Analyses of a Panel of Transcripts Identified From a Small Sample Size and Construction of RNA Networks in Hepatocellular Carcinoma

Zhiyong Sheng[1,2], Xiaolin Wang[2], Geliang Xu[1]*, Ge Shan[2,3]* and Liang Chen[2]*

[1] Anhui Province Key Laboratory of Hepatopancreatobiliary Surgery, First Affiliated Hospital of University of Science and Technology of China, Hefei, China, [2] Hefei National Laboratory for Physical Sciences at Microscale, The Chinese Academy of Sciences (CAS) Key Laboratory of Innate Immunity and Chronic Disease, School of Life Sciences, University of Science and Technology of China, Hefei, China, [3] Chinese Academy of Sciences (CAS) Centre for Excellence in Molecular Cell Science, Shanghai Institutes for Biological Sciences, CAS, Shanghai, China

Hepatocellular carcinoma (HCC) is one of the most common cancers in the world. Dysregulation of mRNAs and non-coding RNAs (ncRNAs) plays critical roles in the progression of HCC. Here, we investigated HCC samples by RNA-seq and identified a series of dysregulated RNAs in HCC. Various bioinformatics analyses established long non-coding RNA (lncRNA)-mRNA co-expression and competing endogenous RNA (ceRNA) networks in circRNA-miRNA-mRNA axis, indicating the potential *cis* and/or *trans* regulatory roles of lncRNAs and circRNAs. Moreover, GO pathway analysis showed that these identified RNAs were associated with many biological processes that were related to tumorigenesis and tumor progression. In conclusion, we systematically established functional networks of lncRNA-mRNA, circRNA-miRNA-mRNA to further unveil the potential interactions and biological processes in HCC. These results provide further insights into gene expression network of HCC and may assist future diagnosis of HCC.

Keywords: hepatocellular carcinoma, lncRNA, miRNA, circRNA, network

INTRODUCTION

Hepatocellular carcinoma (HCC) is one of the most frequent malignancies worldwide, and the highest incidence rates in the world are reported in Asia and Africa, with China alone accounting for half of the cases in Asia and Africa (McGlynn and London, 2011). Clinical investigation shows that HCC is the sixth most common cancer and third main cause of cancer mortality in the world (Forner et al., 2018). The overall 5 years survival rate is less than 20% due to lack of early and timely detection and treatment (Lee et al., 2012). So far, the available treatment approaches for HCC mainly include resection, liver transplantation, image-guided tumor ablation, and systemic therapy. The 5 years survival rates of 60–70% can be achieved in early-stage HCC patients (Llovet and Bruix, 2003; Llovet et al., 2003). However, systematic therapies cannot improve the survival rates of the patients with advanced stage of HCC due to limited specific and effective biomarkers or targets for clinical treatments (Llovet and Bruix, 2003; Llovet et al., 2003; Marquardt et al., 2012).

Dysregulated gene expression is a common theme underneath human disease (Adams and Cory, 2007; Aguirre-Ghiso, 2007; Amaravadi and Thompson, 2007; Azam et al., 2010; Hanahan and Weinberg, 2011). The resulted abnormal mRNAs and non-coding RNAs (ncRNAs) play roles in the occurrence and progress of diseases such as cancers (Prensner et al., 2011; Kung et al., 2013; Iyer et al., 2015; Rusk, 2015; Klingenberg et al., 2017). Many mRNAs are encoded by tumor genes or tumor suppressor genes, and ncRNAs such as microRNAs (miRNAs), long ncRNAs (lncRNAs), and circular RNAs (circRNAs) are also known to be associated with tumorigenesis and tumor progression (Panzitt et al., 2007; Wang et al., 2010; Wei et al., 2012; Zhang et al., 2012; Li G. et al., 2015; Palazzo and Lee, 2015; Zhou et al., 2016; Cui et al., 2017; Shen et al., 2018; Zhang and Wu, 2018).

In HCC, tumor genes such as procollagen-lysine, 2-oxoglutarate 5-dioxygenase 3 (PLOD3), and splicing factor 3b subunit 4 (SF3B4) are overexpressed in HCC tumor tissues compared with adjacent tissues, and knocking down their expression levels can effective suppress tumor invasion (Shen et al., 2018). Conversely, tumor suppressor genes such as Alcohol dehydrogenase 4 (ADH4) is significantly downregulated in tumor tissues, and patients with lower ADH4 expression levels have worse prognosis and lower overall survival rate (Wei et al., 2012).

In HCC, many miRNAs are found to play oncogenic or tumor suppressive roles (Zhang et al., 2012; Li G. et al., 2015; Zhou et al., 2016). For example, miR-421 can promote proliferation and migration in HCC cell by downregulating farnesoid X receptor (FXR), which is usually highly expressed in normal liver (Zhang et al., 2012). On the other hand, miR-708 is markedly downregulated in HCC tumor tissues compared with adjacent tissues. Low miR-708 level is related to tumor node metastasis (TNM) in advanced stage patients, and over expression of miR-708 can suppress the invasion and migration of in HCC cell lines *in vitro* (Li G. et al., 2015).

LncRNAs are also correlated to HCC (Chen et al., 2018). The lncRNA HULC is upregulated in HCC with high specificity (Panzitt et al., 2007; Wang et al., 2010). It can be detected both in tumor tissues and blood of HCC patients. Aberrantly elevated HULC promotes HCC invasion and progression by increasing expression of the oncogene HMGA2. The lncRNA PVT1 enhances the tumorigenesis and metastasis of HCC by acting as a competing endogenous RNA (ceRNA) for miR-186-5p, and the knockdown of PVT1 can effectively impede HCC invasion (Karreth et al., 2015; Cui et al., 2017).

CircRNAs are a large class of transcripts in multicellular organisms with emerging importance (Chen and Shan, 2015; Chen et al., 2015; Haque and Harries, 2017). One of functional mechanisms of circRNAs is acting as ceRNA (Hansen et al., 2013; Memczak et al., 2013). For example, the circRNA CDR1as functions as sponge of miR-7 in neuronal tissues (Memczak et al., 2013; Piwecka et al., 2017). Alterations in circRNA expression are found to be related to aberrant physiology and human diseases (Hansen et al., 2013; Memczak et al., 2013; Chen and Shan, 2015; Chen et al., 2015; Holdt et al., 2016; Piwecka et al., 2017).

In this study, we have identified differentially expressed mRNAs, miRNAs, lncRNAs, and circRNAs in fresh HCC tissues through high-throughput RNA sequencing. Functional networks of lncRNA-mRNA and circRNA-miRNA-mRNA have been established to provide new insights for biomarkers and treatments in HCC.

MATERIALS AND METHODS

Clinical Samples
All fresh HCC patient tumor samples and adjacent tissues were collected from The First Affiliated Hospital of University of Science and Technology of China, which was approved by the Human Research Ethics Committee of University of Science and Technology of China (USTCEC201700007). Written informed consent was obtained from each patient for this study. All samples were rinsed with DEPC water and then kept in RNAhold (TransGen) within 30 min after removing from the operation. HCC patient tumor sample and adjacent tissue pairs were collected from 21 patients (12 males and 9 females with advanced stage HCC, all of them were HBsAg positive, and did not have anti-tumor therapy before surgery).

Total RNA Extraction
The clinical samples were cut into small pieces and homogenized in TRizol reagent (Life Technologies) by homogenizer. Total RNA was extracted by using TRizol reagent according to the manufacturer's instructions.

Transcriptome Data Analysis
Total RNAs from four pairs of HCC patient tumor tissues were extracted for high-throughput sequencing. Whole transcriptome libraries were constructed by the TruSeq Ribo Profile Library Prep Kit (Illumina, United States), according to the manufacturer's instructions. In brief, 10 μg total RNA was depleted rRNA with an Illumina Ribo-Zero Gold kit and purified for end repair and 5′-adaptor ligation. Then, reverse transcription was performed with random primers containing 3′ adaptor sequences and randomized hexamers. Finally, the cDNAs were purified and amplified with thermo cycler. The PCR products of 300–500 bp were purified, quantified and stored at -80°C before sequencing. The libraries were subjected to 151 nt paired-end sequencing with an Illumina Nextseq 500 system (Novogene, China). Each library was generated a depth of 50–100 million read pairs and then adapters were removed with cutadapt to obtain clean reads. For mRNA and lncRNA analyses, the expression levels were calculated using TopHat2 and Cufflinks followed by the annotation references of Refseq and Ensemble transcript databases with the genome release Homo sapiens, hg19. The differentially expressed mRNA and lncRNA were determined by DEseq2 with the corresponding cutoff ($P < 0.001$, RPKM $\geq$ 10, $|\log2(\text{fold change})| \geq 1$ for mRNA and $P < 0.05$, RPKM $\geq$ 1, $|\log2(\text{fold change})| \geq 1$ for lncRNA). For circular RNA (circRNA) prediction, we identified the candidates with find_circ (Memczak et al., 2013) and the junction reads were calculated as Transcripts Per Kilobase Million (TPM). A criterion of $P \leq 0.05$, TPM ≥ 0.1 and $|\log2(\text{fold change})| \geq 1$ among four pairs of samples was used to identify differentially expressed

TABLE 1 | Primers for RT-qPCR validation.

