Identifying Circular RNAs in HepG2 Expressing Genotype IV Swine Hepatitis E Virus ORF3 Via Whole Genome Sequencing

Hanwei Jiao[1,2,3] ⬤, Yu Zhao[1,4], Zhixiong Zhou[1], Wenjie Li[1], Bowen Li[1], Guojing Gu[1], Yichen Luo[1,2,3], Xuehong Shuai[1,2,3], Cailiang Fan[5], Li Wu[1,3], Jixuan Chen[1,3], Qingzhou Huang[1,3], Fengyang Wang[6], and Juan Liu[1,2,3]

Abstract

Swine hepatitis E (SHE) is a new type of zoonotic infectious disease caused by swine hepatitis E virus (SHEV). Open reading frame 3 (ORF3) is a key regulatory and virulent protein of SHEV. Circular RNAs (circRNAs) are a special kind of non-coding RNA molecule, which has a closed ring structure. In this study, to identify the circRNA profile in host cells affected by SHEV ORF3, adenovirus ADV4-ORF3 mediated the overexpression of ORF3 in HepG2 cells, whole genome sequencing was used to investigate the differentially expressed circRNAs, GO and KEGG were performed to enrichment analyze of differentially expressed circRNA-hosting gene, and Targetscan and miRanda softwares were used to analyze the interaction between circRNA and miRNA. The results showed adenovirus successfully mediated the overexpression of ORF3 in HepG2 cells, 1,105 up-regulation circRNAs and 1,556 down-regulation circRNAs were identified in ADV4-ORF3 infection group compared with the control. GO function enrichment analysis of differentially expressed circRNAs-hosting genes classified three main categories (cellular component, biological process and molecular function). KEGG pathway enrichment analysis scatter plot showed the pathway term of top20. The circRNAs with top10 number of BS sites for qRT-PCR validation were selected to confirmed, the results indicated that the up-regulated hsa_circ_0001423 and hsa_circ_0006404, and down-regulated of hsa_circ_0004833 and hsa_circ_0007444 were consistent with the sequencing data. Our findings first preliminarily found that ORF3 protein may affect triglyceride activation (GO:0006642) and riboflavin metabolism (ko00740) in HepG2 cells, which provides a scientific basis for further elucidating the effect of ORF3 on host lipid metabolism and the mechanism of SHEV infection.

Keywords

[1] College of Veterinary Medicine, Southwest University, Chongqing, China
[2] Immunology Research Center, Medical Research Institute, Southwest University, Chongqing, China
[3] Chongqing Veterinary Scientific Engineering Research Center, Southwest University, Chongqing, China
[4] Institute of Animal Husbandry and Veterinary Medicine of Guizhou Academy of Agricultural Science, Guiyang, China
[5] Rongchang Animal Epidemic Prevention and Control Center, Chongqing, Rongchang, China
[6] Hainan Key Lab of Tropical Animal Reproduction and Breeding and Epidemic Disease Research, College of Animal Science and Technology, Hainan University, Haikou, China

Corresponding Authors:
Hanwei Jiao, Southwest University college Road 160# Chongqing, 402460 China.
Email: jiaohanwei@swu.edu.cn

Fengyang Wang, Hainan University Haikou, Hainan China.
Email: fywang68@163.com

Juan Liu, Southwest University Chongqing, China.
Email: liujrc@163.com

Introduction

Hepatitis E (HE) is an acute infectious disease caused by hepatitis E virus (HEV), since the first outbreak of HE in India in 1955, it has been prevalent in India, Nepal, Sudan, Kyrgyzstan, and China[1–3]. HEV is a single strand positive strand RNA virus, which is a spherical particle without envelope. HEV was divided into 8 genotypes, 1 to 8. The most popular strains in China were type 1 (formerly known as Myanmar strain), followed by type 4. At present, cell culture of HEV is not available, which can infect monkeys in laboratory[4–6]. The resistance of HEV to external environment is not strong. The clinical manifestations of HE are recessive infection, acute hepatitis, nonchronic hepatitis, and so on, it is highly prevalent in young people and adults aged 15–39 years old, and there is no HE vaccine[7–9].

Swine Hepatitis E (SHE) is a newly discovered zoonotic infectious disease caused by swine hepatitis E virus (SHEV). The clinical features of SHEV infection in pigs are almost no symptoms except jaundice, but it has fatal infection to humans, SHE seriously threatens the development of animal husbandry and the safety of human public health[10–12]. SHEV open reading frame 3 (ORF3) is an important virulent protein, which may affect the viral replication and release, lipid metabolism and the occurrence of hepatitis[13,14].

Circular RNAs (circRNAs) are a class of noncoding RNAs, which is also the latest research hotspot in RNA field[15]. CircRNA is a closed ring structure, which is not affected by RNA exonuclease, and its expression is more stable and not easy to degrade[16]. In terms of function, recent studies have shown that circRNA molecules are rich in microRNA (miRNA) binding sites, which play the role of miRNA sponge in cells, thus relieving the inhibition of miRNA on its target genes and increasing the expression level of target genes[17]. This mechanism is called competitive endogenous RNA (ceRNA) mechanism. CircRNAs plays an important regulatory role in diseases through the interaction of miRNAs associated with diseases[18].

In an attempt to identify the differentially expressed circRNAs and predict the pathways we are interested in using bioinformatics analysis in HepG2 cells expressing genotype IV Swine Hepatitis E Virus ORF3, which may help to elucidate the circRNAs play an important role in the SHEV ORF3-host cell interactions and explain the pathogenesis of SHEV.

Materials and Methods

Cell Culture

HepG2 cells were obtained from Shanghai cell bank of Chinese Academy of Science, cultured in Dulbecco's modification of Eagle's medium Dulbecco (DMEM) (Life technology, USA) containing 10% fatal bovine serum (FBS) (Gibco, USA) and placed in a 37°C, 5% CO_2 incubator for routine culture.

Packaging Preparation for the Over Expressed Adenovirus of ADV4-ORF3 and ADV4-NC

Genotype IV swine hepatitis E virus (SHEV) open reading frame 3 (ORF3) gene was synthesized by Shanghai genepharma Co., Ltd. The full length of SHEV-ORF3 gene was 345 bp, and was cloned into vector of ADV4 to construct the overexpressed adenovirus plasmid of ADV4-SHEV-ORF3, the restriction sites of restriction enzyme were EcoRI and BamHI. The ADV4-SHEV-ORF3 plasmid and skeleton plasmids of pGP-Ad-Pac vector were prepared, extracted with high purity and endotoxin free respectively, and were co-transfected into 293A cells to prepare high titer over expressed adenovirus of ADV4-ORF3 and ADV4-NC (Genepharma, Shanghai, China).

Preparation of Anti-ORF3 Polyclonal Antibody and Western Blot

Genotype IV SHEV-ORF3 gene was synthesized by Sangon Biotech (Shanghai) Co., Ltd. The SHEV-ORF3 gene was cloned into PET-28a vector, and the Recombinant Prokaryotic expression vector PET-28a-SHEV-ORF3 was constructed, the correct reading frame was confirmed by sequencing. The detailed preparation of anti-ORF3 polyclonal antibody was previously described[19]. The whole blood was collected from heart and serum was separated. The antibody titer was detected by indirect ELISA.

HepG2 cells were inoculated into 12 well cell culture plates with a density of 1×10^6 per well for 12 h, dilution of adenovirus ADV4-ORF3 and ADV4-NC with 1 mL complete medium containing 0.8 μL polybrene, the above fresh medium was added to the cells, and the plate was gently shaked by cross. 24 h after adenovirus infection, 1.2 ml trizol (Life technology, USA) was added into each well, put it at room temperature for 2 min, and gently blew the cells with a pipette to clear solution, HepG2 cells were harvested after ADV4-NC and ADV4-ORF3 infection for 24 h, and Western blotting was performed as described before[20]. 40μg total protein was used for loading, β-actin was used for an internal control. The primary antibody were a rabbit anti-ORF3 (1:400 dilution) and monoclonal mouse anti-β-actin (1:2000 dilution; Santa Cruz Biotechnology, USA). Secondary antibodies were horseradish peroxides (HRP)-labeled rabbit goat anti-rabbit IgG (1:5000 dilution; Abcam, USA) and mouse-IgGk BP-HRP (1:5000 dilution; Santa Cruz Biotechnology, USA).

