol reagent kit (Invitrogen; Thermo Fisher Scientific Inc., USA) according to the manufacturer's instructions, and the RNA molecules measuring 18-30 nucleotides were enriched by polyacrylamide gel electrophoresis. The library was constructed using an NEBNext® Multiplex Small RNA Library Prep Kit for Illumina (E7300; New England BioLabs, Inc., MA, USA). The 3' adapters were then added and the 36-44 nucleotide RNAs were enriched; the 5' adapters were then ligated to the RNAs. The ligation products were reverse transcribed by PCR amplification, and the 140-160-base-pair-sized PCR products were enriched to generate a cDNA library. The quality of the library was assessed on a Bioanalyzer 2100 system (Agilent Technologies Inc., CA, USA), and an ABI StepOnePlus Real-Time PCR System (Life Technologies Inc., CA, USA) was used for quantification. A 3 nM cDNA library was sequenced from 5' to 3' using Illumina Hiseq 2500 SE50 with sequencing strategy of single-end 50 bp reads (Illumina Inc., CA, US). Libraries were then amplified and sequenced using HiSeq Rapid SBS Kit V2 and HiSeq Rapid SRCluster Kit V2 (Illumina, Inc.). The procedure was performed as previously described [16].

Reads were further filtered to remove the low-quality reads containing adapters, reads with unknown nucleotides or reads measuring <18 nucleotides. All clean tags were searched against the miRBase database (Release 21) to identify existing murine miRNAs and alignment with other species to identify known miRNAs. All the unannotated tags were aligned with the reference genome, and the novel miRNAs were identified according to their genome positions and the hairpin structures predicted by Mireap_v0.2 software.

Total miRNA consists of existing miRNA, known miRNA and novel miRNA, based on their expression in each sample. The miRNA expression level was calculated and normalized to transcripts per million. miRNAs with a fold change of ≥2 and P <0.05 in a comparison were identified as differentially expressed miRNAs. RNAhybrid (version 2.1.2)+svm_light (version 6.01), Miranda (version 3.3a) and TargetScan (version 7.0) were used to predict the target genes of the existing miRNAs, known miRNAs and novel miRNAs, and the functions of miRNA target genes were analyzed via functional enrichment analysis as described below.

**Library construction, lncRNA and mRNA sequencing analysis**

LncRNA isolation, library construction, and lncRNA sequencing were performed by Gene Denovo Biotechnology Company (CA, USA). Total RNA was extracted using a TRIzol reagent kit (Invitrogen; Thermo Fisher Scientific Inc.), according to the manufacturer's instructions. RNA quality was assessed on an Agilent 2100 Bioanalyzer (Agilent Technologies, Inc.) and checked using RNase-free agarose gel electrophoresis. After total RNA was extracted, rRNAs were removed to retain mRNAs and ncRNAs using a Ribo-Zero™ Magnetic kit (Epicentre; Lucigen Corporation, WI, USA). The enriched mRNAs and ncRNAs were broken into short fragments using fragmentation buffer and reverse-transcribed into cDNA using random primers. Second-strand cDNA was synthesized by DNA polymerase I, RNase H, dNTP (dUTP instead of dTTP) and buffer. Next, the cDNA fragments were purified with a QIAquick PCR extraction kit (Qiagen AB, Sweden) and end-repaired. Poly (A) was then added, and the cDNA fragments were ligated to Illumina sequencing adapters (Illumina Inc.). Uracil-N-glycosylase was then used to digest the second-strand cDNA. The digested products underwent size selection by agarose gel electrophoresis and were amplified by PCR. The quality of the library was assessed using an Agilent High Sensitivity DNA kit on a Bioanalyzer 2100 system (Agilent Technologies

Inc., CA, USA), with ABI StepOnePlus Real-Time PCR System (Life Technologies Inc.) finally used for quantification. A 3 nM cDNA library was sequenced from 5' to 3' using Illumina HiSeq 2500-PE150 with sequencing strategy of Paired-end 150 bp reads. The kit used in the sequencing process was HiSeq Rapid SBS Kit v2 (Illumina Inc.).

Reads were further filtered to remove low-quality reads containing adapters, reads with >10% of unknown nucleotides or reads containing >50% of low-quality (Q-value≤20) bases. All clean reads were mapped with rRNA database using Bowtie2 (version 2.2.8, http://bowtie-bio.sourceforge.net/bowtie2/index.shtml) [17]. The rRNA mapped reads were then removed. The remaining reads were mapped to the reference genome by TopHat2 (version 2.1.1, http://ccb.jhu.edu/software/tophat/index.shtml) [18]. The transcripts were reconstructed using Cufflinks software [19], and further downstream differential expression analysis was performed. The differentially expressed transcripts of mRNAs and lncRNAs were analyzed. The edgeR package (http://www.r-project.org/) was used to identify differentially expressed transcripts in groups, and transcripts with a fold change of ≥2 and a false discovery rate (FDR) of <0.05 were identified as differentially expressed transcripts. Differentially expressed mRNAs were then subjected to enrichment analysis as shown below. The interaction between antisense lncRNAs and mRNAs was analyzed using RNAplex software (http://www.tbi.univie.ac.at/RNA/RNAplex.1.html) [20]. The lncRNA *trans*-regulation analysis was evaluated using Pearson's correlation coefficient (PCC).

## Gene ontology (GO) and the Kyoto encyclopedia of genes and genomes (KEGG) pathway analysis

GO is a gene functional classification system that defines the properties and functions of genes in different organisms. GO has 3 ontologies: biological process, cellular component and molecular function. All source genes were mapped to GO terms in the GO database (http://www.geneontology.org/), and gene numbers were calculated for every term, with an FDR of ≤0.05 as the threshold. KEGG pathway analysis (https://www.kegg.jp/) identifies significantly enriched pathways in source genes, as compared with the whole genome, which further helps understand the functions of genes. The calculating formula is the same as that for GO analysis, and pathways with an FDR of ≤0.05 were defined as significantly enriched. The procedure of GO and KEGG analysis was the same as described in our previous study [16].

## qRT-PCR

The mice were killed by cervical dislocation, and cessation of breathing and loss of heartbeat were confirmed as death of the mice. Total RNA was extracted from the liver samples at P0 using TRIzol reagent (Invitrogen, Thermo Fisher Scientific Inc.), and subjected to reverse transcription using the PrimeScriptTM RT Reagent kit (Takara Bio Inc., Japan). All mRNA primers are listed in Supplementary Table S2, and all miRNAs and lncRNAs primers in Supplementary Table S3 (Sangon Biotech Co., Ltd., Shanghai, China). qRT-PCR of mRNAs was performed using the SYBR Premix Ex Taq kit (Takara Bio Inc., Shiga, Japan) with the PikoReal PCR system (Thermo Fisher Scientific Inc.) under the following conditions: 95˚C for 30 s, followed by 40 cycles at 95˚C for 5 s, 60˚C for 20 s and 65˚C for 15 s. GAPDH was used as a reference gene. qRT-PCR of miRNAs and lncRNAs was performed using a LightCycler480 system (Roche Diagnostics, Switzerland) under the following conditions: 95˚C for 30 s, followed by 40 cycles at 95˚C for 5 s and 60˚C for 30 s, then 95˚C for 5 s, and 60°C for 1 min. GAPDH was used as a reference gene for lncRNAs, and U6 as a reference gene for miRNAs.

## Hematoxylin and eosin (H&E) staining

The mouse liver tissue at E18.5, P0 and P3 was fixed in 4% paraformaldehyde (Leagene Biotechnology Co. Ltd., Beijing, China) at 4°C overnight, dehydrated with ascending series of alcohol and embedded in paraffin. H&E staining was performed on 4-μm-thick sections, stained with hematoxylin for 3 min and eosin for 20 sec at room temperature (Merck KGaA, NJ, USA). Images of H&E staining were captured using an automated quantitative pathology system (PerkinElmer Automated Quantitative Pathology System; PerkinElmer Inc., MA, USA).

### Detection of triglyceride (TG) and total cholesterol (TC) in the liver of mice

The liver sample at P3 was prepared into a homogenate by adding isotonic buffer, with weight:volume ratio of 1:9, and the homogenate was centrifuged (1500 ×g for 10 min). The supernatant was used to determine TG and TC using their respective kits (cat. no. A110-1-1 for TG; cat. no. A111-1-1 for TC; Nanjing Jiancheng Bioengineering Institute, China) according to the manufacturer's instructions.