Name	Primer sequence 5′–3′	Product size (bp)
C12orf75	F: GCAGGAGCAGCCAAAGATGT	113
	R: CTGGACACCATATTGACAGC	
ZIC5	F: AGCCCTTCAAGGCCAAATAC	129
	R: GCTTGTGGATCTTGAGGTTC	
PZP	F: TCAAGCTAGAAGCTGGCATC	137
	R: AAGCACAAATTCCTCCACGG	
FAM65C	F: TGACATCGCCGACTTCTTCA	177
	R: TACAAGGAGCCCTTCCTGCT	
C1QL1	F: CACCTACCATGTCCTCATGC	120
	R: GCGTAGTCGTAGTTCTGGTC	
TMEM74	F: GGAGCTCGCCATCAGCAAAA	142
	R: TTCCCTGAAGACGTGGCTTC	
GNAZ	F: CATCGCCGCAGCTGACTATA	132
	R: CCACGTCCACCATCTTGAAG	
LNC-TOB2P1	F: GTGACTACAGGAAGGTTAGC	161
	R: GAGAAGAGGCTGCATGAGAG	
LOC100499489	F: TAGAGATGAGACGATGCAGG	110
	R: CAGCTCCACCCACGTGGTTT	
LINC01093	F: AAGAGACTGGAGCCCTGAGT	92
	R: GCTGTAACCCCAAAAACTCC	
LOC100130899	F: GAGGTTTGCTGCAAATCAGG	101
	R; ACATACTCACGGCTGTGCCT	
LOC200772	F: GAGGTTTGCTGCAAATCAGG	135
	R; ACATACTCACGGCTGTGCCT	
LNC-FENDRR	F: AGGACCATTCAGCTCTAGAG	141
	R: ATGGGGTGAGCAAAGCGCAT	
LNC-NBPF22P	F: CCATGTCACCACCAAGAACA	152
	R: CCATGTCACCACCATCAATG	
LNC-NMRAL2p	F: TGCAACACCGAAGGATTTCC	148
	R: CAAGTTCCGAGGCTTCTCAT	
ACTB	F: CCAACACAGTGCTGTCTGG	130
	R: GAGTACTTGCGCTCAGGAG	
PLOD3	F: GTTTGTGGATAGCTACGACG	72
	R: CCACTCTGGACGAACTTCTT	
SF3B4	F: GGACCCTGAGATTGATGAGA	91
	R: GTCAGGGTCCCGCATAATTT	
SNHG7	F: TGCTCACTGGAGATGACACG	92
	R: CACTGGAAGTCCATCACAGG	
HULC	F: AGGAAGAGTCGTCACGAGAA	124
	R: TATTCCGGCCTTTACTTCAG	
ADH4	F: ATCAACAATGCCAAGGTCAC	113
	R: CCTATGATTCTGGAAGCTCC	
COLEC10	F: GGCAATATTGGCAAGACTGG	77
	R: TCTCCAGGTATTCCAAGCAA	
circAKR1B10	F: CCTTGTGAGGAAAGCCTTTG	199
	R: CAGCACACTTAGAGGAAGCT	
circAKR1C3	F: AGGATTGGCCAAGTCCATCG	211
	R: CCATCGTTTGTCTCGTTGAG	
circHMGCS1	F: CTCGGATGTTGCTGAATGAC	197
	R: CGGTCTAATGCACTGAGGTA	
circC3P1	F: GGTTGCAAGAACAGGAATCG	137
	R: TCACATGCTGTTGTCTGGAG	
miR-10b-3p	SL: GTCGTATCCAGTGCAGGGTCCGAG	
	GTATTCGCACTGGATACGACATTCCC	
	F: CACGCAACAGATTCGATT	
	R: CCAGTGCAGGGTCCGAGGTA	
miR-421	SL GTCGTATCCAGTGCAGGGTCCGAGGT	

(Continued)

TABLE 1 | Continued

Name	Primer sequence 5′–3′	Product size (bp)
	ATTCGCACTGGATACGACGCGCCC	
	F: CACGCAATCAACAGACATTA	
	R: CCAGTGCAGGGTCCGAGGTA	
miR-761	SL GTCGTATCCAGTGCAGGGTCCGAGG	
	TATTCGCACTGGATACGACTGTGTC	
	F: CACGCAGCAGCAGGGTGAAA	
	R: CCAGTGCAGGGTCCGAGGTA	
miR-200a-3p	SL: GTCGTATCCAGTGCAGGGTCCGA	
	GGTATTCGCACTGGATACGACACATCG	
	F: CACGCATAACACTGTCTGGT	
	R: CCAGTGCAGGGTCCGAGGTA	
miR-200b-3p	SL: GTCGTATCCAGTGCAGGGTCCGAG	
	GTATTCGCACTGGATACGACTCATCA	
	F: CACGCATAATACTGCCTGGT	
	R: CCAGTGCAGGGTCCGAGGTA	
miR-139-5p	SL: GTCGTATCCAGTGCAGGGTCCGAG	
	GTATTCGCACTGGATACGACACTGGA	
	F: CACGCATCTACAGTGCACGT	
	R: CCAGTGCAGGGTCCGAGGTA	

SL, Stem loop RT primer.

circRNAs. RNA-seq data were deposited in NCBI with the GEO accession code GSE128274.

Small RNA Data Analysis

For small RNA (sRNA) sequencing, eight sRNA libraries were generated with TruSeq small RNA (Illumina, United States) according to the manufacturer's instructions. Then the prepared libraries were sequenced with an Illumina Nextseq 500 system (Novogene, China). After filtering out the reads shorter than 15 nt, the remaining reads were mapped to the human genome (hg19) and the miRNA database in miRBase with bowtie (-v 1). The differentially expressed miRNAs were determined by DEseq2 with the cutoff of $P < 0.05$, TPM ≥ 1, $|\log2(\text{fold change})| \geq 1$.
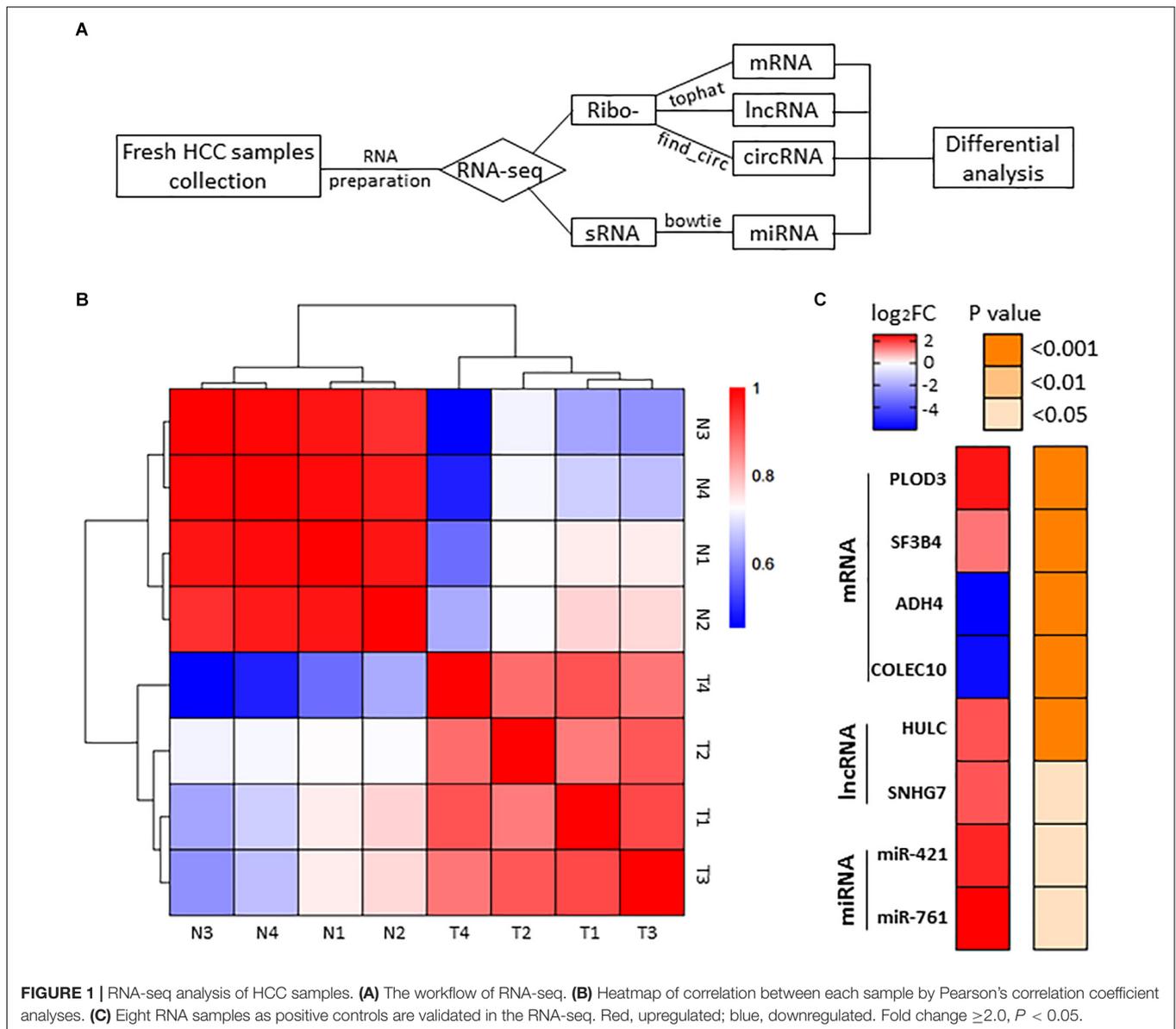
Construction of Co-expression and CeRNA Network