High Quality RNA Isolation and RNA Library Construction and Whole Genome Sequencing

High quality total RNAs of Ad_GFP ($n = 3$) and Ad_ORF3 ($n = 3$) isolation and RNA library construction were previously described[19]. At last, the average insert size for the final cDNA library was 300 bp (± 50 bp). We performed the

paired-end sequencing on an Illumina Hiseq 4000 (LC Bio, China) following the vendor's recommended protocol.

circRNAs Bioinformation Analysis

Firstly, Cutadapt[21] was used to remove the reads that contained adaptor contamination, low quality bases and undetermined bases. Then sequence quality was verified using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). We used Bowtie[22] and Hisat[23] to map reads to the genome of species. Remaining reads (unmapped reads) were still mapped to genome using tophat-fusion[24]. CIRCExplorer[25,26] and CIRI[27] was used to denovo assemble the mapped reads to circular RNAs at first; Then, back splicing reads were identified in unmapped reads by tophat-fusion. All samples were generated unique circular RNAs. The differentially expressed circRNAs were selected with log2 (fold change) > 1 or log2 (fold change) < -1 and with statistical significance (P value < 0.05) by R package-edgeR[28].

Sequencing Experimental Design and circRNA Expression Profile Identifying

High quality total RNAs were extracted from Ad_GFP ($n = 3$) and Ad_ORF3 ($n = 3$), a chain-specific library was constructed to remove ribosomal RNA (rRNA depletion) from the total RNA. Subsequently, RNA-seq libraries were constructed and sequenced. High-throughput sequencing data were used to conduct a quality assessment and aligned to the Homo sapiens reference genome (ftp://ftp.ensembl.org/pub/release-96/fasta/homo_sapiens/dna/). According to the structural characteristics of circRNAs and the characteristics of splicing sequence, we used CIRCExploter2 and CIRI software to predict circRNAs, and integrate the results of the two software according to the starting and terminating position of circRNA. The circRNAs criteria includes mismatch ≤ 2, back-spliced junctions reads ≥ 1 and the distance between two splice sites was less than 100 kb. Cause the special looping mechanism of circRNAs. Starting with hsa-circ was the known circRNAs contained in the circbase database. Circbase is a database which collects and integrates the published circRNA data, including the circRNAs information of human, mouse and other species.

The types of circRNAs can be divided into following four categories according to their sources: circRNA of all exons, EicircRNA of intron and exon combinations, ciRNA composed of introns, circRNA produced by cyclization of virus RNA genome, tRNA, rRNA, snRNA, and so on. Bioinformatics can be used to analyze the ciRNA of exon and intron. More than 80% of circRNAs contain exons. Different types of circRNAs have different biological functions. circRNAs or EIcircRNAs located in the nucleus are mainly involved in transcriptional regulation. At the same time, if the absolute value of log2 fold change ≥ 1 and P value ≤ 0.05, the circRNA was marked as yes, if not, it was marked as no.

In this study, we selected the circRNAs with |log2 fold-change| ≥ 1.5 as the research target.

Gene Ontology (GO) and KEGG Enrichment Analysis of Differentially Expressed circRNAs-Hosting Genes

In our sequencing results, we annotate and enrich the host genes of circRNAs. At present, there is no direct evidence that there is a direct link between circRNA and the functional annotation of its hosting gene, the function of circRNAs are reflected through their hosting genes. Therefore, we used GO and KEGG to perform functional enrichment analysis of differentially expressed circRNAs-hosting genes. To better reflect the cluster expression mode, we used Z-value to display the differential hosting genes expression of srpbm for biological repeat.

GO enrichment analysis histogram of differential expression circRNAs-hosting genes reflects the number distribution of differential genes on GO terms enriched by biological process, cellular component and molecular function. GO enrichment analysis scatter plot of differential expression circRNAs-hosting genes was displayed by ggplot2, ggplot2 was used to display the results of GO enrichment analysis in the form of scatter plot, in which the abscissa rich factor represented the number of differential circRNAs located in the GO/the total number of circRNAs located in the GO (rich factor = S gene number/b gene number), the larger the rich factor was, the higher the degree of GO enrichment, the ordinate was GO_term.

KEGG enrichment analysis scatter plot was displayed by ggplot2, ggplot2 was used to display the results of KEGG enrichment analysis in the form of scatter plot, in which the abscissa rich factor represented the number of differential circRNAs located in the KEGG/the total number of circRNAs located in the KEGG (rich factor = S gene number/b gene number). The larger the rich factor was, the higher the degree of KEGG enrichment was.

Analysis of Interaction Between circRNA and miRNA

There are five possible regulatory mechanisms of circRNAs: participating in transcription regulation in the nucleus, competing with mRNA precursor in transcription, target binding site of competing miRNA in the cytoplasm, containing ribosome entry site, it can translate and express effective polypeptide or egg white, and interaction between circRNA and protein. In this study, we focused on the third regulatory mechanism, that was, circRNAs as an endogenous competitive RNA (ceRNA) affects the post transcriptional regulatory function of miRNA. Two softwares, Targetscan (http://www.targetscan.org/mamm_31/) and miRanda (http://miranda.org.uk/), were used to analyze the interaction between circRNA and miRNA. Targetscan predicts miRNA target based on seed region. miRanda is mainly based on the binding free energy of circRNA and miRNA, the smaller the free energy, the stronger the binding ability.

qRT-PCR Validation

To validate the sequencing data, and research has shown that we should focus on the number of back splitting junction

Table 1. The primers of circRNA with top10 Back Splitting (BS) Sites for qRT-PCR Validation.

CricRNA_gene_name	Primer Sequence (5′-3′)
hsa_circ_0001423	F: ACTTGAAAGTGCCTGCCAAA
	R: CTGGTTGCGTCTTTCCTTCT
hsa_circ_0007444	F: TTCTGGGGATGTTTCAAATGT
	R: ACACACTGGCAGCAGAACAG
hsa_circ_0039927	F: CGCTACTGCAAGCCAAGAG
	R: TGGTAAGCAAAGTGGTGTGG
hsa_circ_0018992	F: ACATGTGGGATGAAGAAGGC
	R: CCATACTTTGGCAACTTGGAA
hsa_circ_0093996	F: AAGATTCTGAACTGCCCACCT
	R: TGGGTTCAATTCCAATTTTTG
hsa_circ_0006404	F: GTGCTAAGCAGGCCTCATCT
	R: TCTTGCCAGTTCCCTCATTC
hsa_circ_0004833	F: TGTTTTTGTGCTTGGCTGAC
	R: CCCATCGGAGGACTTTATCA
circRNA5648	F: GACTTTGGACTGCTACGATCC
	R: CGGTTCTCATGCACACTGAC
circRNA5642	F: GATGTTTATGGGGTCAGGCG
	R: TGAGTCTGAACCCGAAGCTT
circRNA5459	F: AGGAGAGCTGATAGTATTCATTCCA
	R: GCAATGCATGACTATAGTTAAAAGC

(BS) for the later data screening and mining, average the number of BS sites of circRNAs identified in each sample, then arrange the average number of BS sites in descending order, and select the differential expression of circRNAs with top10 BS sites for qRT-PCR validation. The primers were designed by NCBI primer online software (Table 1), GAPDH was used as an internal control. Total RNAs extracted from Ad_GFP and Ad_ORF3 were used for reverse transcription by PrimeScriptTM RT reagent kit (TaKaRa, Japan), qRT-PCR was performed by TB Green® Premix Ex Taq™ II (Tli RNaseH Plus) (TaKaRa, Japan). The relative expression level of each circRNA was calculated using the $2^{-\Delta\Delta Ct}$ method.