### Integrated analysis of lncRNA-miRNA-mRNA

RNAhybrid (version 2.1.2) + svm_light (version 6.01), Miranda (version 3.3a) and TargetScan (version: 7.0) were used to analyze which mRNA and which lncRNAs could be targeted by miRNA, and the miRNA-mRNA and miRNA-lncRNA targeting relationship was then integrated into the lncRNA-miRNA-mRNA network. The lncRNA-miRNA-mRNA network was constructed by assembling all co-expression competing triplets and was visualized using Cytoscape software (version 3.6.0; http://www.cytoscape.org/).

### ceRNAs network analysis

Expression correlation between mRNA-miRNA or lncRNA-miRNA was evaluated using Spearman's rank correlation coefficient (SCC). Pairs with an SCC of <-0.7 were selected as negatively co-expressed lncRNA-miRNA pairs or mRNA-miRNA pairs. The correlation between the expression of lncRNA and that of mRNA was evaluated using PCC. Pairs with a PCC of >0.9 were selected as co-expressed lncRNA-mRNA pairs, and both mRNA and lncRNA in this pair were targeted and co-expressed negatively with a common miRNA. A hypergeometric cumulative distribution function test was performed to check whether the common miRNA sponges between the two genes were significant, and only the gene pairs with a $P<0.05$ were selected. The lncRNA-miRNA-mRNA network was constructed by assembling all co-expression competing triplets and was visualized using Cytoscape software (version 3.6.0; http://www.cytoscape.org/).

### Statistical analysis

Statistical differences were determined using SPSS 23.0 software. Data were expressed as the mean±SE. A t-test was performed between two groups. Each experiment was repeated at least 3 times. $P<0.05$ was considered to indicate a statistically significant difference.

## RESULTS

### Overview of miRNAs in the livers of EpCAM$^{-/-}$ mice

According to the sequence data, 72851959 clean reads (total 38958436 and average 12986145 for EpCAM$^{-/-}$ mice; total 33893523 and average 11297841 for WT mice) were generated. Following the removal of the low-quality and adapter- and poly-N-containing reads, 71836336 high-quality clean reads (total 38388525 and average 12796175 from EpCAM$^{-/-}$ mice, and a total 33447811 and average 11149270 from WT mice were obtained. Total miRNAs consist of existing miRNAs, known miRNAs and novel miRNAs. The number of differentially expressed existing miRNAs, known miRNAs and novel miRNAs are shown in Supplementary Fig. S1, S2 and S3, respectively, and a total of 24 upregulated and 13 downregulated miRNAs were identified in the EpCAM$^{-/-}$ mouse group compared with the WT group (Fig. 1A and B). The target genes of the differentially expressed miRNAs were mapped to terms in the GO database (Fig. 1C). With regards to biological process, most genes were related to "cellular process", "single-organism process" and "metabolic process". With regard to the cellular component, most genes were related to the "cell", "cell part" and "organelle part". With regards to molecular function, most genes were related with "binding", "catalytic activity" and "signal transducer activity". The top 20 pathways of the KEGG assignments of the target genes of differentially expressed miRNAs are shown in Fig. 1D, including "olfactory transduction", "Ras signaling pathway", "calcium signaling pathway", "Th17 cell differentiation", "TNF signaling pathway, "Th1 and Th2 cell differentiation", "cAMP signaling pathway", "signaling pathways regulating pluripotency of stem cells", "Rap1 signaling pathway", "Wnt signaling pathway", "natural killer cell mediated

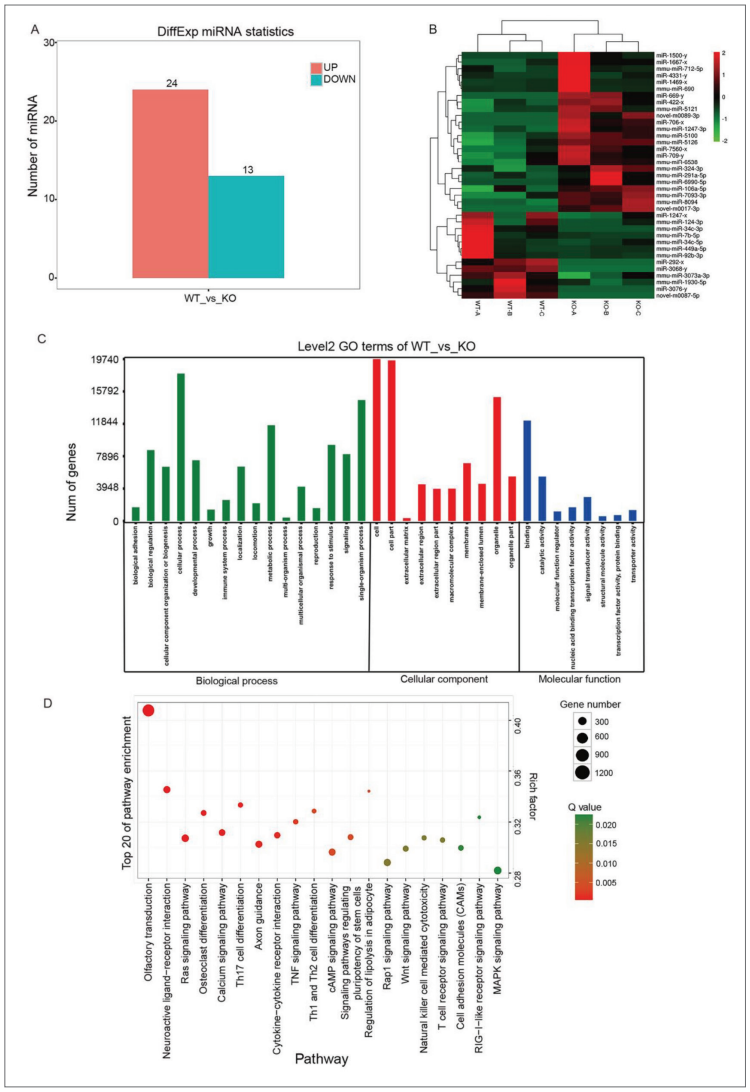


**Fig. 1**. miRNA expression profile in the liver of *EpCAM*$^{-/-}$ mice. **A** – A total of 24 upregulated and 13 downregulated miRNAs were identified. **B** – Heatmap of differentially expressed miRNAs. **C** – GO classification of the target genes of the differentially expressed miRNAs. **D** – KEGG assignments of the target genes of the differentially expressed miRNAs. EpCAM – epithelial cell adhesion molecule; miRNAs – microRNAs; GO – gene ontology; KEGG – Kyoto Encyclopedia of Genes and Genomes.

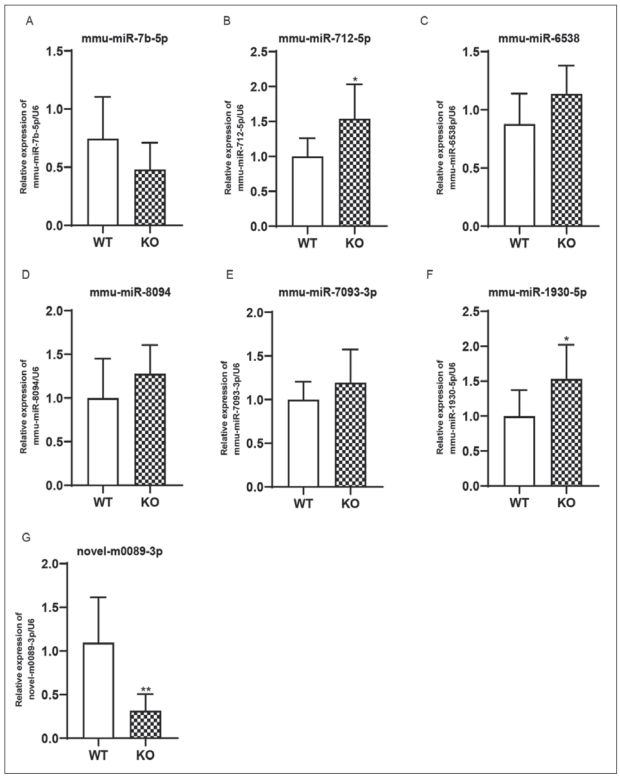


**Fig. 2.** The expression of some differentially expressed miRNAs was validated by qRT-PCR. **A** – mmu-miR-7b-5p; **B** – mmu-miR-712-5p; **C** – mmu-miR-6538; **D** –mmu-miR-8094; **E** – mmu-miR-7093-3p; **F** – mmu-miR-1930-5p; **G** –novel-m0089-3p. $^{*}$P<0.05 and $^{**}$P<0.01 vs WT. miRNAs – microRNAs; WT – wild-type.

cytotoxicity", "T cell receptor signaling pathway", "cell adhesion molecules (CAMs)", "RIG-I-like receptor signaling pathway" and "MAPK signaling pathway". The GO and KEGG analysis of the target genes of the differentially expressed existing miRNAs, known miRNAs and novel miRNAs are demonstrated in Supplementary Figs. S1, S2 and S3, respectively. In order to confirm the differentially expressed miRNAs predicted by sequencing, 7 miRNAs, including mmu-miR-7b-5p, mmu-miR-712-5p, mmu-miR-6538, mmu-miR-8094, mmu-miR-7093-3p, mmu-miR-1930-5p and novel-m0089-3p, were randomly selected to examine the expression level by RT-qPCR (Fig. 2). Results showed that the expression of most miRNAs was consistent with the sequencing data. The expression of mmu-miR-712-5p and mmu-miR-1930-5p was significantly increased (P<0.05), and the novel-m0089-3p was significantly decreased (P<0.01), which was in agreement with the sequencing data. The expression of mmu-miR-6538, mmu-miR-8094 and mmu-miR-7093-3p was increased, and that of mmu-miR-7b-5p was decreased, although not significantly, which was also consistent with the sequencing results.