For the co-expression network of significantly dysregulated lncRNAs and mRNAs, Pearson's correlations were calculated the co-expression analysis according to the expression levels in eight samples. A criterion of the coefficient parameter R-squared more than 0.99 was used for the remaining RNAs to further construct the network. For the competing endogenous RNAs (ceRNA) network of significantly dysregulated circRNAs and mRNAs, the miRNA/mRNA and miRNA/circRNA interaction were predicted with TargetScanHuman7.2. The above networks were both performed with Cytoscape[1].

GO Analysis

The significantly dysregulated mRNAs in co-expression and ceRNA network were both analyzed using GOrilla web-server with default parameters (Eden et al., 2009).

---

[1]www.cytoscape.org

FIGURE 1 | RNA-seq analysis of HCC samples. **(A)** The workflow of RNA-seq. **(B)** Heatmap of correlation between each sample by Pearson's correlation coefficient analyses. **(C)** Eight RNA samples as positive controls are validated in the RNA-seq. Red, upregulated; blue, downregulated. Fold change ≥2.0, $P < 0.05$.

Statistical Analysis

In all experiments, Student's t-tests were used to calculate P-values, as indicated in the figure legends. The values reported in the graphs represent averages of actual number of independent experiments, with error bars showing SD. After analysis of variance with F-tests, the statistical significance and P-values were evaluated with Student's t-tests.

## Reverse Transcription and Real-Time Quantitative PCR

cDNA was prepared using GoScript Reverse Transcription System (Promega) according to the manufacturer's protocol. Notably, for the first strand cDNA synthesis of miRNA, stem-loop method was used (Kramer, 2011). Quantitative real-time PCR was performed with GoTaq SYBR Green qPCR

Master Mix (Promega) on a PikoReal 96 real-time PCR system followed by 40 amplification cycles (Thermo Fisher Scientific) according to standard procedures. Actually, all amplification curves already reached stationary stage before 35 amplification cycles, and the readings of Ct value were obtained at the exponential stage. Relative RNA expression was normalized to ACTB expression level. All primers are shown in **Table 1**.
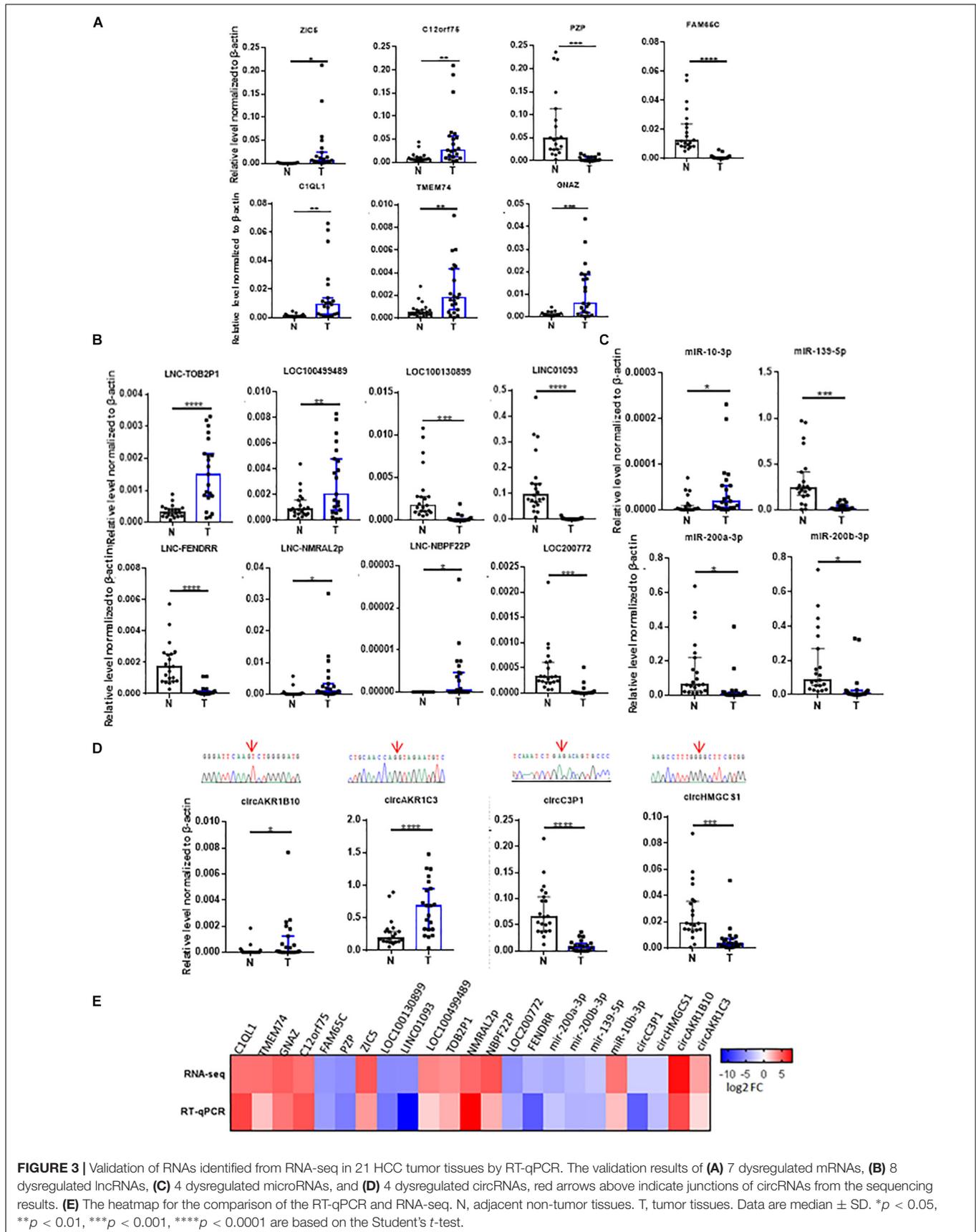
RESULTS
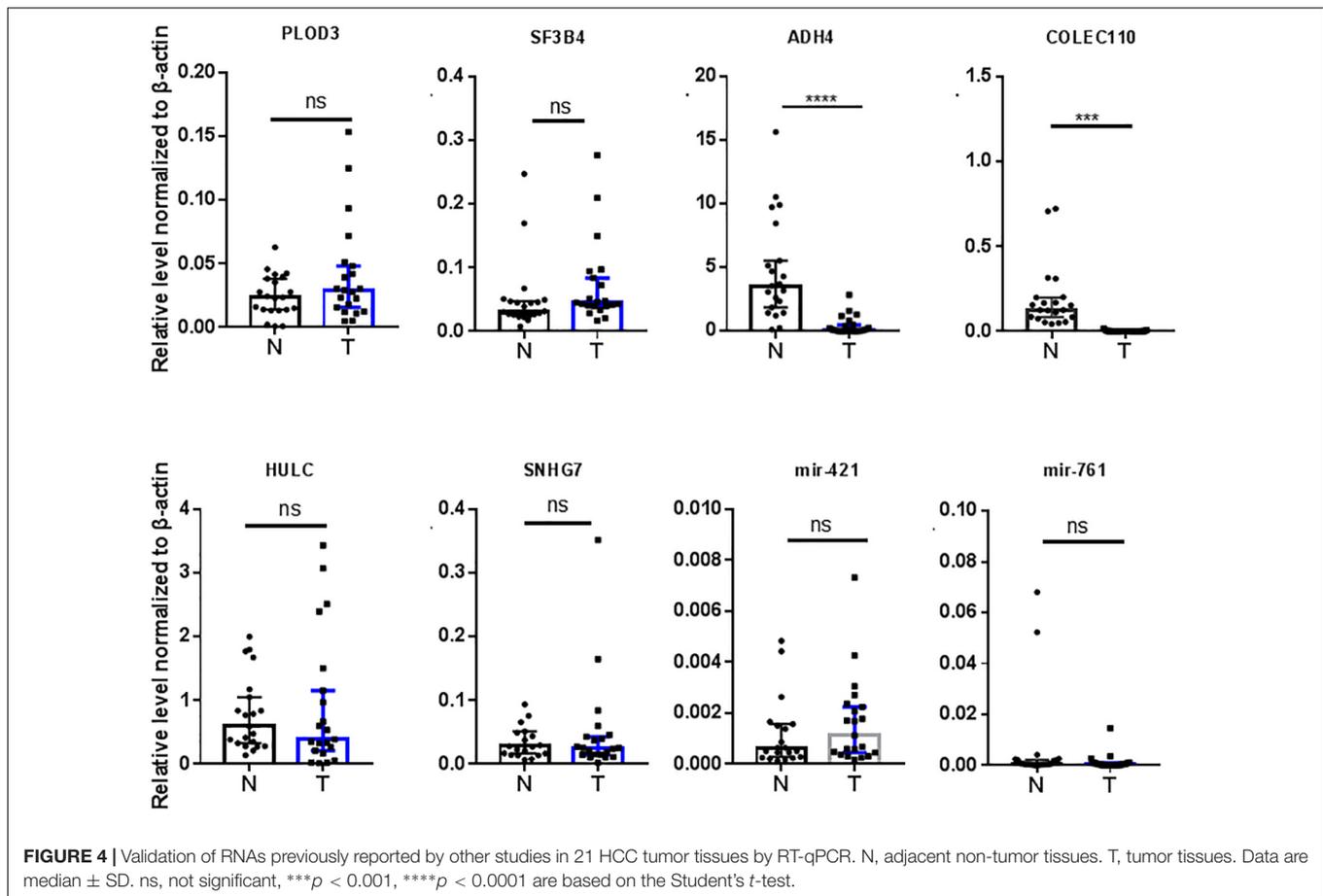
Identification of Differentially Expressed RNAs in HCC Samples