Results

Adenovirus ADV4-ORF3 Successfully Mediated the Overexpression of ORF3 in HepG2 Cells

HepG2 cells were infected adenovirus ADV4-ORF3 and ADV4-NC (MOI=5:1) for 24 h, respectively. The results of fluorescence microscopy showed that the cells basically expressed green fluorescence compared with the HepG2 blank cells (Fig. 1A–C). The indirect ELISA results showed that the anti-ORF3 titer reached 1:12800 (data not shown).
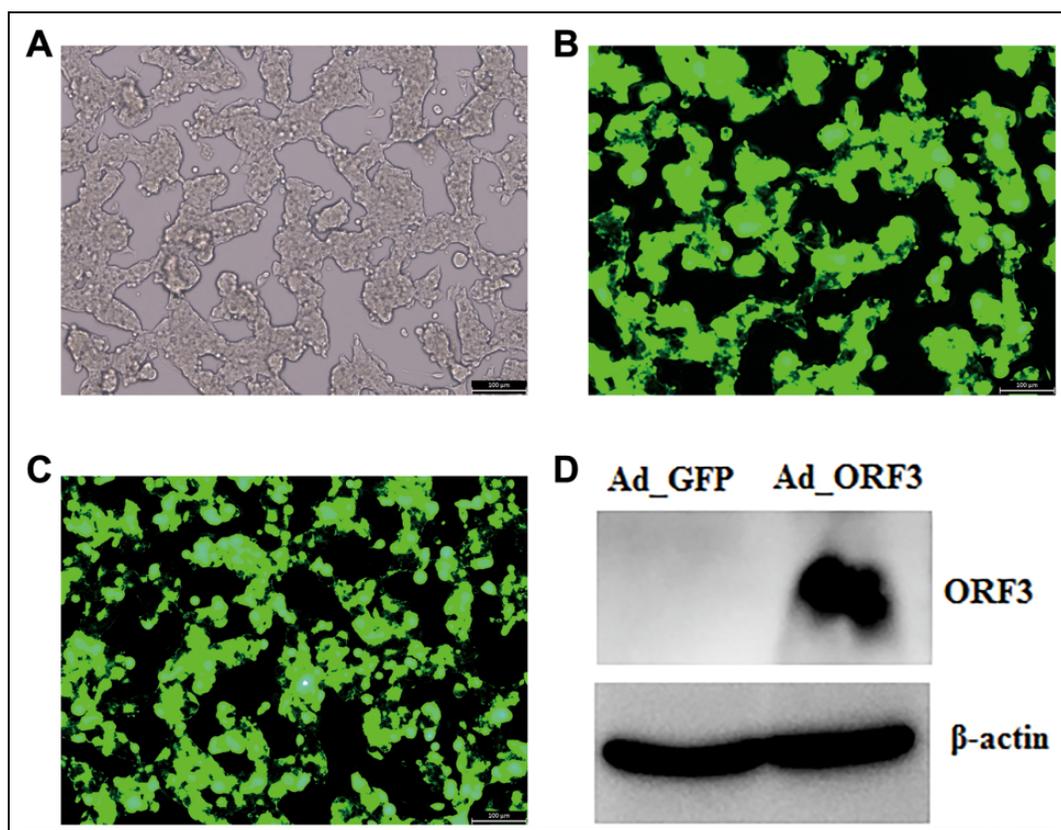
Figure 1. Overexpression of ORF3 mediated by adenovirus ADV4-ORF3 (Ad_ORF3) in HepG2 cells. (A) HepG2 blank cells. (B) HepG2 cells infected with ADV4-NC (Ad_GFP) for 24 h as control group. (C) HepG2 cells infected with ADV4-ORF3 for 24 h as experimental group. Scale bar: 100 μm. (D) Western blot analysis of HepG2 infected with ADV4-NC and ADV4-ORF3 for 24 h by rabbit polyclonal anti-ORF3 and monoclonal mouse anti-β-actin.
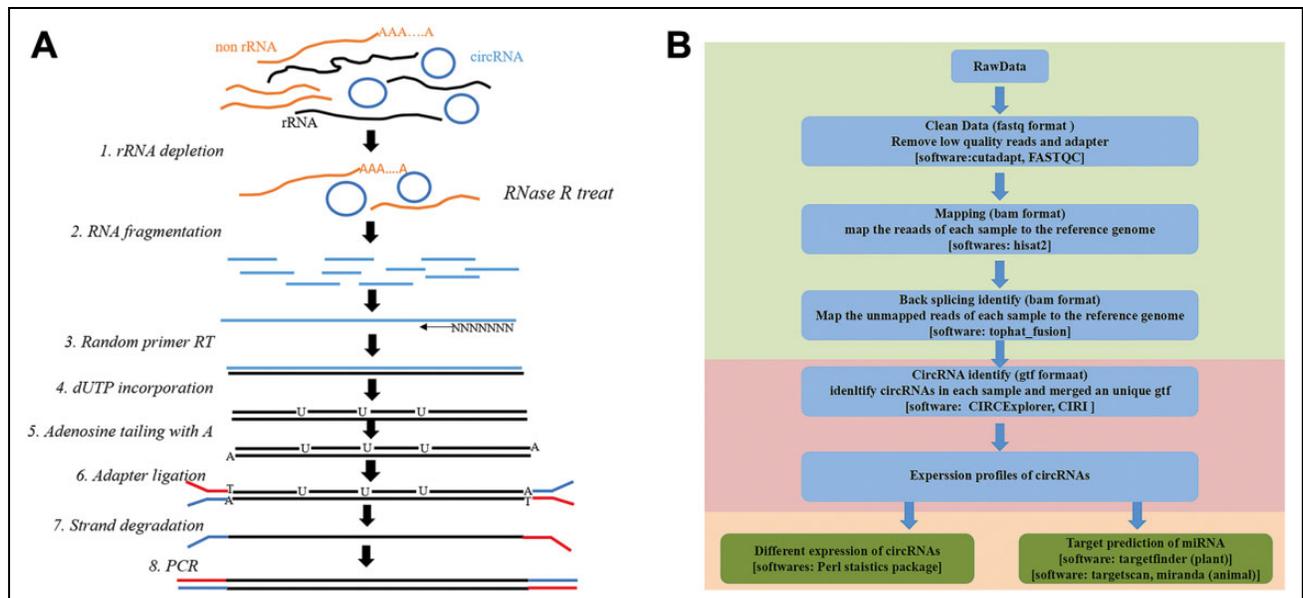
Figure 2. The overview of circRNA high-throughput sequencing experiment workflow. (A) The workfow of construction library of circRNA sequencing. High quality total RNAs were extracted from Ad_GFP ($n = 3$) and Ad_ORF3 ($n = 3$), the ribosomal RNA (rRNA depletion) was removed from the total RNA, and the circRNAs were sequenced by Illumina HiSeq 4000, and the length of the sequence was 2×150 bp. (B) circRNA bioinformatics analysis workflow.

Table 2. The Mapped Region in Six Samples.

Sample	Ad_GFP_1	Ad_GFP_2	Ad_GFP_3	Ad_ORF3_1	Ad_ORF3_2	Ad_ORF3_3
Valid reads	76422050	72863272	70886348	80331434	79550392	74340958
Mapped reads	73759465(96.52%)	70532973(96.80%)	68469619(96.59%)	77506773(96.48%)	76865926(96.63%)	71334229(95.96%)
Unique Mapped reads	56104320(73.41%)	55768628(76.54%)	52692847(74.33%)	58300247(72.57%)	58773001(73.88%)	52126112(70.12%)
Multi Mapped reads	17655145(23.10%)	14764345(20.26%)	15776772(22.26%)	19206526(23.91%)	18092925(22.74%)	19208117(25.84%)
PE Mapped reads	69425224(90.84%)	67188456(92.21%)	63807292(90.01%)	71865924(89.46%)	70599754(88.75%)	67093056(90.25%)
Reads map to sense strand	33787842(44.21%)	32724994(44.91%)	31288110(44.14%)	34819881(43.35%)	34947808(43.93%)	31996502(43.04%)
Reads map to antisense strand	33802519(44.23%)	32752695(44.95%)	31302334(44.16%)	34848116(43.38%)	34970273(43.96%)	32006611(43.05%)
Non-splice reads	42656503(55.82%)	44818091(61.51%)	41797270(58.96%)	45268941(56.35%)	47561774(59.79%)	41361120(55.64%)
Splice reads	24933858(32.63%)	20659598(28.35%)	20793174(29.33%)	24399056(30.37%)	22356307(28.10%)	22641993(30.46%)
Unmapped reads	2662585(3.48%)	2330299(3.20%)	2416729(3.41%)	2824661(3.52%)	2684466(3.37%)	3006729(4.04%)

Western blotting results indicated that the specific band at the molecular weight about 12kD for cell extracts from Ad_ORF3 was found, but no ORF3 expressing in Ad_GFP was found (Fig. 1D), which demonstrated that Ad-ORF3 successfully mediated the overexpression of ORF3 in HepG2 cells.