## Overview of lncRNAs and mRNAs in the livers of *EpCAM*$^{-/-}$ mice

For lncRNAs, 2268 known transcripts and 854 novel transcripts were identified. A total of 3 upregulated

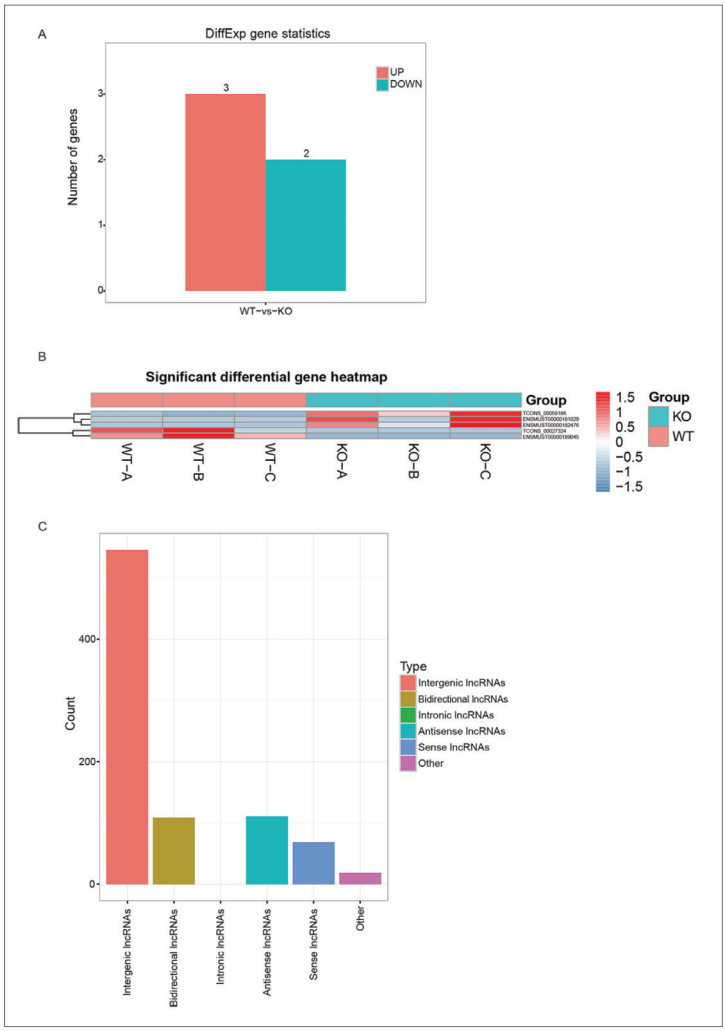


**Fig. 3**. lncRNA expression profile in the liver of *EpCAM*$^{-/-}$ mice. **A** – A total of 3 upregulated and 2 downregulated lncRNAs were identified. **B** – Heatmap of differentially expressed lncRNAs. **C** – Distribution of the different types of novel lncRNAs. lncRNA – long non-coding RNAs; EpCAM – epithelial cell adhesion molecule.

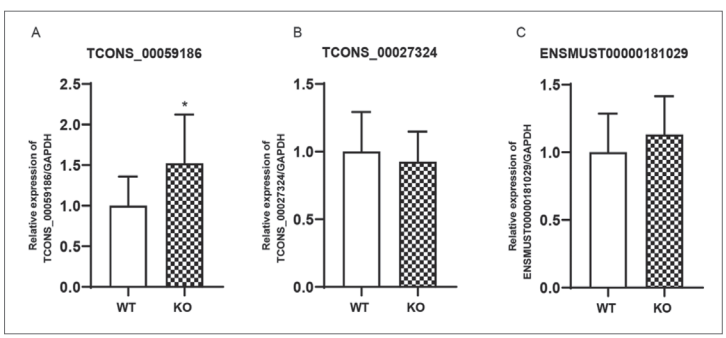


**Fig. 4.** The expression of some differentially expressed lncRNAs was validated by qRT-PCR. **A** – TCONS_00059186; **B** – TCONS_00027324; **C** –ENSMUST00000181029. *P<0.05 vs. WT. lncRNAs – long non-coding RNAs; WT – wild-type.

and 2 downregulated lncRNAs were identified in the *EpCAM*$^{-/-}$ mouse group compared with the WT group (Fig. 3A); the heat map of the 5 differentially expressed lncRNAs is presented in Fig. 3B. lncRNAs were categorized into five classes according to their location relative to protein-coding genes as follows: intergenic lncRNAs, bidirectional lncRNAs, intronic lncRNAs, antisense lncRNAs and sense overlapping lncRNAs. The distribution of the different types of novel lncRNAs is shown in Fig. 3C. In order to confirm the differentially expressed lncRNAs predicted by the sequencing, the expression levels of 3 randomly selected lncRNAs, including TCONS_00059186, TCONS_00027324 and ENSMUST00000181029, were examined using RT-qPCR (Fig. 4). The result demonstrated that the expression of TCONS_00059186 was significantly increased ($P<0.05$). The expression of TCONS_00027324 was decreased, and that of ENSMUST00000181029 was increased, although not significantly. These results were in agreement with the sequencing data.

According to the sequencing data, 39595 known transcripts and 8633 novel transcripts were identified in the mRNAs. A total of 19 upregulated and 11 downregulated mRNAs were identified in the *EpCAM*$^{-/-}$ mouse group compared with the WT mouse group (Fig. 5A). The heat map of the 30 differentially expressed mRNAs is presented in Fig. 5B. The differentially expressed mRNAs were mapped to terms in the GO database (Fig. 5C). With regards to biological process, the most upregulated genes were related to "cellular process", "developmental process", "single-organism process" and "metabolic process". As regards the cellular component, the most upregulated genes were related to the "cell", "cell part" and "organelle part". With regard to molecular function, the most upregulated genes were related with "binding" and "catalytic activity". The top 20 pathways of the KEGG assignments of the differentially expressed genes are shown in Fig. 5D, including the "PI3K-Akt signaling