To identify HCC-related RNAs, we used four HCC patients' fresh tumor tissues and paired adjacent non-tumor tissues for RNA sequencing. Differential expression of mRNAs, lncRNAs,
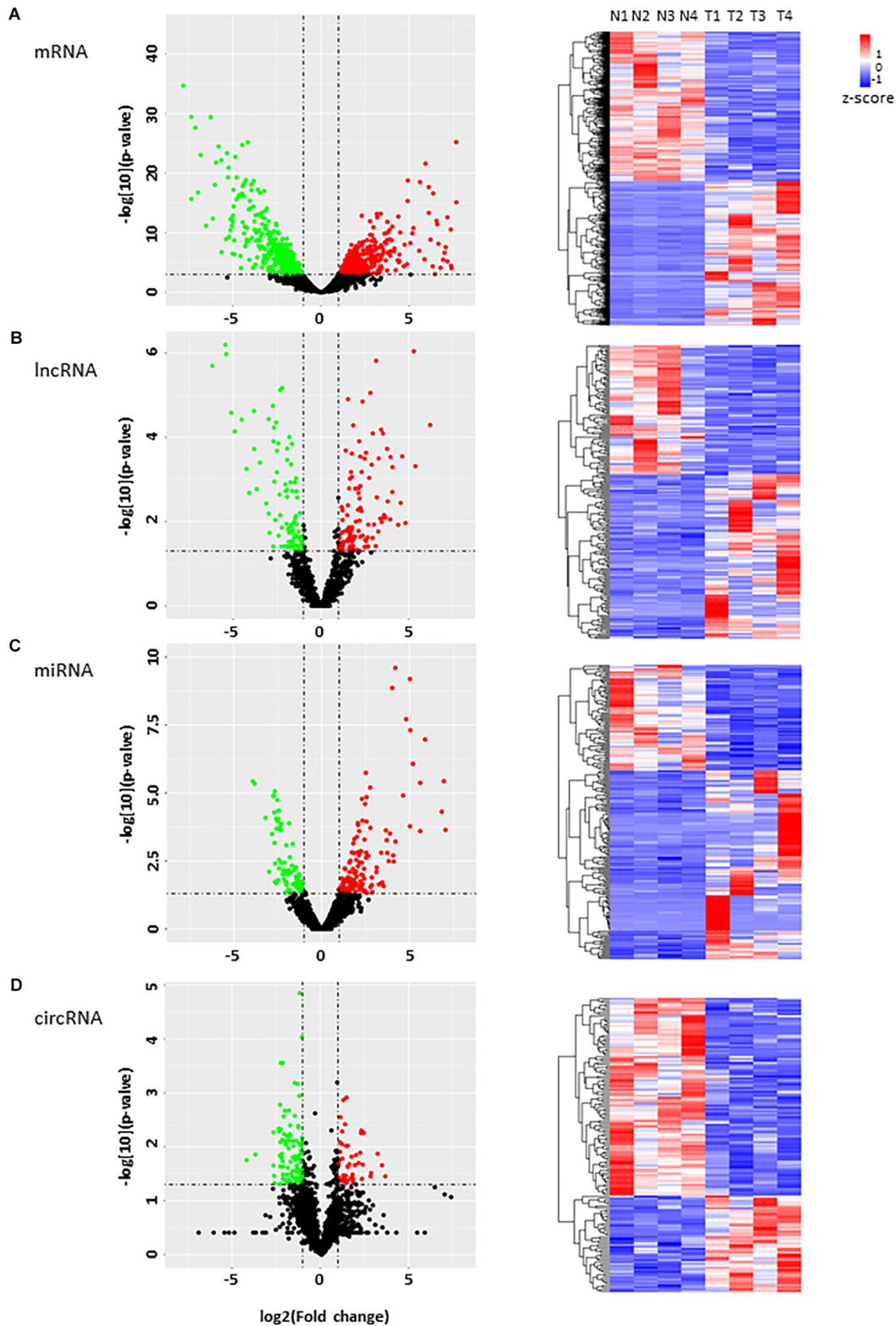
FIGURE 2 | Analysis of the RNA expression profile in HCC. **(A–D)** volcano plots (left) display differentially expressed transcripts, and the hierarchical cluster analyses on the right of each panel display differential expression of RNAs in four HCC tumor tissues and their controls. **(A)** mRNAs, **(B)** lncRNAs, **(C)** miRNAs, and **(D)** circRNA. Green and red dots represent downregulated and upregulated RNAs respectively in tumor tissues compared with adjacent tissues. Black dots indicate no significant difference.

FIGURE 3 | Validation of RNAs identified from RNA-seq in 21 HCC tumor tissues by RT-qPCR. The validation results of **(A)** 7 dysregulated mRNAs, **(B)** 8 dysregulated lncRNAs, **(C)** 4 dysregulated microRNAs, and **(D)** 4 dysregulated circRNAs, red arrows above indicate junctions of circRNAs from the sequencing results. **(E)** The heatmap for the comparison of the RT-qPCR and RNA-seq. N, adjacent non-tumor tissues. T, tumor tissues. Data are median ± SD. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$ are based on the Student's t-test.