Overview of High-Throughput Sequencing Experimental Design

The overview of the circRNA high-throughput sequencing experiment workflows are presented in Fig. 2. The sequencing data of all experimental samples in the fastq format have been submitted to the Sequence Read Archive of NCBI under the accession number GSE147129.

circRNA Expression Analysis in Different Samples

The mapped region in six samples as shown in Table 2. Bioinformatics can be used to analyze the ciRNAs of exon

and intron. More than 80% of circRNAs contain exons. In this study, the GC content of reads and distribution of reads length of each sample as shown in Fig. 3A, B. Determining the type of circRNA is helpful for later functional research, and further screening the type of circRNA on the basis of statistical screening. In this study, circRNA, ciRNA, and intergenic three types of circRNAs were identified.

Differential Expression Analysis of circRNAs

Statistics of up and down-regulation frequency of differentially expressed circRNAs in different groups showed that 1,105 circRNAs were up-regulated, and 1,556 circRNAs were down-regulated in Ad_GFP vs Ad_ORF3 (Fig. 4A and Table 3). In the differential expression circRNAs volcano, Red represented the up-regulated circRNAs, blue represented the down regulated circRNAs, and gray represented the nonsignificant circRNAs (Fig. 4B). Cluster analysis of differential expression circRNAs, the abscissa was the
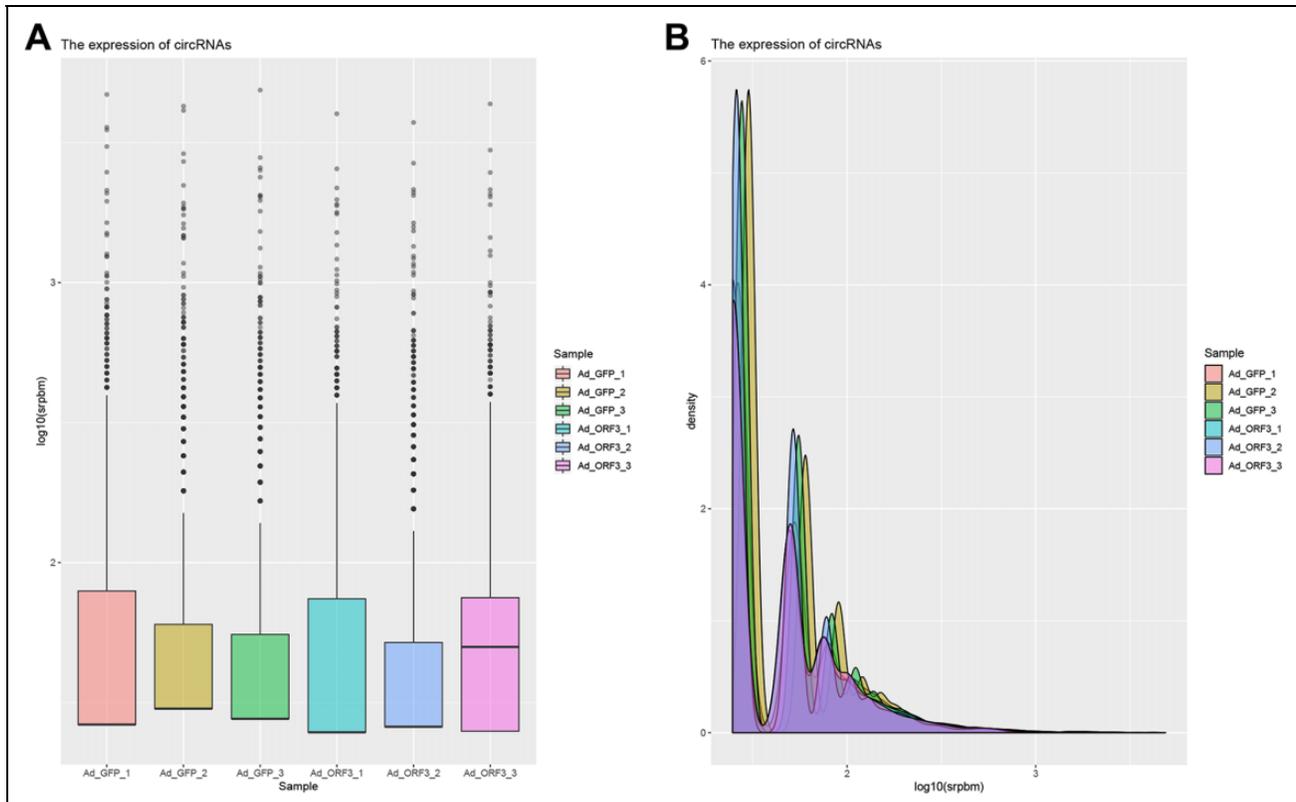
Figure 3. Analysis the expression of circRNAs in different samples. (A) The boxplot of the circRNA expression in Ad_GFP1, Ad_GFP2, Ad_GFP3, Ad_ORF3_1, Ad_ORF3_2, and Ad_ORF3_3. (B) The density of the circRNA expression in Ad_GFP1, Ad_GFP2, Ad_GFP3, Ad_ORF3_1, Ad_ORF3_2, and Ad_ORF3_3.

different sample, the ordinate was the screened differential expression circRNAs (top100), different colors represented different circRNAs expression levels, the color from blue through white to red represented the expression quantity from low to high, red represented the high expression circRNAs, and blue represented the low expression circRNAs (Fig. 4C). These data show that 1,105 circRNAs were up-regulated, and 1,556 circRNAs were down-regulated in HepG2 cells expressing genotype IV Swine Hepatitis E Virus ORF3.

GO and KEGG Enrichment Analysis of Differentially Expressed circRNAs-Hosting Genes

The histogram of GO enrichment analysis results indicated the number distribution of the 1,105 up-regulated and 1,556 down-regulated circRNAs-hosting genes on GO term enriched by biological_process, cellular_component and molecular_function (Fig. 5A). The scatter diagram of GO enrichment analysis was based on the GO term of top20 according to the significance of enrichment (*P* value), they were ventricular septum morphogenesis (GO:0060412), type II transforming growth factor beta receptor binding (GO:0005114), triglyceride mobilization (GO:0006642), transcription corepressor activity (GO:0003714), RNA polymerase II CTD heptapeptide repeat phosphatase activity

(GO:0008420), RNA polymerase core enzyme binding (GO:0043175), protein N-terminus binding (GO:0047485), negative regulation of protein export from nucleus (GO:0046826), negative regulation of neuroblast proliferation (GO:0007406), morphogenesis of a branching structure (GO:0001763), miRNA catabolic process (GO:0010587), metal ion binding (GO:0046872), ionotropic glutamate receptor complex (GO:0008328), intrinsic apoptotic signaling pathway in response to DNA damage (GO:0042771), intracellular protein transport (GO:0006886), histone methyltransferase activity (H3-K27) specific (GO:0046976), endoplasmic reticulum exit site (GO:0070971), COPII-coated vesicle budding (GO:0090114), cell-cell adhesion (GO:0098609), and animal organ regeneration (GO:0031100) (Fig. 5B).