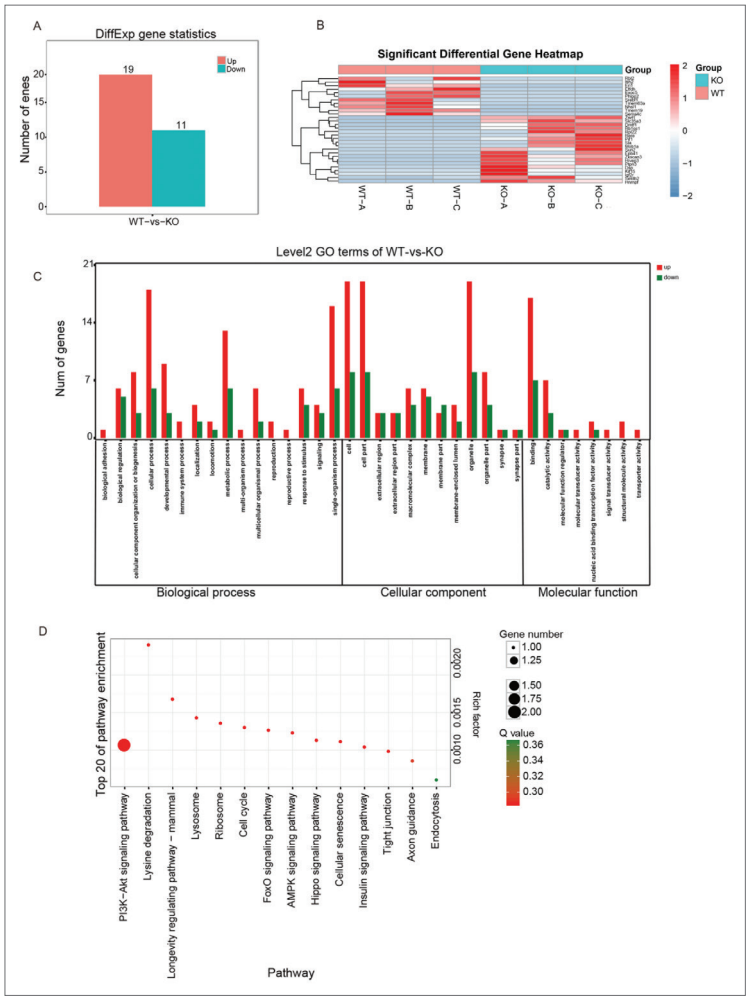


**Fig. 5**. mRNA expression profile in the liver of EpCAM$^{-/-}$ mice. **A** – A total of 19 upregulated and 11 downregulated mRNAs were identified. **B** – Heatmap of differentially expressed mRNAs. **C** – GO classification of the differentially expressed mRNAs. **D** – KEGG assignments of the differentially expressed mRNAs. EpCAM – epithelial cell adhesion molecule; GO – gene ontology; KEGG – Kyoto Encyclopedia of Genes and Genomes.

pathway", "longevity regulating pathway", "forkhead box O (FoxO) signaling pathway", "AMP-activated protein kinase (AMPK) signaling pathway", "Hippo signaling pathway", "cell cycle", "cell senescence", "insulin signaling pathway" and "tight junction".

## Validation of mRNA expression profiles in the liver of *EpCAM*$^{-/-}$ mice

A total of 30 differentially expressed mRNAs were identified by sequencing, and 15 genes were randomly selected to examine the expression level by RT-qPCR. The results showed that the expression of SET domain bifurcated 2 (Setdb2) was significantly increased (Fig. 6A; P<0.001), which was in agreement with the sequencing result. The expression of most of the other 14 genes was also consistent with the sequencing results, although it was not statistically significant. The expression of exocyst complex component 5 (*Exoc5*) and electron transfer flavoprotein dehydrogenase (*Etfdh*) was decreased, but not significantly, and the expression of heterogeneous nuclear ribonucleoprotein F (*Hnrnpf*) and cyclin-D-binding Myb-like protein 1 (*Dmtf1*) exhibited almost no change between the *EpCAM*$^{-/-}$ mouse group and the WT mouse group (Fig. 6B-6O). The results of RT-qPCR confirmed that the sequencing data were reliable. *Setdb2* is a metabolism-related gene [21], and the TG and TC levels in the liver were further examined in the present study. The results showed that the TG level was significantly decreased in the liver of the *EpCAM*$^{-/-}$ mice at the P3 stage. The TC level was also decreased in the liver of the *EpCAM*$^{-/-}$ mice, although not significantly (Supplementary Fig. S4A). There was no difference in liver weight and liver index between the EpCAM$^{-/-}$ and WT mice at the P0 stage (Supplementary Fig. S4B). The HE staining results revealed no marked differences between the livers of the EpCAM$^{-/-}$ and WT mice at P0, with the livers of both *EpCAM*$^{-/-}$ and WT mice demonstrating abundant glycogen accumulation (Supplemental Fig. S4C), similar to what has been described in previous studies [22]. At P3, there was less glycogen accumulation in the liver of the WT mice, and almost all glycogen accumulation disappeared in the liver of the *EpCAM*$^{-/-}$ mice (Supplementary Fig. S4C).

## Correlation analysis of lncRNAs and mRNAs

LncRNAs play an important role in regulating gene expression at the posttranscriptional level, and the antisense-regulation, *cis*-regulation and *trans*-regulation of lncRNAs were further analyzed in the present study. The results of the GO and KEGG enrichment analysis of the target genes of antisense-regulation, *cis*-regulation and *trans*-regulation are shown in Fig. 7. The results of the GO analysis of the target genes of the antisense-regulation, *cis*-regulation and *trans*-regulation of lncRNAs were similar to those of the

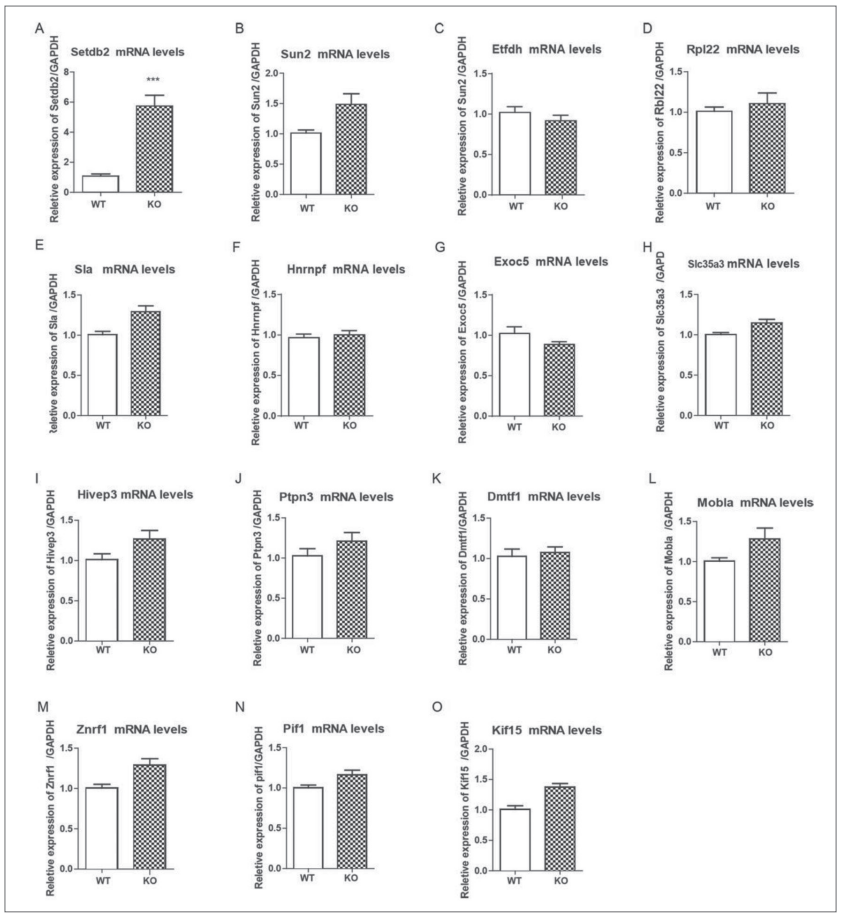


**Fig. 6**. The expression of some differentially expressed mRNAs validated by qRT-PCR. **A** – Setdb2; **B** – Sun2; **C** – Etfdh; **D** – Rpl22; **E** – Sla; **F** – Hnrnpf; **G** –Exoc5; **H** – Slc35a3; **I** – Hivep3; **J** – Ptpn3; **K** – Dmtf1; **L** – Mobla; **M** – Znrf1; **N** –Pif1; **O** – Kif15. ***P<0.001 vs. WT. Setdb2 – SET domain bifurcated 2; Sun2 – SAD1/UNC84 domain protein-2; Etfdh – electron transfer flavoprotein dehydrogenase; Rpl22 – RP ribosomal protein L22; Sla – swine leukocyte antigen; Hnrnpf – heterogeneous nuclear ribonucleoprotein F; Exoc5 – exocyst complex component 5; Slc35a3 – solute carrier family 35 member A3; Hivep3 – human immunodeficiency virus type I enhancer binding protein 3; Ptpn3 – protein tyrosine phosphatase non-receptor 3; Dmtf1 – cyclin D binding Myb-like protein 1; Mobla – MOB kinase activator 1a; Znrf1 – zinc and ring finger 1; Kif15 - Kinesin family member 15.

target genes of differentially expressed miRNAs and those of the differentially expressed mRNAs. With regard to biological process, most genes were related to "cellular process", "single-organism process" and "metabolic process". With regard to cellular component, most genes were related to the "cell", "cell part" and "organelle part". With regard to molecular function, most genes were related with "binding" and "catalytic activity" (Fig. 7A-C). The results of the KEGG analysis of the target genes of the antisense-regulation demonstrated that the top pathways focused on "glutamatergic synapse", "selenocompound metabolism", "oocyte meiosis", "cholinergic synapse" and "dopaminergic synapse" (Fig. 7D). The results of the KEGG analysis of the target genes of the *cis*-regulation demonstrated that the top pathways included "CAMs", the "hypoxia inducible factor (HIF)-1 signaling pathway", "hematopoietic cell lineage", "intestinal immune network for immunoglobulin A (IgA) production", "B cell receptor (BCR) signaling pathway", "mammalian target of rapamycin (mTOR) signaling pathway", "interleukin (IL)-17 signaling pathway" and "T-helper type 17 (Th17) cell differentiation" (Fig. 7E). The results of the KEGG analysis of the target genes of the trans-regulation demonstrated that the top pathways included "insulin secretion", "cAMP signaling pathway", "calcium signaling pathway", "MAPK signaling pathway", "CAMs", "regulation of lipolysis in adipocyte" (Fig. 7F).