FIGURE 4 | Validation of RNAs previously reported by other studies in 21 HCC tumor tissues by RT-qPCR. N, adjacent non-tumor tissues. T, tumor tissues. Data are median ± SD. ns, not significant, ***p < 0.001, ****p < 0.0001 are based on the Student's t-test.

circRNAs, and miRNAs were then analyzed (**Figure 1A**). Pearson's correlation coefficient analysis showed that tumor tissues were positively correlated with each other, and control tissues also showed strong positive correlation (**Figure 1B**). It seems evident that the 4 tumor samples differ more from normal samples than among themselves, maybe due to the tumor heterogeneity. We noticed that previously reported transcripts with aberrant levels in HCC such as mRNAs (PLOD3, SF3B4, ADH4, and COLEC10), lncRNAs (HULC and SNHG7), and miRNAs (miR-421 and miR-761) were also identified in our RNA-seq (**Figure 1C**; Panzitt et al., 2007; Wang et al., 2010; Wei et al., 2012; Zhang et al., 2012; Li G. et al., 2015; Zhou et al., 2016; Shen et al., 2018; Zhang and Wu, 2018).
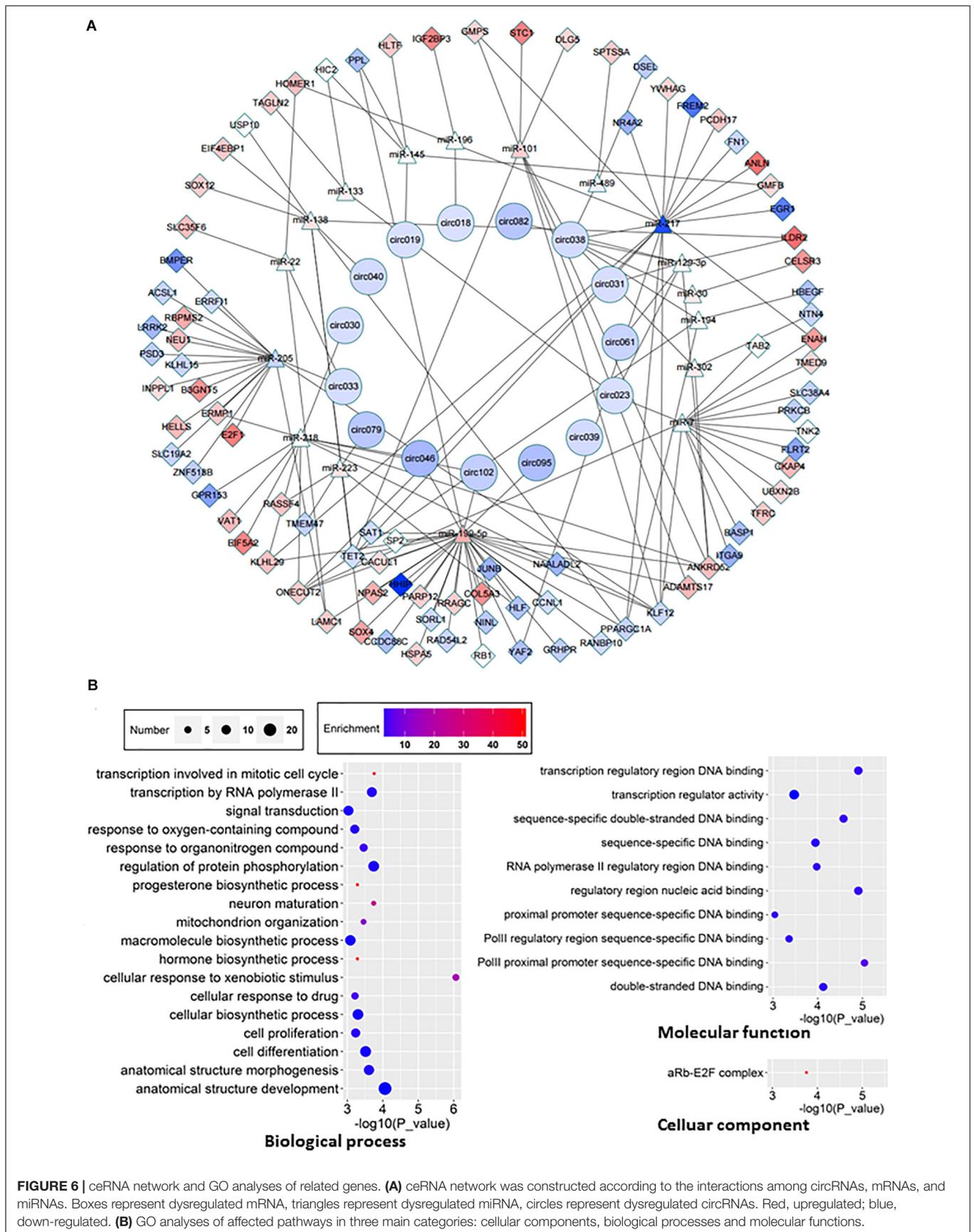
Next, we performed volcano plots for mRNAs, lncRNAs, miRNAs and circRNAs in HCC paired tissues (fold change ≥ 2.0, P < 0.05) (**Figures 2A–D**). 919 differentially expressed mRNAs, 207 lncRNAs, 216 miRNAs, and 152 circRNAs were identified (**Figures 2A–D**). Among them, 452 mRNAs (49.18%), 116 lncRNAs (50.04%), 138 miRNAs (63.89%), and 50 circRNAs (32.89%) were upregulated, while 467 mRNAs (50.82%), 91 lncRNAs (43.96%), 78 miRNAs (36.11%), and 102 circRNAs (67.11%) were downregulated (**Figures 2A–D**). We also conducted a hierarchical cluster analysis to display the differential expression of four types of RNAs across eight samples (**Figures 2A–D**). Tumor and adjacent

non-tumor samples respectively were classified into different branch (**Figures 2A–D**).

RT-qPCR Verification of the Dysregulated RNAs in Clinical Samples

To validate the results of the RNA-seq, we selected 23 dysregulated RNA candidates including mRNAs, lncRNAs, miRNAs and circRNAs for RT-qPCR validation with 21 HCC patient sample pairs (**Figures 3A–D**). Housekeeping gene ACTB was used as the endogenous control. In the mRNA group, ZIC5, C12orf75, C1QL1, TMEM74, and GNAZ were significantly upregulated in tumor tissues compared to the adjacent control; PZP and FAM65C were significantly downregulated compared to the adjacent control (**Figure 3A**). In the lncRNA group, lncRNA-TOB2P1, LOC100499489, lnc-NMRAL2p, and lnc-NBPF22P were markedly upregulated in tumor tissues, while LINC01093, LOC100130899, LOC200772, and lnc-FENDRR were significantly downregulated in tumor tissues, compared to the adjacent controls (**Figure 3B**). In the miRNA group, miR-10-3p was significantly upregulated in tumor tissues, while miR-200a-3p, miR-200b-3p, and miR-139-5p were significantly downregulated in tumor tissues (**Figure 3C**). In the circRNA group, circAKR1B10 and circAKR1C3 were significantly upregulated in tumor tissues, while circHMGCS1 and circC3P1

FIGURE 5 | The lncRNA-mRNA network in HCC and GO analyses of dysregulated genes. **(A)** The co-expression network between differentially expressed lncRNA and mRNA (correlation coefficient absolute value ≥0.99). The boxes represent mRNA, the circles represent lncRNA. Red, upregulated; blue, downregulated. **(B)** GO analyses and annotation of dysregulated genes in three main categories: biological processes, cellular components and molecular functions.

FIGURE 6 | ceRNA network and GO analyses of related genes. (A) ceRNA network was constructed according to the interactions among circRNAs, mRNAs, and miRNAs. Boxes represent dysregulated mRNA, triangles represent dysregulated miRNA, circles represent dysregulated circRNAs. Red, upregulated; blue, down-regulated. (B) GO analyses of affected pathways in three main categories: cellular components, biological processes and molecular functions.

were markedly downregulated, compared to the adjacent controls (**Figure 3D**). For these 23 RNAs examined with 21 pairs of patient samples, the RT-qPCR verification was in accordance with RNA-seq results with high confidence (**Figure 3E**).

We also validated those transcripts with aberrant levels in HCC such as mRNAs previously identified by other studies (**Figure 1C**), with the same 21 HCC patient sample pairs with RT-qPCR (**Figure 4**). Among the eight transcripts (PLOD3, SF3B4, ADH4, COLEC10, HULC, SNHG7, miR-421, and miR-761), only two (ADH4 and COLEC10) transcripts were in accordance with RNA-seq results and previous reports (**Figure 4**; Wei et al., 2012; Zhang and Wu, 2018).