The vertical coordinate was pathway term. According to the significance of enrichment (*P* value), KEGG enrichment analysis scatter plot was based on the pathway term of top20, they were ubiquitin mediated proteolysis (ko04120), transcriptional misregulation in cancer (ko05202), RNA degradation (ko03018), riboflavin metabolism (ko00740), progesterone-mediated oocyte maturation (ko04914), platinum drug resistance (ko01524), phototransduction (ko04744), pancreatic cancer (ko05212), mTOR signaling pathway (ko04150), hippo signaling pathway (ko04390), hepatocellular carcinoma (ko05225), hepatitis B (ko05161), gastric cancer (ko05226), endocrine and other factor-
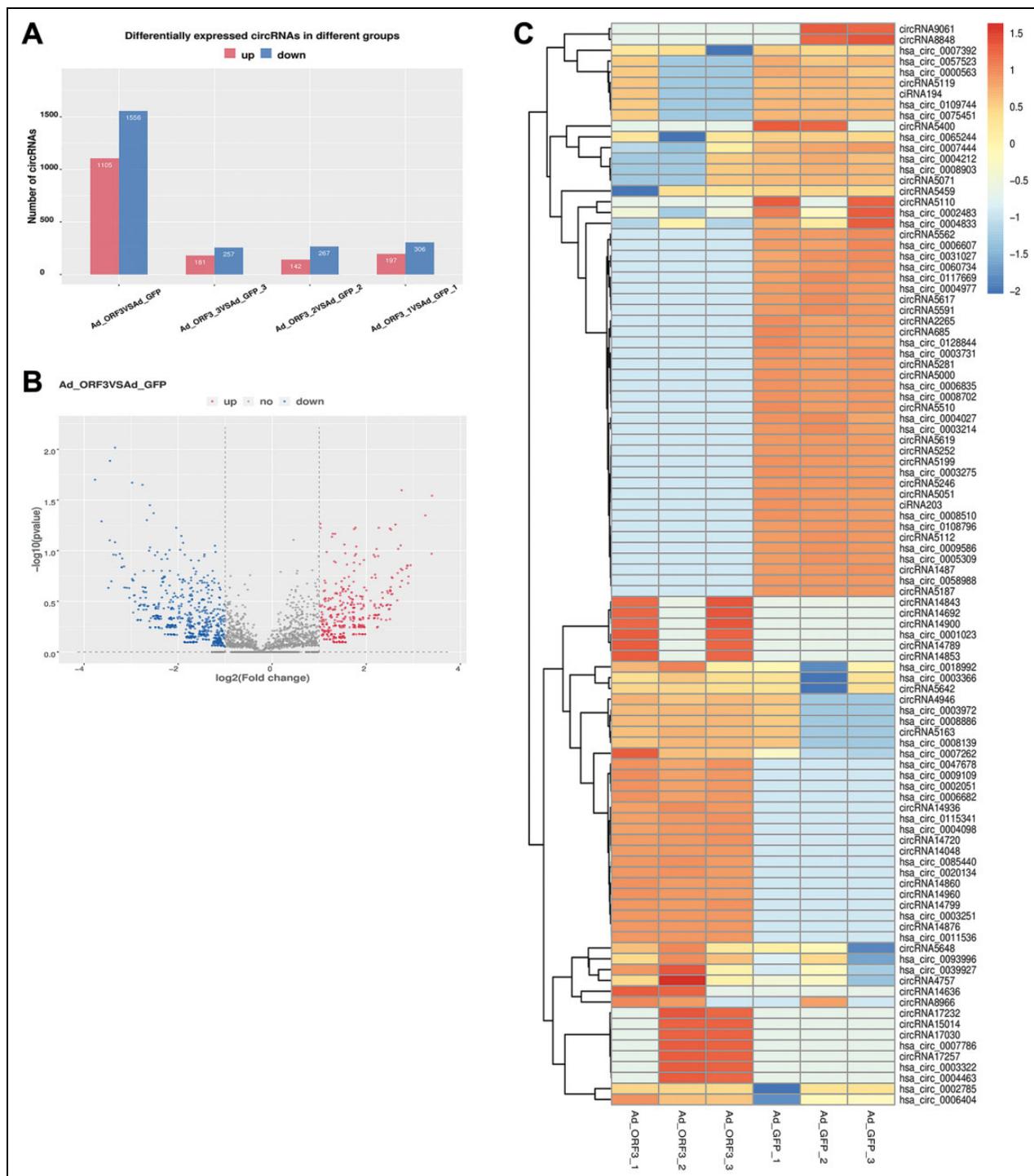
Figure 4. Identification of the differentially expressed circRNAs. (A) Statistics of up and down-regulation frequency of differentially expressed circRNA in different groups. The histogram showed that the red column represented the up-regulated circRNAs frequency, and the blue column represented the down-regulated circRNAs frequency, which directly showed the number of differentially expressed circRNAs and the up-regulated situation in different comparison groups. (B) Differential expression circRNA volcano, X-axis represented log2 (fold change), Y-axis represented -log10 (P value). (C) Cluster analysis of differentially expressed circRNA, X-axis represented the different samples, Y-axis represented the differentially expressed circRNAs.

regulated calcium reabsorption (ko04961), colorectal cancer (ko05210), cAMP signaling pathway (ko04024), amoebiasis (ko05146), aldosterone-regulated sodium reabsorption

(ko04960), AGE-RAGE signaling pathway in diabetic complications and adherens junction (ko04933) (Fig. 5C). The specific biological functions of the circRNAs selected to be

Table 3. The Summary of High-Throughput Sequencing Data used in this Study.

Sample	Raw Data		Valid Data		Valid Ratio(reads)	Q20%	Q30%	GC content%
	Read	Base	Read	Base				
Ad_GFP_1	91669794	13.75G	76422050	11.46G	83.37	99.99	98.39	51
Ad_GFP_2	86905270	13.04G	72863272	10.93G	83.84	99.99	98.36	50
Ad_GFP_3	85198608	12.78G	70886348	10.63G	83.20	99.99	98.25	51
Ad_ORF3_1	97563154	14.63G	80331434	12.05G	82.34	99.99	98.52	51.50
Ad_ORF3_2	96477864	14.47G	79550392	11.93G	82.45	99.99	98.41	50.50
Ad_ORF3_3	89299078	13.39G	74340958	11.15G	83.25	99.99	97.95	52

verified through a series of functional verification in the later stage in-depth study, we will focus on the functional enrichment analysis results of circRNAs-hosting gene.

Predicting circRNAs Involved in Triglyceride Mobilization and Riboflavin Metabolism Affected by ORF3

GO functional enrichment analysis predicted that 6 circRNAs were involved in triglyceride mobilization in HepG2 cells affected by ORF3, which is closely related to SHEV ORF3 function, they were hsa_circ_0002229, circRNA14701, hsa_circ_0093888, circRNA3581, circRNA3580, and circRNA3582 (Fig. 6A). KEGG pathway analysis predicted that 4 circRNAs were involved in riboflavin metabolism in HepG2 cells affected by ORF3, which is closely related to SHEV ORF3 function, they were ciRNA410, hsa_circ_0130711, circRNA13511, and hsa_circ_0130715 (Fig. 6B).

Analysis of Interaction Between circRNA and miRNA

Based on the literature reported, there were five possible regulatory mechanisms of circRNAs. circRNA as an ceRNA affected the post transcriptional regulation function of miRNA, which provided the basis for subsequent experimental verification. Two softwares, Targetscan and miRanda, were used to analyze the circRNA-miRNA network, the results showed there were a total of 31,778 circRNA-miRNA pairs of network (Supplementary file).

Validation of Sequencing Results by qRT-PCR

According to the number of BS sites of circRNAs, we selected circRNAs with top10 average the number of BS sites for qRT-PCR validation (Table 4). The results showed that hsa_circ_0001423, hsa_circ_0004833, hsa_circ_0006404 and hsa_circ_0007444 were consistent with the sequencing results, and comfirmed that hsa_circ_0001423 and hsa_circ_0006404 were up-regulated, and hsa_circ_0004833 and hsa_circ_0007444 were down-regulated in Ad_ORF3 compared with Ad_GFP (Fig. 7).