## Regulatory networks of lncRNA-miRNA-mRNA

A total of 19 upregulated and 11 downregulated mRNAs were identified in the *EpCAM*$^{-/-}$ mouse group compared with the WT group, and the lncRNA-miRNA-mRNA regulatory network of 10 differentially expressed mRNAs associated with metabolism, cellular processes and signal transduction was analyzed (Fig. 8). Metabolism-related gene *Setdb2* was predicted to be regulated by mmu-miR-7093-3p, mmu-miR-8094 and mmu-miR-1930-5p. The expression of mmu-miR-7093-3p, mmu-miR-8094 and mmu-miR-1930-5p was examined by RT-qPCR, and the results showed that the expression of mmu-miR-8094 and mmu-miR-7093-3p was increased (Fig.2), though not significantly, and the expression of mmu-miR-1930-5p was increased significantly (Fig.2; P<0.05). The regulatory network between *Setdb2* and miRNAs requires further study. PH domain and leucine-rich repeat protein phosphatase 2 (Phlpp2) were predicted to be regulated by mmu-miR-690, mmu-miR-449a-5p

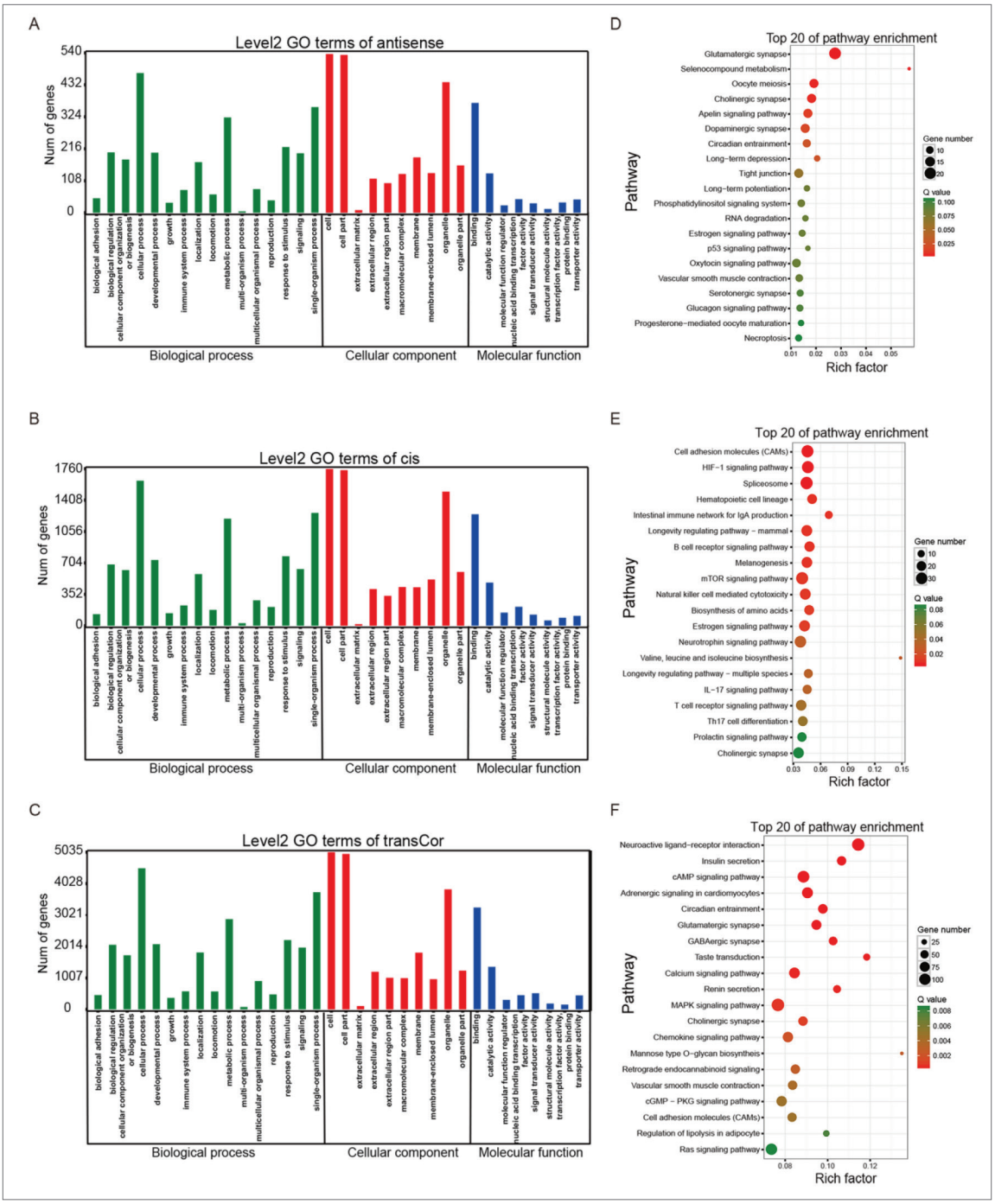


**Fig. 7**. GO and KEGG analysis of the target genes of antisense-regulation, *cis*-regulation and *trans*-regulation of lncRNAs. **A**-**C** – GO analysis of the target genes of antisense-regulation, *cis*- and *trans*-regulation of lncRNAs, respectively. **D**-**F** – KEGG analysis of the target genes of the antisense-regulation, *cis*- and *trans*-regulation of lncRNAs respectively. GO - gene ontology; KEGG - Kyoto Encyclopedia of Genes and Genomes; lncRNAs - long non-coding RNA.

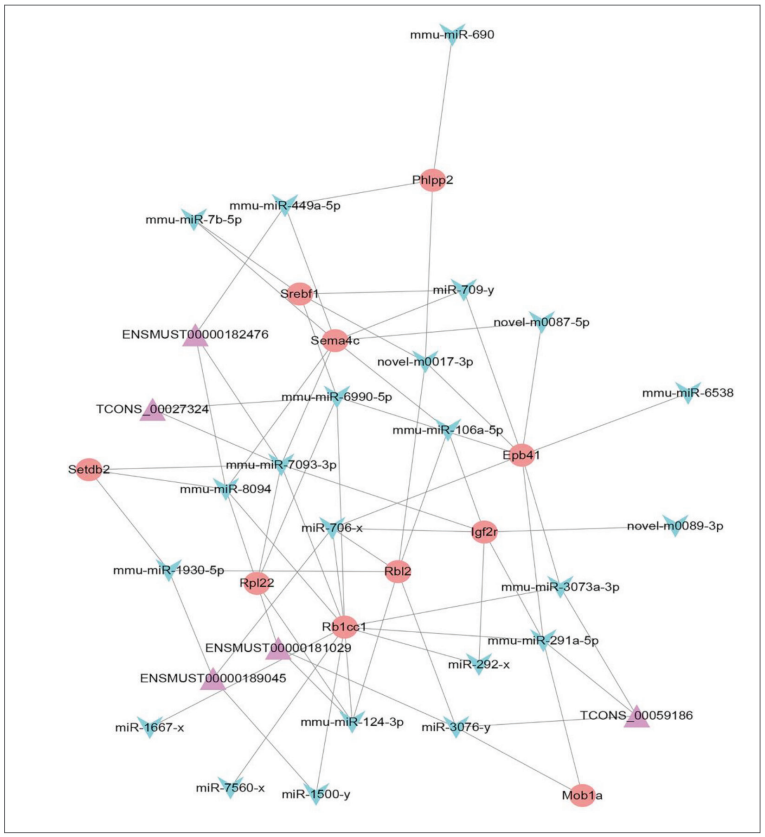


**Fig. 8**. lncRNA-miRNA-mRNA regulatory networks of 10 differentially expressed mRNAs. lncRNA – long non-coding RNA; miRNA – microRNA.