DISCUSSION

Co-expression Network of lncRNAs/mRNAs and GO Analysis

We then set out to investigate the identified lncRNAs associated with HCC. LncRNA functions can be predicted based on the functions of their co-expressed protein-coding genes, and alterations in the associations between these genes in clinical samples can be used to identify key lncRNAs in HCC (Cabili et al., 2011; Liao et al., 2011; Guo et al., 2013). We constructed a co-expression network of lncRNAs and co-expressed mRNAs based on the RNA-seq data to investigate their interactions (**Figure 5A**). Our analysis demonstrated that 98 lncRNAs interacted with 175 mRNAs (**Figure 5A**). The results of GO pathway analyses of the differentially RNAs showed that most co-expressed lncRNAs were closely related to several important pathways, including biological processes such as gene silencing, chromatin silencing, and response to stress, cellular components such as extracellular region, and molecular functions such as protein dimerization activity (**Figure 5B**).

Construction of CeRNA Network

One of molecular functions of circRNAs is ceRNA (Hansen et al., 2013; Memczak et al., 2013; Piwecka et al., 2017). In order to investigate the potential circRNAs acting as ceRNAs in HCC through regulating miRNAs and consequently modulating mRNAs, we constructed a ceRNA network among differentially expressed circRNAs, miRNAs, and mRNAs in HCC (**Figure 6A**). 15 circRNAs (**Table 2**), 17 miRNAs, and 89 mRNAs were found to be correlated in this ceRNA network. The results of GO pathway analyses showed that mRNAs in this network were correlated to regulations of biological processes such as protein phosphorylation, signal transduction, and cell proliferation, molecular functions such as transcription regulator activity and double-stranded DNA binding, and cellular components such as Rb-E2F complex (**Figure 6B**).

Survival Curves of Identified Genes

To explore the relationship between our identified targets and clinical observations, we then examined with survival curves in online database, UALCAN analysis[2]. Total of 9 mRNAs, 7 of them

[2]http://ualcan.path.uab.edu/analysis.html

TABLE 2 | Information of circRNAs in **Figure 5A**.

circRNA	Genomic position	Gene symbol	circRNA ID
circ018	chr2 120885263 120932580	GLI2	Novel
circ019	chr13 33109905 33111164	N4BP2L2	hsa_circ_0100273
circ023	chr8 18725183 18730243	PSD3	hsa_circ_0136098
circ030	chr16 18852886 18856973	SMG1	hsa_circ_0006434
cir031	chr10 13322975 13337606	PHYH	hsa_circ_0093118
circ033	chr5 139819703 139828890	ANKHD1	hsa_circ_0001541
circ038	chr21 16386664 16415895	NRIP1	hsa_circ_0004771
circ039	chr12 12672795 12674397	DUSP16	hsa_circ_0003855
circ040	chr7 70227855 70236630	AUTS2	hsa_circ_0080420
circ046	chr6 161006077 161012133	LPA	hsa_circ_0131238
circ061	chr19 41382431 41414558	CYP2A7	Novel
circ079	chr5 78324255 78329231	DMGDH	hsa_circ_0129713
cir082	chr9 136302868 136303486	ADAMTS13	hsa_circ_0089372
circ095	chr6 161010584 161016567	LPA	Novel
circ0102	chr6 161006077 161016567	LPA	Novel

identified in our RNA-seq and qRT-PCR verification and 2 of them previously identified and further verified by this study, were analyzed with survival curves. 6 (ZIC5, C12orf75, PZP, FAM65C, ADH4, and COLEC110) out of the 9 mRNAs were correlated to survival curves with significance in HCC (**Figure 7**). Those genes we found upregulated in HCC (ZIC5, C12orf75) were positively correlated with survival curves (patients with higher expression levels in HCC survive shorter). Genes downregulated in HCC (PZP, FAM65C, ADH4, and COLEC110) were negatively correlated with survival curves (patients with higher expression levels in HCC survive longer). For other identified genes including significantly upregulated/downregulated mRNAs, lncRNAs, miRNAs, and circular RNAs, the UALCAN database does not have HCC-relevant information about them.

The incidence of HCC has been increasing, and the consequent mortality is also rising for the past decades (Forner et al., 2018). For the early-stage patients, it is amenable to potentially curative treatments such as resection, liver transplantation, image-guided tumor ablation and systemic therapy, which can increase the 5 years survival rates to 60–70% (Llovet and Bruix, 2008; Yu, 2016). However, patients who are diagnosed at the advanced stage of HCC are treated poorly due to lack of effective treatment options for potential liver disease. Early diagnosis and effective surveillance are required for the treatment of HCC patients to reduce the disease-related mortality. Future diagnosis and treatments call for novel HCC biomarkers and potential targets.

RNA biomarkers, as measurable clinical indicators, can be used to predict and detect some diseases state and symptoms outside the body of patient with unique advances. To provide effective treatment for HCC patients and insights for future diagnosis, several potential RNA biomarkers for HCC have also been investigated (Klingenberg et al., 2017). Accumulating evidences have shown that lncRNA and miRNAs are suitable potential markers for HCC (Xie et al., 2016; Birgani et al., 2018). It has been reported that more than half of the miRNAs genes are located in cancer-associated genomic regions or in fragile sites
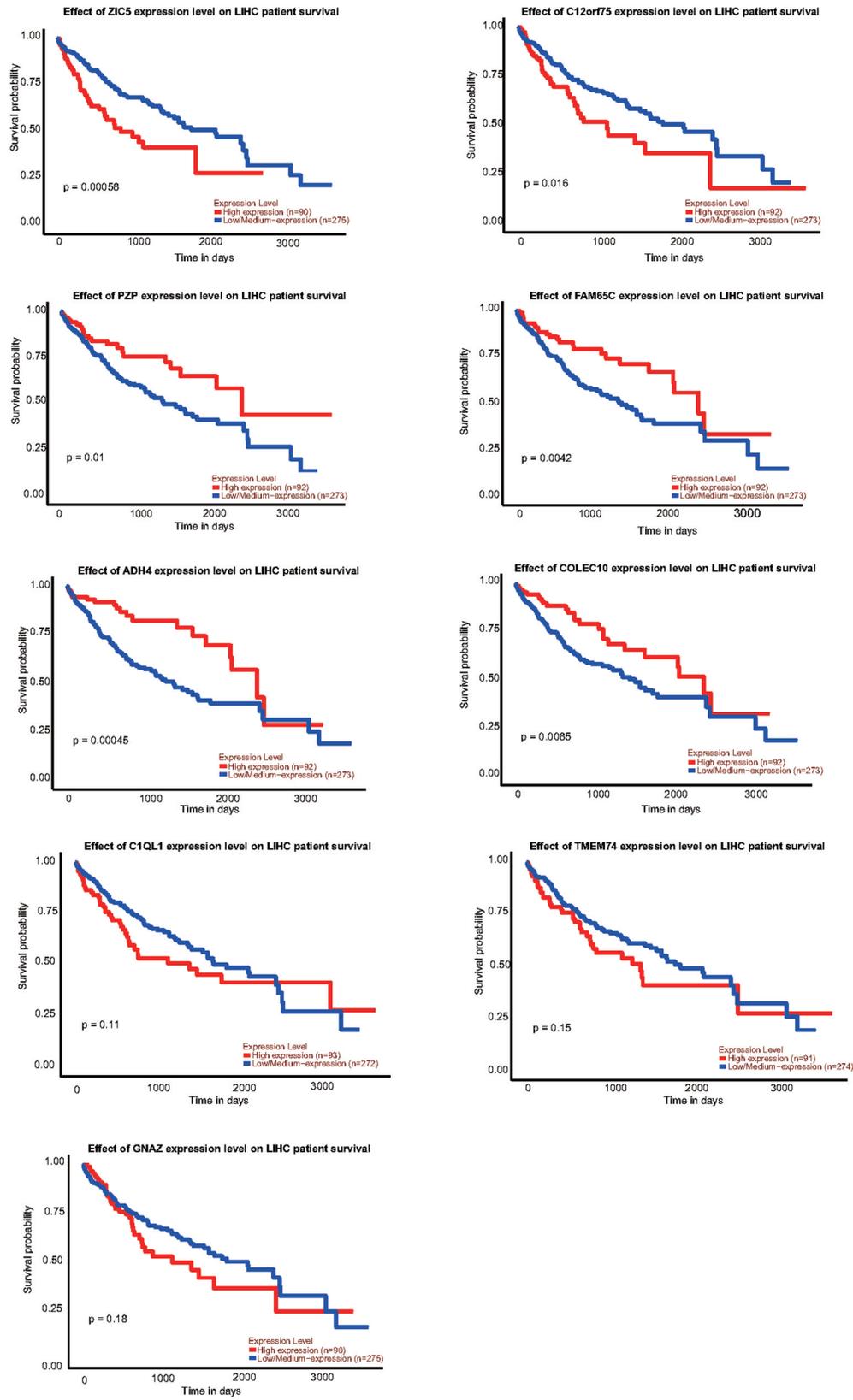
FIGURE 7 | Survival curves of identified genes analyzed by UALCAN analysis; *p*-value was calculated by log-rank test.