Discussion

High throughput sequencing, also known as "next generation" sequencing technology, is characterized by being able to sequence hundreds of thousands to millions of DNA molecules in parallel at one time and generally having short reading length, which provides a powerful method to identify diferentially expressed circRNAs[29–31]. The distribution of ORF3 epitopes of genotype 4 hepatitis E virus affects the proline rich C-terminal domain, which has an effect on the immune activity of downstream V105 dlp108[14]. In the northeast of China, some strains of hepatitis E virus that were prevalent in human and pig populations were all genotype 4 swine hepatitis E virus[32–34]. SHEV ORF3 is an important virulent protein, which affects the replication and release of virus, it is also an important viral regulatory protein, which can affect transcription factor activity and cytoplasmic signal transduction pathway[35–37].

Our previous research found SHEV ORF3 may mediate the whole membrane protein and basement membrane protein expression of HepG2 cells by down regulating the expression of CLDN6 and FREM1; may cause the apoptosis of HepG2 cells by down regulating the expression of NLRP1; and may regulate the process of lipid metabolism of HepG2 cells by down regulating the expression of APOC3, SCARA3, and DKK1, which laid a foundation for elucidating the important molecular mechanism of the interaction between SHEV and host[38]. We also found that cyclin dependent kinase inhibitor 1B (p27 kip1) was expressed and regulated in HEK293 cells by miR-221 and miR-222, which provided a new entry point for further study of the regulatory mechanism of miRNAs in the infection of SHEV[39]. However, there is no report about the effect of SHEV on the circRNA of host cells.

Recent studies have shown that circRNA molecules are rich in miRNA binding sites, which play the role of miRNA sponge in cells, thereby relieving the inhibition of miRNA on its target genes and increasing the expression level of target genes, this mechanism is called ceRNA mechanism[40–43]. circRNAs plays an important regulatory role in diseases through the interaction of miRNAs associated with diseases[44,45]. Hsa_circRNA_0068871, as the sponge of miR-181a-5p, can target and regulate the expression of FGFR3, activate STAT3 and promote the development of bladder cancer, which can be
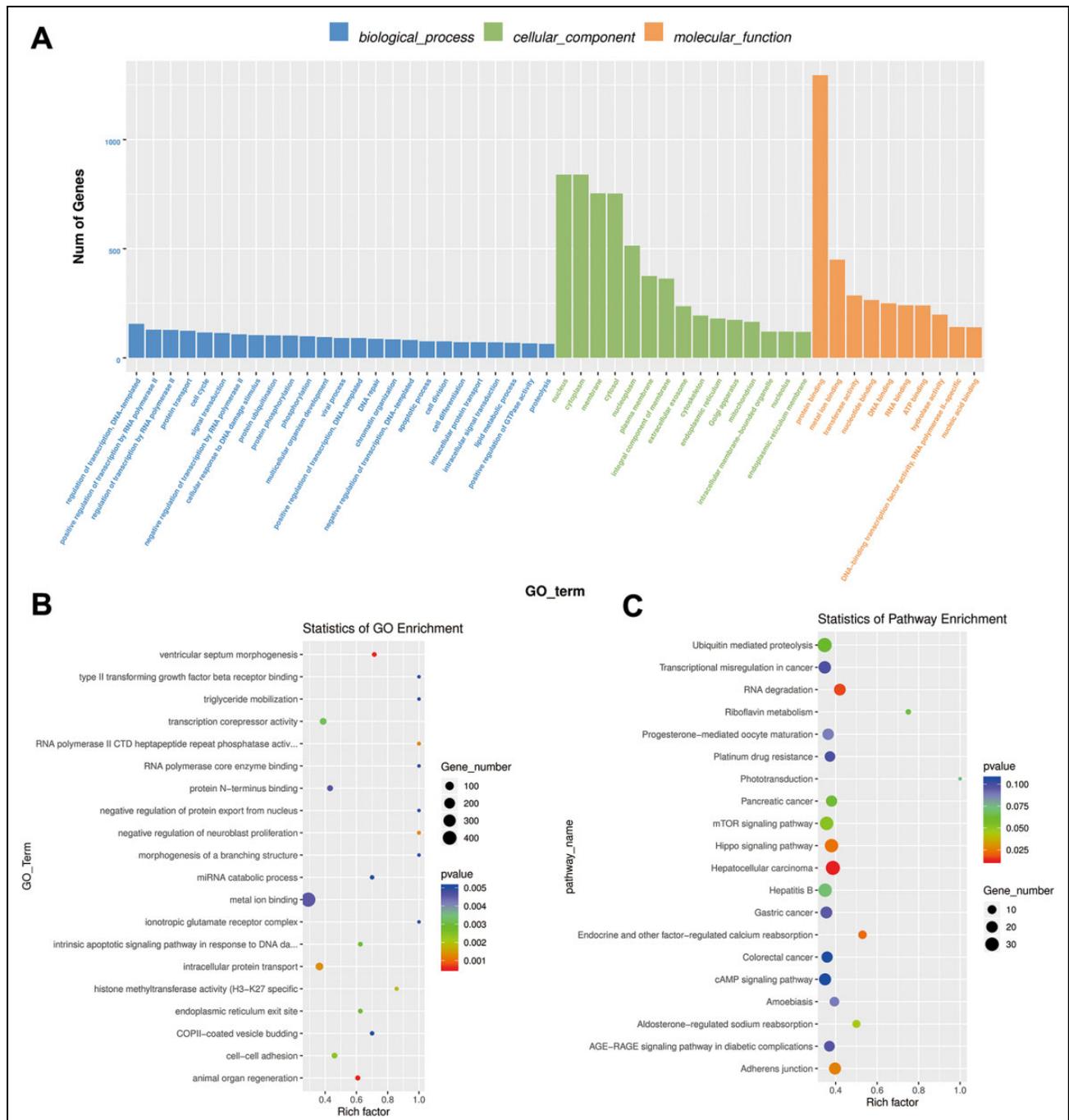
Figure 5. GO and KEGG enrichment analysis of differentially expressed circRNA-hosting gene. (A) The histogram of GO enrichment analysis results represented the number distribution of differential circRNAs-hosting genes on GO term enriched by biological_process, cellular_component and molecular_function. (B) Scatter plot of GO enrichment analysis of differentially expressed circRNAs-hosting genes. X-axis represented Rich factor, Y-axis represented GO_term. (C) Scatter plot of KEGG enrichment analysis differential expression of circRNAs-hosting genes. The size of the dot represented the number of genes with significant difference matching S gene number to a single KEGG, the color of the dot represented the P value. X-axis represented Rich factor, Y-axis represented pathway_name.

used as a potential biomarker[46]. It is found that hsa_-circ_0027089 can be used as a biomarker of hepatitis B virus-related liver cancer, which will be of great significance for the diagnosis of hepatitis B virus-related liver cancer[47]. As the sponge of miR-141-3p, hsa_circRNA_100338 regulated the expression of RHEB and activated mTOR signaling pathway, which affected the poor prognosis of hepatitis B-related hepatocellular carcinoma (HCC) patients[48]. circRNA_100338 and miR-141-3p play key antagonistic roles in the regulation of invasion potential of HCC cells, indicating that circRNA is

a biomarker and a new target for the diagnosis and treatment of HCC in clinic[49] . Then, in this study, we performed KEGG functional enrichment analysis of differential circRNAs and found that there was a pathway of HCC (ko05225) in top20, which suggests that our next research will explore the important role of key circRNAs in the process of liver cancer
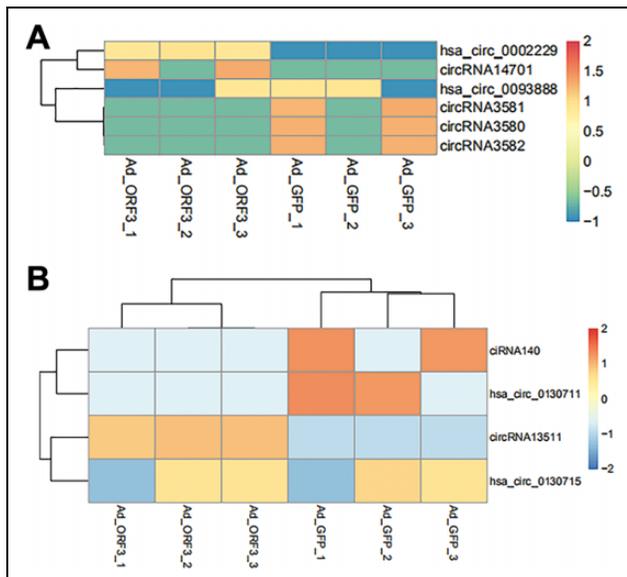
Figure 6. GO and KEGG functional enrichment analysis to predict the circRNAs involved in triglyceride mobilization and riboflavin metabolism affected by ORF3. (A) Heatmap of the 6 circRNAs of triglyceride mobilization affected by ORF3. (B) Heatmap of the 4 circRNAs of riboflavin metabolism affected by ORF3.

mediated by SHEV ORF3. CircRNAs as an endogenous competitive RNA (ceRNA) affects the post transcriptional regulation function of miRNA, providing a basis for subsequent experimental verification.