and novel-m0017-3p. Retinoblastoma-like 2 protein (Rbl2) might be regulated by novel-m0017-3p, mmu-miR-106a-5p, mmu-miR-1930-5p, mmu-miR-124-3p and miR-3076-y. MOB kinase activator 1a (Mob1a) might be regulated by mmu-miR-291a-5p and miR-3076-y. Sterol-regulatory element-binding transcription factor 1 (Srebf1) was predicted to be regulated by miR-709-y, novel-m0017-3p and mmu-miR-6990-5p. Semaphorin 4c (Sema4c) might be regulated by novel-m0087-5p, miR-709-y, mmu-miR-449a-5p, mmu-miR-7b-5p, mmu-miR-106a-5p, mmu-miR-7093-3p and mmu-miR-8094. Protein 4.1R (Epb41) was regulated by mmu-miR-106a-5p, novel-m0017-3p, miR-709-y, novel-m0087-5p, mmu-miR-6538, miR-706-x, mmu-miR-3073a-3p and mmu-miR-291a-5p. Insulin-like growth factor-2 receptor (Igf2r) was regulated by miR-706-x, TCONS_00027324, mmu-miR-106a-5p, novel-m0089-3p, miR-292-x and mmu-miR-291a-5p. RP ribosomal protein L22 (Rpl22) was regulated by mmu-miR-8094, mmu-miR-7093-3p, mmu-miR-6990-5p, ENSMUST00000181029 and mmu-miR-124-3p. RB1-inducible coiled-coil 1 (Rb1cc1) was predicted to be regulated by mmu-miR-8094, mmu-miR-7093-3p, miR-706-x, mmu-miR-6990-5p, mmu-miR-3073a-3p, mmu-miR-291a-5p, miR-292-x, mmu-miR-124-3p, miR-1500-y, miR-7560-x and miR-1667-x. The ceRNA regulatory networks in the liver of *EpCAM*$^{-/-}$ mice were predicted in Supplementary Fig. S5.

## DISCUSSION

The expression of *EpCAM* in the development of the liver is dynamic, since it is expressed in fetal liver, while mature hepatocytes do not express *EpCAM* [23]. Moreover, liver regeneration is associated with *EpCAM*, which is highly expressed in hepatic stem/progenitor cells, but whose expression is decreased during differentiation into hepatocytes [23]. It was reported that the *EpCAM*-positive population of proliferating ductal cells, which were derived from bile ductules, might be the cellular origin of HCC, suggesting the existence of stem/progenitor-derived hepatocarcinogenesis [24]. It was reported that the expression of *EpCAM* was significantly increased in the liver patients with alcoholic hepatitis (AH), as compared with the healthy control [25]. The expression of *EpCAM* was increased in HBV-mediated HCCs and, and HBV X protein, an important factor in HBV-mediated hepatocarcinogenesis, could induce the expression of *EpCAM* by activating DNA demethylation [26]. Therefore, *EpCAM* is involved in several biological and pathological processes of the liver, but the underlying mechanisms remain unclear. In the present study, the expression profiles of miRNAs, lncRNAs and mRNAs in the liver of EpCAM$^{-/-}$ mice were established, and the interactions of microRNA-lncRNA-mRNA regulatory networks were analyzed, providing important information for a systematic understanding of the functions and mechanisms of *EpCAM* in the liver.

The present results demonstrate that 24 upregulated and 13 downregulated miRNAs were identified

in the *EpCAM*$^{-/-}$ mouse group when compared with the WT group. GO enrichment analysis showed that the target genes of the differentially expressed miRNAs were related to "cellular process", "single-organism process", "metabolic process", "binding", "catalytic activity" and "signal transducer activity". The pathways of the KEGG assignments of the target genes of differentially expressed miRNAs mainly focused on the "Ras signaling pathway", "calcium signaling pathway", "TNF signaling pathway", "cAMP signaling pathway", "signaling pathways regulating pluripotency of stem cells", "Rap1 signaling pathway", "Wnt signaling pathway", "RIG-I-like receptor signaling pathway", "MAPK signaling pathway", "T cell receptor signaling pathway", "natural killer cell mediated cytotoxicity", "Th17, Th1 and Th2 cell differentiation" and "cell adhesion molecules (CAMs)". It was reported that the knockdown of *EpCAM* inhibits the growth and metastasis of breast cancer cells by inhibiting the Ras/Raf/ERK signaling pathway and matrix metalloproteinase (MMP)-9 [27]. Autocrine Wnt signaling is likely a distinct feature of *EpCAM*$^{+}$ cells from advanced cirrhosis and HCC, while it is absent from noncancerous/noncirrhotic *EpCAM*$^{+}$ cells; *EpCAM*$^{+}$ cells in advanced cirrhosis included a population of cancer stem cell (CSC)-like cells that could be used for the early diagnosis of HCC [28]. *EpCAM* could also affect the immune reaction [28]. *EpCAM* has been reported to be one of the most common tumor-associated antigens in colon cancer, and *EpCAM*-reactive Th2 cells could promote the growth of *EpCAM*-expressing tumors [29]. The sequencing results of the present study confirmed these results.

A total of 19 upregulated and 11 downregulated mRNAs were identified in the *EpCAM*$^{-/-}$ mouse group compared with the WT group, and KEGG analysis revealed that most of the differentially expressed mRNAs were focused on the "PI3K-Akt signaling pathway", "longevity regulating pathway", the "FoxO signaling pathway", "AMPK signaling pathway", "Hippo signaling pathway", "cell cycle", "cell senescence", "insulin signaling pathway" and "tight junction". The expression levels of some mRNAs were validated by RT-qPCR, and the expression of Setdb2 mRNA was significantly increased in the livers of *EpCAM*$^{-/-}$ mice compared to those of the WT. It is herein reported that the expression of *Setdb2* is significantly increased in the fasted state of mice, that its expression is regulated by glucocorticoids, and that the protein Setdb2 interacts with the nuclear glucocorticoid receptor (GR); *SETDB2* knockdown attenuated the glucocorticoid-mediated induction of several GR target genes, such as *trans*-activator of transcription (*Tat*) gene, cytochrome P450 2B10 (*Cyp2b10*), insulin induced gene 2 (*Insig2*), insulin like growth factor binding protein 1 (*Igfbp1*) and growth differentiation factor 15 (*Gdf15*); GR-Setdb2-induced *Insig2a*, contributed to the negative regulation of lipogenesis during fasting, linking glucocorticoids and GR to lipid metabolism [30]. The effects of *EpCAM* knockdown on GR signaling may be investigated using Setdb2$^{-/-}$ mice in future studies. The TG and TC levels in the liver of mice were further examined in the present study. The TG level was found to be significantly decreased in the liver of the *EpCAM*$^{-/-}$ mice, and the TC level was also clearly decreased. The H&E staining results also demonstrated an interesting phenomenon: at P0, the liver of both *EpCAM*$^{-/-}$ and WT mice exhibited abundant glycogen accumulation, similar to that described in previous studies [16,22]; at P3, glycogen accumulation was clearly reduced in the liver of WT mice and almost disappeared in the liver of *EpCAM*$^{-/-}$ mice. It is hypothesized that there was abundant glycogen accumulation in the liver of newborn mice, which were mainly dependent on glycogen accumulation in the liver for energy. When the newborn mice began to feed on breast milk, liver glycogen accumulation was no longer the main energy source for the mice and so it gradually decreased. However, the *EpCAM*$^{-/-}$ mice exhibiting intestinal defects had difficulties in digesting and absorbing food and energy, which might affect glycogen accumulation in the liver. This phenomenon and hypothesis require further investigation. The results of the KEGG analysis of the target genes of the *trans*-regulation of lncRNAs also demonstrated that the top pathways included insulin secretion and regulation of lipolysis in adipocyte. Therefore, the above results indicated that EpCAM was also involved in the metabolism of glucose and lipid in the liver, which might demonstrate a new function of EpCAM in the biological processes of the liver. However, the underlying mechanisms need further study.