(Greene et al., 2017). Aberrant expression of lncRNAs, miRNAs, and circRNAs along with mRNAs may directly or indirectly lead to the progression of cancers due to their massive involvement (Ryan et al., 2010; Cheetham et al., 2013; Guarnerio et al., 2016).

RNA-seq is a powerful tool to study and detect the global transcriptome in tissues and cells (Sharma et al., 2010; Adey et al., 2013). In this study, we have identified 919 differentially expressed mRNA, 207 lncRNAs, 216 miRNAs, and 152 circRNAs in HCC through RNA-seq, these dysregulated RNAs, especially those validated with 21 patient HCC samples can be highlighted as potential biomarkers or therapeutic targets for HCC (**Figure 3**).

Although increasing pieces of evidence have demonstrated the role of aberrant expression of mRNA, miRNA, lncRNA, and circRNA in HCC, not many studies have systematically investigated the crosstalk among transcripts in this context. The co-expression network between lncRNA and mRNA (**Figure 5A**) and the ceRNA network of differential expression circRNA-miRNA-mRNA (**Figure 6A**) in our study provide insights for further investigation. Of course, both networks have their own limitation. For example, circRNAs can function not only as ceRNAs but also as transcriptional regulators (Li Z. et al., 2015; Hu and Zhou, 2018). Another interesting point is that circRNAs seem to be more often downregulated in tumor tissues as shown in this study as well as several other studies (**Figure 2D**; Zheng et al., 2016; Greene et al., 2017).

We examined the 8 RNAs reported in other studies, however only two were found consistent with our RT-qPCR results from 21 patients (**Figure 4**). The inconsistence may be due to the fact that all patients in this study were HBsAg positive with advanced stage HCC, and our patient cohort may be distinct from previous studies. We did not have an opportunity to investigate the potential exposure of the main HCC carcinogen aflatoxin of these patients, which may be a weakness of this study. Another limitation of the present study is that, just like most Chinese HCC patients, all patients in this study are already in the advanced stage upon their first diagnosis, due to limited coverage of preclinic screening. We also explored the correlation of the RNAs identified in our study with patient survival curves. 6 out of the 9 mRNAs were correlated to survival curves of HCC, indicating multiple transcripts identified in this study may play critical roles in the tumorigenesis and advance of HCC (**Figure 7**).

CONCLUSION

In conclusion, we have provided a comprehensive identification and analyses of the differentially expressed mRNAs,

miRNAs, lncRNAs, and circRNAs using RNA-seq, and some of these transcripts have been verified with clinic HCC samples. Functional network of lncRNA-mRNA and circRNA-miRNA-mRNA ceRNA network have been systematically established to further indicate potential interactions in HCC. GO pathway analyses also facilitate future studies on the specific mechanisms of HCC. We expect this work will serve as a valuable resource in future clinical diagnosis and therapy of HCC.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the Human Research Ethics Committee of University of Science and Technology of China (USTCEC201700007) with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Human Research Ethics Committee of University of Science and Technology of China (USTCEC201700007).

AUTHOR CONTRIBUTIONS

GX, GS, and LC designed and initiated this project. GS provided the major funding. GS and LC supervised the experiments. ZS and XW performed all the experiments. ZS, XW, LC, and GS analyzed the data and wrote the manuscript. All authors have discussed the results and made comments on the experiment.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Adams, J. M., and Cory, S. (2007). The Bcl-2 apoptotic switch in cancer development and therapy. *Oncogene* 26, 1324–1337. doi: 10.1038/sj.onc.1210220

Adey, A., Burton, J. N., Kitzman, J. O., Hiatt, J. B., Lewis, A. P., Martin, B. K., et al. (2013). The haplotype-resolved genome and epigenome of the

aneuploid HeLa cancer cell line. *Nature* 500, 207–211. doi: 10.1038/nature12064

Aguirre-Ghiso, J. A. (2007). Models, mechanisms and clinical evidence for cancer dormancy. *Nat. Rev. Cancer* 7, 834–846. doi: 10.1038/nrc2256

Amaravadi, R. K., and Thompson, C. B. (2007). The roles of therapy-induced autophagy and necrosis in cancer treatment. *Clin. Cancer Res.* 13, 7271–7279. doi: 10.1158/1078-0432.CCR-07-1595

Azam, F., Mehta, S., and Harris, A. L. (2010). Mechanisms of resistance to antiangiogenesis therapy. *Eur. J. Cancer* 46, 1323–1332. doi: 10.1016/j.ejca.2010.02.020

Birgani, M. T., Hajjari, M., Shahrisa, A., Khoshnevisan, A., Shoja, Z., Motahari, P., et al. (2018). Long non-coding rna snhg6 as a potential biomarker for hepatocellular carcinoma. *Pathol. Oncol. Res.* 24, 329–337. doi: 10.1007/s12253-017-0241-3

Cabili, M. N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A., et al. (2011). Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* 25, 1915–1927. doi: 10.1101/gad.17446611

Cheetham, S. W., Gruhl, F., Mattick, J. S., and Dinger, M. E. (2013). Long noncoding RNAs and the genetics of cancer. *Br. J. Cancer* 108, 2419–2425. doi: 10.1038/bjc.2013.233

Chen, L., Dzakah, E. E., and Shan, G. (2018). Targetable long non-coding RNAs in cancer treatments. *Cancer Lett.* 418, 119–124. doi: 10.1016/j.canlet.2018.01.042

Chen, L., Huang, C., Wang, X., and Shan, G. (2015). Circular RNAs in eukaryotic cells. *Curr. Genomics* 16, 312–318. doi: 10.2174/1389202916666150707161554

Chen, L., and Shan, G. (2015). Circular RNAs remain peculiarly unclear in biogenesis and function. *Sci. China Life Sci.* 58, 616–618. doi: 10.1007/s11427-015-4855-y

Cui, H., Zhang, Y., Zhang, Q., Chen, W., Zhao, H., and Liang, J. A. (2017). comprehensive genome-wide analysis of long noncoding RNA expression profile in hepatocellular carcinoma. *Cancer Med.* 6, 2932–2941. doi: 10.1002/cam4.1180

Eden, E., Navon, R., Steinfeld, I., Lipson, D., and Yakhini, Z. (2009). GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* 10:48. doi: 10.1186/1471-2105-10-48

Forner, A., Reig, M., and Bruix, J. (2018). Hepatocellular carcinoma. *Lancet* 391, 1301–1314. doi: 10.1016/S0140-6736(18)30010-2

Greene, J., Baird, A. M., Brady, L., Lim, M., Gray, S. G., McDermott, R., et al. (2017). Circular RNAs: biogenesis, function and role in human diseases. *Front. Mol. Biosci.* 4:38. doi: 10.3389/fmolb.2017.00038

Guarnerio, J., Bezzi, M., Jeong, J. C., Paffenholz, S. V., Berry, K., Naldini, M. M., et al. (2016). Oncogenic role of fusion-circRNAs derived from cancer-associated chromosomal translocations. *Cell* 165, 289–302. doi: 10.1016/j.cell.2016.07.035

Guo, X., Gao, L., Liao, Q., Xiao, H., Ma, X., Yang, X., et al. (2013). Long non-coding RNAs function annotation: a global prediction method based on bi-colored networks. *Nucleic Acids Res.* 41:e35. doi: 10.1093/nar/gks967

Hanahan, D., and Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell* 144, 646–674. doi: 10.1016/j.cell.2011.02.013

Hansen, T. B., Jensen, T. I., Clausen, B. H., Bramsen, J. B., Finsen, B., Damgaard, C. K., et al. (2013). Natural RNA circles function as efficient microRNA sponges. *Nature* 495, 384–388. doi: 10.1038/nature11993

Haque, S., and Harries, L. W. (2017). Circular RNAs (circRNAs) in health and disease. *Genes* 8:E353. doi: 10.3390/genes8120353

Holdt, L. M., Stahringer, A., Sass, K., Pichler, G., Kulak, N. A., Wilfert, W., et al. (2016). Circular non-coding RNA ANRIL modulates ribosomal RNA maturation and atherosclerosis in humans. *Nat. Commun.* 7:12429. doi: 10.1038/ncomms12429

Hu, Q., and Zhou, T. (2018). EIciRNA-mediated gene expression: tunability and bimodality. *FEBS Lett.* 592, 3460–3471. doi: 10.1002/1873-3468.13253