Triglycerides are the components of lipids. They are the fat molecules formed by glycerol and three long-chain fatty acids, the triglycerides synthesized in vivo are mainly in the liver[50,51]. ORF3 protein is an important regulatory protein of SHEV, which affects the activity of transcription factors and cytoplasmic signal transduction pathway, and may cause lipid metabolism disorder of HepG2 cells[38]. Lipid metabolism disorder will cause jaundice, which is the typical clinical symptoms of swine hepatitis E. In HepG2 cells, we used bioinformatics methods for the first time to predict that there were hsa_circ_0002229, circRNA14701, hsa_circ_0093888, circRNA3581, circRNA3580, and circRNA3582 involved in triglyceride mobilization. When SHEV infection leads to jaundice, liver function is damaged and bile excretion is blocked, and lipid digestion, absorption and metabolism are affected[34,52,53]. The molecular mechanism of ORF3 affecting host lipid metabolism is not completely clear, we will further explore it. It was found that riboflavin can affect the synthesis, transport and decomposition of lipids, and maintain the normal transport of fat in the liver[54,55]. When riboflavin is deficient, the secretion of apolipoprotein B in HepG2 cells is down regulated, which leads to the disruption of lipid balance in HepG2 cells, and finally leads to lipid metabolism disorder[56,57]. Based on the important role of riboflavin in lipid metabolism, the use of riboflavin in the treatment of lipid metabolism related diseases has become a new research hotspot, we also used bioinformatics methods

Table 4. The Summary of the Back-Spliced Junction reads and circRNAs-Hosting Genes Number.

Sample	Ad_GFP_1	Ad_GFP_2	Ad_GFP_3	Ad_ORF3_1	Ad_ORF3_2	Ad_ORF3_3
Candidate back-spliced junction reads	252701(0.33%)	221793(0.30%)	240977(0.34%)	269371(0.34%)	257282(0.32%)	266920(0.36%)
Confident post reads	12202(0.02%)	10372(0.01%)	10980(0.02%)	11033(0.01%)	10049(0.01%)	10140(0.01%)
CircRNA number	5007	4554	4754	4711	4602	4254
CircRNA-hosting gene num	2684	2563	2605	2574	2550	2386
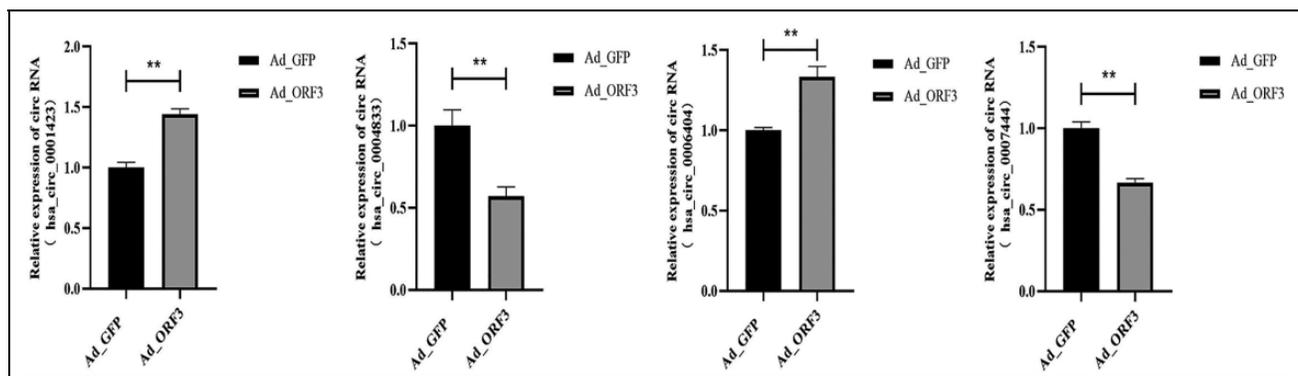
Figure 7. qRT-PCR validation for circRNAs with top10 back splitting junction (BS) sites. "*" indicated *P*-value <0.05, "**" indicated *P*-value <0.01.

for the first time to predict that there were ciRNA410, hsa_circ_0130711, circRNA13511, and hsa_circ_0130715 involved in riboflavin metabolism.

In this work, 1,105 up-regulation and 1,556 down-regulation circRNAs were identified, we investigated the function of differentially expressed circRNAs by GO and KEGG, and found that circRNAs were involved in some very interesting pathways. These findings may contribute to our understanding of the function of SHEV ORF3 and mechanism of SHEV infection.

Conclusion

Taken together, we preliminarily found that ORF3 protein may affect triglyceride activation (GO:0006642) and riboflavin metabolism (ko00740) in HepG2 cells, which provides a scientific basis for further elucidating the effect of ORF3 on host lipid metabolism and the mechanism of SHEV infection.

Authors' Contributions

HJ, FW, and JL contributed to conception and design of the study. HJ, YZ, and ZZ contributed to data acquisition and data analysis. BL, GG, and WL contributed to data interpretation. ML, XG, HZ, and YX drafted the manuscript. HJ, YL, XS, and QH revised the manuscript. All authors have read and approved the final version of the manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical Approval

This study was approved by the Institutional Animal Care and Use Committee of the University of Southwest University (Permit No. IACUC-2015-0219-03).

Statement of Human and Animal Rights

A six week old male New Zealand white rabbit was used in the animal experiment. All animal procedures were reviewed and approved by Institutional Animal Care and Use Committee of Southwest University (Permit No. IACUC-2015-0219-03).

Statement of Informed Consent

Not applicable.

ORCID iD

Hanwei Jiao https://orcid.org/0000-0001-8001-8697

Supplemental Material

Supplemental material for this article is available online.

References

1. Izopet J, Tremeaux P, Marion O, Migueres M, Capelli N, Chapuy-Regaud S, Mansuy JM, Abravanel F, Kamar N, Lhomme S. Hepatitis E virus infections in europe. J Clin Virol. 2019; 120:20–26.

2. Lin KY, Lin PH, Sun HY, Chen YT, Su LH, Su YC, Ho SY, Liu WC, Chang SY, Hung CC, Chang SC. Hepatitis e virus infections among human immunodeficiency virus-positive individuals during an outbreak of acute hepatitis a in taiwan. Hepatology. 2019;70(6):1892–1902.

3. Geng Y, Wang Y. Epidemiology of hepatitis e. Adv Exp Med Biol. 2016;948:39–59.

4. Karlsson M, Norder H, Bergstrom M, Park PO, Karlsson M, Wejstal R, Alsio A, Rosemar A, Lagging M, Mellgren A. Hepatitis E virus genotype 3 is associated with gallstone-related disease. Scand J Gastroenterol. 2019;54(10): 1269–1273.

5. Dogadov DI, Korzaya LI, Kyuregyan KK, Karlsen AA, Kichatova VS, Potemkin IA, Goncharenko AM, Mikhailov MI, Lapin BA. Natural infection of captive cynomolgus monkeys (Macaca fascicularis) with hepatitis E virus genotype 4. Arch Virol. 2019;164(10):2515–2518.