A total of 10 mRNAs with clear signaling pathways were selected to analyze the lncRNA-miRNA-mRNA regulation network. Phlpp2 and Rbl2 were associated with the PI3K-Akt signaling pathway [31,32],

and silencing EpCAM was found to repress the PI3K/Akt/mTOR signaling pathway, thus inhibiting the progression of AH [33]. Mob1a was one of the core components of the Hippo pathway, and Mob1a/1b double deficiency in mouse liver could lead to the hyperplasia of oval cells, immature cholangiocytes, infiltration of inflammatory cells and fibrosis [34]. *Sema4c* is the target of several miRNAs and is involved in the progression of several types of cancer [35]. The expression of *Sema4c* was increased in the adipose tissue of patients with nonalcoholic steatohepatitis (NASH) [36]. *Srebf1* is an important gene associated with obesity and type 2 diabetes mellitus (T2DM) [37, 38], and it has been reported that bisphenol A could induce hepatic steatosis and lipid accumulation and lead to nonalcoholic fatty liver disease (NAFLD) by regulating the miR-192-Srebf1 axis [39]. The polymorphism and circulating Igf2r was associated with T2DM and could inhibit the invasion of tumorigenic liver cells by preventing the pericellular action of mannose 6-phosphate-modified proteins [40,41]. *Epb41* is an HCC susceptibility gene and the expression of *Epb41* was significantly decreased in HCC tissue as compared to normal liver tissues [42]. These mRNAs were shown to be involved in liver diseases, such as hepatitis, HCC, NAFLD and NASH, indicating that EpCAM was also involved in the processes of these diseases. Moreover, several mRNAs were associated with glucose and lipid metabolism, such as Setdb2, Sema4c, Srebf1 and Igf2r, suggesting that EpCAM might play an important role in the regulation of glucose and lipid metabolism, which is in agreement with the results of RT-qPCR, TG and TC detection and H&E staining. Our previous study also reported that the expression of certain glycogen-related genes was altered in the livers of *EpCAM*$^{-/-}$ mice [16]. The described lncRNA-miRNA-mRNA regulatory network provides further directions for studying the functions and mechanisms of EpCAM in the liver.

## CONCLUSIONS

The expression profiles of miRNAs, lncRNAs and mRNAs in the liver of *EpCAM*$^{-/-}$ mice were established, and the interactions of microRNA-lncRNA-mRNA regulatory networks were analyzed. The expression of some differentially expressed mRNAs was confirmed, and the expression of *Setdb2*, a metabolism-related gene was observed to be significantly increased in the liver of *EpCAM*$^{-/-}$ mice. The TG and TC levels were clearly decreased in the liver of EpCAM$^{-/-}$ mice as compared with that of WT mice. The sequencing data of KEGG analysis and miRNA-lncRNA-mRNA regulatory networks suggested that EpCAM might play an important role in glucose and lipid metabolism in the liver. Several functions of EpCAM in biological and pathological processes and their underlying mechanisms have not been identified, therefore the results of the present study provide important information and directions for a systematical understanding of the functions and mechanisms of EpCAM in the developing liver as well as providing important information and directions for the future study of novel targets for the treatment of liver diseases.

**Funding:** This work was supported by the National Natural Science Foundation of China (Grant Nos. 81803912, 82171855 and 31671520), and the Opening Foundation of the Key Laboratory of Regenerative Biology, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences (Grant No. KLRB201807).

**Author contributions:** Zili Lei and Yuting Lei contributed equally to this work. Yanhong Yang and Zili Lei conceived and designed the study, analyzed and interpreted the results of RNA sequencing. Yanhong Yang wrote the first draft of the manuscript and revised it critically. Yuting Lei, Guibin Chen and Li Huang established the mouse model. Yuting Lei, Shaomin Liu, Wanwan Liu, Lanxiang Yang and Huijuan Wu performed the molecular experiments, including RT-qPCR and H&E staining. Yuting Lei created the figures. All authors read and approved the final manuscript.

**Conflict of interest disclosure:** The authors declare that there is no conflict of interests.

**Data availability:** The lncRNA sequencing, miRNA sequencing, and mRNA sequencing data have been uploaded to figshare (https://doi.org/10.6084/m9.figshare.13668647). Anyone interested in the study of EpCAM and related fields can download the data.

## REFERENCES

1. Lei Z, Maeda T, Tamura A, Nakamura T, Yamazaki Y, Shiratori H, Yashiro K, Tsukita S, Hamada H. EpCAM contributes to formation of functional tight junction in the intestinal epithelium by recruiting claudin proteins. Dev Biol. 2012;371(2):136-45. https://doi.org/10.1016/j.ydbio.2012.07.005
2. Huang L, Yang Y, Yang F, Liu S, Zhu Z, Lei Z, Guo J. Functions of EpCAM in physiological processes and diseases (Review). Int J Mol Med. 2018;42(4):1771-85. https://doi.org/10.3892/ijmm.2018.3764

3. Das B, Okamoto K, Rabalais J, Marchelletta RR, Barrett KE, Das S, Niwa M, Sivagnanam M. Congenital tufting enteropathy-associated mutant of epithelial cell adhesion molecule activates the unfolded protein response in a murine model of the disease. Cells. 2020;9(4):946. https://doi.org/10.3390/cells9040946
4. Das B, Okamoto K, Rabalais J, Kozan PA, Marchelletta RR, McGeough MD, Durali N, Go M, Barrett KE, Das S, Sivagnanam M. Enteroids expressing a disease-associated mutant of EpCAM are a model for congenital tufting enteropathy. Am J Physiol Gastrointest Liver Physiol. 2019;317(5):G580-91. https://doi.org/10.1152/ajpgi.00098.2019
5. Jiang L, Shen Y, Guo D, Yang D, Liu J, Fei X, Yang Y, Zhang B, Lin Z, Yang F, Wang X, Wang K, Wang J, Cai Z. EpCAM-dependent extracellular vesicles from intestinal epithelial cells maintain intestinal tract immune balance. Nat Commun. 2016;7:13045. https://doi.org/10.1038/ncomms13045
6. Tanaka M, Okabe M, Suzuki K, Kamiya Y, Tsukahara Y, Saito S, Miyajima A. Mouse hepatoblasts at distinct developmental stages are characterized by expression of EpCAM and DLK1: drastic change of EpCAM expression during liver development. Mech Dev. 2009;126(8-9):665-76. https://doi.org/10.1016/j.mod.2009.06.939
7. Yousaf M, Tayyeb A, Ali G. Expression profiling of adhesion proteins during prenatal and postnatal liver development in rats. Stem Cells Cloning. 2017;10:21-28. https://doi.org/10.2147/SCCAA.S139497
8. Mani SKK, Zhang H, Diab A, Pascuzzi PE, Lefrançois L, Fares N, Bancel B, Merle P, Andrisani O. EpCAM-regulated intramembrane proteolysis induces a cancer stem cell-like gene signature in hepatitis B virus-infected hepatocytes. J Hepatol. 2016;65(5):888-98. https://doi.org/10.1016/j.jhep.2016.05.022
9. Zhou L, Zhu Y. The EpCAM overexpression is associated with clinicopathological significance and prognosis in hepatocellular carcinoma patients: a systematic review and meta-analysis. Int J Surg. 2018;56:274-80. https://doi.org/10.1016/j.ijsu.2018.06.025
10. Caiment F, Gaj S, Claessen S, Kleinjans J. High-throughput data integration of RNA-miRNA-circRNA reveals novel insights into mechanisms of Benzo[a]pyrene-induced carcinogenicity. Nucleic Acids Res. 2015;43(5):2525-34. https://doi.org/10.1093/nar/gkv115
11. Chen W, Liu D, Li Q, Zhu H. The function of ncRNAs in rheumatic diseases. Epigenomics. 2019;11(7):821-33. https://doi.org/10.2217/epi-2018-0135
12. Anastasiadou E, Jacob LS, Slack FJ. Non-coding RNA networks in cancer. Nat Rev Cancer. 2018;18(1):5-18. https://doi.org/10.1038/nrc.2017.99
13. Esteller M. Non-coding RNAs in human disease. Nat Rev Genet. 2011;12(12):861-74. https://doi.org/10.1038/nrg3074
14. Liang H, Yu T, Han Y, Jiang H, Wang C, You T, Zhao X, Shan H, Yang R, Yang L, Shan H, Gu Y. LncRNA PTAR promotes EMT and invasion-metastasis in serous ovarian cancer by competitively binding miR-101-3p to regulate ZEB1 expression. Mol Cancer. 2018;17(1):119. https://doi.org/10.1186/s12943-018-0870-5
15. Li Y, Huo C, Lin X, Xu J. Computational identification of cross-talking ceRNAs. Adv Exp Med Biol. 2018;1094:97-108. https://doi.org/10.1007/978-981-13-0719-5_10
16. Yang Y, Liu S, Lei Z, Chen G, Huang L, Yang F, Lei Y, Liu Y, Yang L, Liu W, Lai L, Guo J. Circular RNA profile in liver tissue of EpCAM knockout mice. Int J Mol Med. 2019;44(3):1063-77. https://doi.org/10.3892/ijmm.2019.4270
17. Langmead B, Salzberg SL. Fast Gapped-Read alignment with Bowtie 2. Nat Methods. 2012;9(4):357-9. https://doi.org/10.1038/nmeth.1923
18. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 2013;14(4):R36. https://doi.org/10.1186/gb-2013-14-4-r36
19. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc. 2012;7(3):562-78. https://doi.org/10.1038/nprot.2012.016
20. Tafer H, Hofacker IL. RNAplex: a fast tool for RNA-RNA interaction search. Bioinformatics. 2008;24(22):2657-63. https://doi.org/10.1093/bioinformatics/btn193
21. Roqueta-Rivera M, Esquejo RM, Phelan PE, Sandor K, Daniel B, Foufelle F, Ding J, Li X, Khorasanizadeh S, Osborne TF. SETDB2 links glucocorticoid to lipid metabolism through Insig2a regulation. Cell Metab. 2016;24(3):474-84. https://doi.org/10.1016/j.cmet.2016.07.025
22. Choi E, Zhang X, Xing C, Yu H. Mitotic checkpoint regulators control insulin signaling and metabolic homeostasis. Cell. 2016;166(3):567-81. https://doi.org/10.1016/j.cell.2016.05.074
23. Dollé L, Theise ND, Schmelzer E, Boulter L, Gires O, Grunsven LA. EpCAM and the biology of hepatic stem/progenitor cells. Am J Physiol Gastrointest Liver Physiol. 2015;308(4):G233-50. https://doi.org/10.1152/ajpgi.00069.2014
24. Matsumoto T, Takai A, Eso Y, Kinoshita K, Manabe T, Seno H, Chiba T, Marusawa H. Proliferating EpCAM-positive ductal cells in the inflamed liver give rise to hepatocellular carcinoma. Cancer Res. 2017;77(22):6131-43. https://doi.org/10.1158/0008-5472.CAN-17-1800
25. Sancho-Bru P, Altamirano J, Rodrigo-Torres D, Coll M, Millán C, Lozano JJ, Miquel R, Arroyo V, Caballería J, Ginès P, Bataller R. Liver progenitor cell markers correlate with liver damage and predict short-term mortality in patients with alcoholic hepatitis. Hepatology. 2012;55(6):1931-41. https://doi.org/10.1002/hep.25614
26. Fan H, Zhang H, Pascuzzi PE, Andrisani O. Hepatitis B virus X protein induces EpCAM expression via active DNA demethylation directed by RelA in complex with EZH2 and TET2. Oncogene. 2016;35(6):715-26. https://doi.org/10.1038/onc.2015.122
27. Gao J, Liu X, Yang F, Liu T, Yan Q, Yang X. By inhibiting Ras/Raf/ERK and MMP-9, knockdown of EpCAM inhibits breast cancer cell growth and metastasis. Oncotarget. 2015;6(29):27187-98. https://doi.org/10.18632/oncotarget.4551
28. Khosla R, Rastogi A, Ramakrishna G, Pamecha V, Mukhopadhyay A, Vasudevan M, Sarin SK, Trehanpati N. EpCAM+ liver cancer stem-like cells exhibiting autocrine Wnt signal-