Iyer, M. K., Niknafs, Y. S., Malik, R., Singhal, U., Sahu, A., Hosono, Y., et al. (2015). The landscape of long non coding RNAs in the human transcriptome. *Nat. Genet.* 47, 199–208. doi: 10.1038/ng.3192

Karreth, F. A., Reschke, M., Ruocco, A., Ng, C., Chapuy, B., Léopold, V., et al. (2015). The BRAF pseudogene functions as a competitive endogenous RNA and induces lymphoma in vivo. *Cell* 161, 319–332. doi: 10.1016/j.cell.2015.02.043

Klingenberg, M., Matsuda, A., Diederichs, S., and Patel, T. (2017). Non-coding RNA in hepatocellular carcinoma: mechanisms, biomarkers and therapeutic targets. *J. Hepatol.* 67, 603–618. doi: 10.1016/j.jhep.2017.04.009

Kramer, M. F. (2011). Stem-loop RT-qPCR for miRNAs. *Curr. Protoc. Mol. Biol.* 95:15. doi: 10.1002/0471142727.mb1510s95

Kung, J. T., Colognori, D., and Lee, J. T. (2013). Long noncoding RNAs: past, present, and future. *Genetics* 193, 651–669. doi: 10.1534/genetics.112.146704

Lee, S. S., Shin, H. S., Kim, H. J., Lee, S. J., Lee, H. S., Hyun, K. H., et al. (2012). Analysis of prognostic factors and 5-year survival rate in patients with hepatocellular carcinoma: a single-center experience. *Korean J. Hepatol.* 18, 48–55. doi: 10.3350/kjhep.2012.18.1.48

Li, G., Yang, F., Xu, H., Yue, Z., Fang, X., and Liu, J. (2015). MicroRNA-708 is downregulated in hepatocellular carcinoma and suppresses tumor invasion and migration. *Biomed. Pharmacother.* 73, 154–159. doi: 10.1016/j.biopha.2015.05.010

Li, Z., Huang, C., Bao, C., Chen, L., Lin, M., Wang, X., et al. (2015). Exon-intron circular RNAs regulate transcription in the nucleus. *Nat. Struct. Mol. Biol.* 22, 256–264. doi: 10.1038/nsmb.2959

Liao, Q., Liu, C., Yuan, X., Kang, S., Miao, R., Xiao, H., et al. (2011). Large-scale prediction of long non-coding RNA functions in a coding–non-coding gene co-expression network. *Nucleic Acids Res.* 39, 3864–3878. doi: 10.1093/nar/gkq1348

Llovet, J. M., and Bruix, J. (2003). Systematic review of randomized trials for unresectable hepatocellular carcinoma: chemoembolization improves survival. *Hepatology* 37, 429–442. doi: 10.1053/jhep.2003.50047

Llovet, J. M., and Bruix, J. (2008). Novel advancements in the management of hepatocellular carcinoma in 2008. *J. Hepatol.* 48, S20–S37. doi: 10.1016/j.jhep.2008.01.022

Llovet, J. M., Burroughs, A., and Bruix, J. (2003). Hepatocellular carcinoma. *Lancet* 362, 1907–1917. doi: 10.1016/S0140-6736(03)14964-1

Marquardt, J. U., Galle, P. R., and Teufel, A. (2012). Molecular diagnosis and therapy of hepatocellular carcinoma (HCC): an emerging field for advanced technologies. *J. Hepatol.* 56, 267–275. doi: 10.1016/j.jhep.2011.07.007

McGlynn, K. A., and London, W. T. (2011). The global epidemiology of hepatocellular carcinoma: present and future. *Clin. Liver Dis.* 15, 223–243. doi: 10.1016/j.cld.2011.03.006

Memczak, S., Jens, M., Elefsinioti, A., Torti, F., Krueger, J., Rybak, A., et al. (2013). Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* 495:333. doi: 10.1038/nature11928

Palazzo, A. F., and Lee, E. S. (2015). Non-coding RNA: what is functional and what is junk? *Front Genet.* 6:2. doi: 10.3389/fgene.2015.00002

Panzitt, K., Tschernatsch, M. M., Guelly, C., Moustafa, T., Stradner, M., Strohmaier, H. M., et al. (2007). Characterization of HULC, a novel gene with striking up-regulation in hepatocellular carcinoma, as noncoding RNA. *Gastroenterology* 132, 330–342. doi: 10.1053/j.gastro.2006.08.026

Piwecka, M., Glažar, P., Hernandez-Miranda, L. R., Memczak, S., Wolf, S. A., Rybak-Wolf, A., et al. (2017). Loss of a mammalian circular RNA locus causes miRNA deregulation and affects brain function. *Science* 357:eaam8526. doi: 10.1126/science.aam8526

Prensner, J. R., Iyer, M. K., Balbin, O. A., Dhanasekaran, S. M., Cao, Q., Brenner, J. C., et al. (2011). Transcriptome sequencing across a prostate cancer cohort identifies PCAT-1, an unannotated lincRNA implicated in disease progression. *Nat. Biotech.* 29, 742–749. doi: 10.1038/nbt.1914

Rusk, N. (2015). Understanding noncoding RNAs. *Nat. Methods* 12:35. doi: 10.1038/nmeth.3235

Ryan, B. M., Robles, A. I., and Harris, C. C. (2010). Genetic variation in microRNA networks: the implications for cancer research. *Nat. Rev. Cancer.* 10, 389–402. doi: 10.1038/nrc2867

Sharma, C. M., Hoffmann, S., Darfeuille, F., Reignier, J., Findeiß, S., Sittka, A., et al. (2010). The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature* 464:250. doi: 10.1038/nature08756

Shen, Q., Eun, J. W., Lee, K., Kim, H. S., Yang, H. D., Kim, S. Y., et al. (2018). Barrier to autointegration factor 1, procollagen-lysine, 2-oxoglutarate 5-dioxygenase 3, and splicing factor 3b subunit 4 as early-stage cancer decision markers and drivers of hepatocellular carcinoma. *Hepatology* 67, 1360–1377. doi: 10.1002/hep.29606

Wang, J., Liu, X., Wu, H., Ni, P., Gu, Z., Qiao, Y., et al. (2010). CREB up-regulates long non-coding RNA, HULC expression through interaction with microRNA-372 in liver cancer. *Nucleic Acids Res.* 38, 5366–5383. doi: 10.1093/nar/gkq285

Wei, R. R., Zhang, M. Y., Rao, H. L., Pu, H. Y., Zhang, H. Z., and Wang, H. Y. (2012). Identification of ADH4 as a novel and potential prognostic marker in hepatocellular carcinoma. *Med. Oncol.* 29, 2737–2743. doi: 10.1007/s12032-011-0126-3

Xie, Q. Y., Almudevar, A., Whitney-Miller, C. L., Barry, C. T., and McCall, M. N. (2016). A microRNA biomarker of hepatocellular carcinoma recurrence

following liver transplantation accounting for within-patient heterogeneity. *BMC Med. Genomics* 9:18. doi: 10.1186/s12920-016-0179-4

Yu, S. J. (2016). A concise review of updated guidelines regarding the management of hepatocellular carcinoma around the world: 2010–2016. *Clin. Mol. Hepatol.* 22, 7–17. doi: 10.3350/cmh.2016.22.1.7

Zhang, B., and Wu, H. (2018). Decreased expression of COLEC10 predicts poor overall survival in patients with hepatocellular carcinoma. *Cancer Manag. Res.* 10, 2369–2375. doi: 10.2147/CMAR.S161210

Zhang, Y., Gong, W., Dai, S., Huang, G., Shen, X., Gao, M., et al. (2012). Downregulation of human farnesoid X receptor by miR-421 promotes proliferation and migration of hepatocellular carcinoma cells. *Mol. Cancer Res.* 10, 516–522. doi: 10.1158/1541-7786.MCR-11-0473

Zheng, Q., Bao, C., Guo, W., Li, S., Chen, J., Chen, B., et al. (2016). Circular RNA profiling reveals an abundant circHIPK3 that regulates cell growth by sponging multiple miRNAs. *Nat. Commun.* 7:11215. doi: 10.1038/ncomms11215

Zhou, X., Zhang, L., Zheng, B., Yan, Y., Zhang, Y., Xie, H., et al. (2016). Micro RNA-761 is upregulated in hepatocellular carcinoma and regulates tumorigenesis by targeting Mitofusin-2. *Cancer Sci.* 107, 424–432. doi: 10.1111/cas.12904