ing potentially originate in cirrhotic patients. Stem Cells Transl Med. 2017;6(3):807-18. https://doi.org/10.1002/sctm.16-0248
29. Ziegler A, Heidenreich R, Braumüller H, Wolburg H, Weidemann S, Mocikat R, Röcken M. EpCAM, a human tumor-associated antigen promotes Th2 development and tumor immune evasion. Blood. 2009;113(15):3494-502. https://doi.org/10.1182/blood-2008-08-175109
30. Roqueta-Rivera M, Esquejo RM, Phelan PE, Phelan PE, Sandor K, Daniel B, Foufelle F, Ding J, Li X, Khorasanizadeh S, Osborne TF. SETDB2 links glucocorticoid to lipid metabolism through insig2a regulation. Cell Metab. 2016;24(3):474-84. https://doi.org/10.1016/j.cmet.2016.07.025
31. Yan X, Li W, Yang L, Dong W, Chen W, Mao Y, Xu P, Li D, Yuan H, Li Y. MiR-135a protects vascular endothelial cells against ventilator-induced lung injury by inhibiting PHLPP2 to activate PI3K/Akt pathway. Cell Physiol Biochem. 2018;48(3):1245-58. https://doi.org/10.1159/000492010
32. Pentimalli F, Forte IM, Esposito L, Indovina P, Iannuzzi CA, Alfano L, Costa C, Barone D, Rocco G, Giordano A. RBL2/p130 is a direct AKT target and is required to induce apoptosis upon AKT inhibition in lung cancer and mesothelioma cell lines. Oncogene. 2018;37(27):3657-71. https://doi.org/10.1038/s41388-018-0214-3
33. Zhang Z, Wen H, Weng J, Feng L, Liu H, Hu X, Zeng F. Silencing of EPCAM suppresses hepatic fibrosis and hepatic stellate cell proliferation in mice with alcoholic hepatitis via the PI3K/Akt/mTOR signaling pathway. Cell Cycle. 2019;18(18):2239-54. https://doi.org/10.1080/15384101.2019.1642067
34. Nishio M, Sugimachi K, Goto H, Wang J, Morikawa T, Miyachi Y, Takano Y, Hikasa H, Itoh T, Suzuki SO, Kurihara H, Aishima S, Leask A, Sasaki T, Nakano T, Nishina H, Nishikawa Y, Sekido Y, Nakao K, Shin-Ya K, Mimori K, Suzuki A. Dysregulated YAP1/TAZ and TGF-β signaling mediate hepatocarcinogenesis in Mob1a/1b-deficient mice. Proc Natl Acad Sci U S A. 2016;113(1):E71-80. https://doi.org/10.1073/pnas.1517188113
35. Jing L, Bo W, Yourong F, Tian W, Shixuan W, Mingfu W. Sema4C mediates EMT inducing chemotherapeutic resistance of miR-31-3p in cervical cancer cells. Sci Rep. 2019;9(1):17727. https://doi.org/10.1038/s41598-019-54177-z
36. Estep JM, Goodman Z, Sharma H, Younossi E, Elarainy H, Baranova A, Younossi Z. Adipocytokine expression associated with miRNA regulation and diagnosis of NASH in obese patients with NAFLD. Liver Int. 2015;35(4):1367-72. https://doi.org/10.1111/liv.12555
37. Zhang S, Lin X, Lynn H, Xu G, Li J, Zhao C, Li M. Dietary cholesterol interacts with SREBF1 to modulate obesity in Chinese children. Mol Nutr Food Res. 2017;61(9):1700105. https://doi.org/10.1002/mnfr.201700105
38. Krause C, Sievert H, Geißler C, Grohs M, Gammal ATE, Wolter S, Ohlei O, Kilpert F, Krämer UM, Kasten M, Klein C, Brabant GE, Mann O, Lehnert H, Kirchner H. Critical evaluation of the DNA-methylation markers ABCG1 and SREBF1 for type 2 diabetes stratification. Epigenomics. 2019;11(8):885-97. https://doi.org/10.2217/epi-2018-0159
39. Lin Y, Ding D, Huang Q, Liu Q, Lu H, Lu Y, Chi Y, Sun X, Ye G, Zhu H, Wei J, Dong S. Downregulation of miR-192 causes hepatic steatosis and lipid accumulation by inducing SREBF1: novel mechanism for bisphenol A-triggered non-alcoholic fatty liver disease. Biochim Biophys Acta Mol Cell Biol Lipids. 2017;1862(9):869-82. https://doi.org/10.1016/j.bbalip.2017.05.001
40. Chanprasertyothin S, Jongjaroenprasert W, Ongphiphadhanakul B. The association of soluble IGF2R and IGF2R gene polymorphism with type 2 diabetes. J Diabetes Res. 2015;2015:216383. https://doi.org/10.1155/2015/216383
41. Puxbaum V, Nimmerfall E, Bäuerl C, Taub N, Blaas P, Wieser J, Mikula M, Mikulits W, Ng KM, Yeoh GCT, Mach L. M6P/IGF2R modulates the invasiveness of liver cells via its capacity to bind mannose 6-phosphate residues. J Hepatol. 2012;57(2):337-43. https://doi.org/10.1016/j.jhep.2012.03.026
42. Yang X, Yu D, Ren Y, Wei J, Pan W, Zhou C, Zhou L, Liu Y, Yang M. Integrative functional genomics implicates EPB41 dysregulation in hepatocellular carcinoma risk. Am J Hum Genet. 2016;99(2):275-86. https://doi.org/10.1016/j.ajhg.2016.05.029

## Supplementary Material:

The Supplementary Material is available at:
https://www.serbiosoc.org.rs/NewUploads/Uploads/Lei%20et%20al_7306_Supplementary%20Material.pdf