6. Sridhar S, Cheng VCC, Wong SC, Yip CCY, Wu S, Lo AWI, Leung KH, Mak WWN, Cai J, Li X, Chan JFW, et al. Donor-derived genotype 4 hepatitis E virus infection, hong kong, China, 2018. Emerg Infect Dis. 2019;25(3):425–433.

7. Lhomme S, Gallian P, Dimeglio C, Assal A, Abravanel F, Tiberghien P, Izopet J. Viral load and clinical manifestations of hepatitis E virus genotype 3 infections. J Viral Hepat. 2019; 26(9):1139–1142.

8. Narayanan S, Abutaleb A, Sherman KE, Kottilil S. Clinical features and determinants of chronicity in hepatitis E virus infection. J Viral Hepat. 2019;26(4):414–421.

9. Wang X, Li M, Lin Z, Pan H, Tang Z, Zheng Z, Li S, Zhang J, Xia N, Zhao Q.Multifaceted characterization of recombinant protein-based vaccines: an immunochemical toolbox for epitope-specific analyses of the hepatitis E vaccine. Vaccine. 2018;36(50):7650–7658.

10. Jemersic L, Prpic J, Brnic D, Keros T, Pandak N, Dakovic Rode O. Genetic diversity of hepatitis E virus (HEV) strains derived from humans, swine and wild boars in croatia from 2010 to 2017. BMC Infect Dis. 2019;19(1):269.

11. Yue N, Wang Q, Zheng M, Wang D, Duan C, Yu X, Zhang X, Bao C, Jin H. Prevalence of hepatitis E virus infection among people and swine in mainland China: a systematic review and meta-analysis. Zoonoses Public Health. 2019;66(3):265–275.

12. Liu B, Chen Y, Sun Y, Nan Y, Li H, Du T, Hiscox JA, Zhao Q, Zhou EM. Experimental infection of rabbit with swine-derived hepatitis E virus genotype 4. Vet Microbiol. 2019;229: 168–175.

13. Kenney SP, Pudupakam RS, Huang YW, Pierson FW, LeRoith T, Meng XJ. The PSAP motif within the ORF3 protein of an avian strain of the hepatitis E virus is not critical for viral infectivity in vivo but plays a role in virus release. J Virol. 2012;86(10):5637–5646.

14. Wang H, Ji F, Liang H, Gu H, Ning Z, Liu R, Zhang G. A Proline-rich domain in the genotype 4 hepatitis E Virus ORF3 C-terminus is crucial for downstream V105DLP108 immunoactivity. PLoS One. 2015;10(7):e0133282.

15. Qu S, Yang X, Li X, Wang J, Gao Y, Shang R, Sun W, Dou K, Li H. Circular RNA: a new star of noncoding RNAs. Cancer Lett. 2015;365(2):141–148.

16. Zhang Z, Yang T, Xiao J. Circular RNAs: promising biomarkers for human diseases. EBioMedicine. 2018;34:267–274.

17. Zang J, Lu D, Xu A.The interaction of circRNAs and RNA binding proteins: an important part of circRNA maintenance and function. J Neurosci Res. 2020;98(1):87–97.

18. Rong D, Sun H, Li Z, Liu S, Dong C, Fu K, Tang W, Cao H. An emerging function of circRNA-miRNAs-mRNA axis in human diseases. Oncotarget. 2017;8(42):73271–73281.

19. Jiao H, Shuai X, Luo Y, Zhou Z, Zhao Y, Li B, Gu G, Li W, Li M, Zeng H, Guo X, et al. Deep insight into long non-coding RNA and mRNA transcriptome profiling in HepG2 cells expressing genotype IV swine hepatitis E virus ORF3. Front Vet Sci. 2021;8:625609.

20. Hanwei J, Nie X, Zhu H, Li B, Pang F, Yang X, Cao R, Yang X, Zhu S, Peng D, Li Y, et al. miR-146b-5p Plays a Critical Role in the Regulation of Autophagy in per Brucella melitensis-Infected RAW264.7 Cells. Biomed Res Int. 2020; 2020:1953242.

21. Kechin A, Boyarskikh U, Kel A, Filipenko M. cutPrimers: a new tool for accurate cutting of primers from reads of targeted next generation sequencing. J Comput Biol. 2017;24(11): 1138–1143.

22. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9(4):357–359.

23. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015; 12(4):357–360.

24. Kim D, Salzberg SL. TopHat-fusion: an algorithm for discovery of novel fusion transcripts. Genome Biol. 2011;12(8):R72.

25. Zhang XO, Dong R, Zhang Y, Zhang JL, Luo Z, Zhang J, Chen LL, Yang L. Diverse alternative back-splicing and alternative splicing landscape of circular RNAs. Genome Res. 2016;26(9): 1277–1287.

26. Zhang XO, Wang HB, Zhang Y, Lu X, Chen LL, Yang L. Complementary sequence-mediated exon circularization. Cell. 2014;159(1):134–147.

27. Gao Y, Wang J, Zhao F. CIRI: an efficient and unbiased algorithm for de novo circular RNA identification. Genome Biol. 2015;16:4.

28. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139–140.

29. Ma L, Jakobiec FA, Dryja TP. A review of next-generation sequencing (NGS): applications to the diagnosis of ocular infectious diseases. Semin Ophthalmol. 2019;34(4):223–231.

30. Bocklandt S, Hastie A, Cao H. Bionano genome mapping: high-throughput, ultra-long molecule genome analysis system for precision genome assembly and haploid-resolved structural variation discovery. Adv Exp Med Biol. 2019;1129:97–118.

31. Chen X, Yang T, Wang W, Xi W, Zhang T, Li Q, Yang A, Wang T. Circular RNAs in immune responses and immune diseases. Theranostics. 2019;9(2):588–607.

32. Tian H, Fu X, Li W, Huang Y, Sun J, Zhou G, Zhou C, Shen Q, Yang S, Zhang W. Genotype 4 hepatitis E virus prevalent in eastern china shows diverse subtypes. Hepat Mon. 2015;15(6) e25367.

33. Wu Q, An J, She R, Shi R, Hao W, Soomro M, Yuan X, Yang J, Wang J. Detection of genotype 4 swine hepatitis E virus in systemic tissues in cross-species infected rabbits. PLoS One. 2017;12(1):e0171277.

34. Shu Y, Chen Y, Zhou S, Zhang S, Wan Q, Zhu C, Zhang Z, Wu H, Zhan J, Zhang L. Cross-sectional seroprevalence and genotype of hepatitis E virus in humans and swine in a high-density pig-farming area in central China. Virol Sin. 2019;34(4): 367–376.

35. Gouttenoire J, Pollan A, Abrami L, Oechslin N, Mauron J, Matter M, Oppliger J, Szkolnicka D, Dao Thi VL, van der Goot FG, Moradpour D. Palmitoylation mediates membrane association of hepatitis E virus ORF3 protein and is required for infectious particle secretion. PLoS Pathog. 2018;14(12): e1007471.

36. Ding Q, Heller B, Capuccino JM, Song B, Nimgaonkar I, Hrebikova G, Contreras JE, Ploss A. Hepatitis E virus ORF3 is a functional ion channel required for release of infectious particles. Proc Natl Acad Sci U S A. 2017;114(5):1147–1152.

37. Lei Q, Li L, Huang W, Qin B, Zhang S. HEV ORF3 downregulatesCD14 and CD64 to impair macrophages phagocytosis through inhibiting JAK/STAT pathway. J Med Virol. 2019; 91(6):111–119.

38. Xu K, Guo S, Zhao T, Zhu H, Jiao H, Shi Q, Pang F, Li Y, Li G, Peng D, Nie X, et al. Transcriptome analysis of HepG2 cells expressing ORF3 from swine hepatitis E virus to determine the effects of ORF3 on host cells. Biomed Res Int. 2016;2016: 1648030.

39. Cheng Y, Du L, Shi Q, Jiao H, Zhang X, Hao Y, Rong H, Zhang J, Jia X, Guo S, Kuang W, et al. Identification of miR-221 and -222 as important regulators in genotype IV swine hepatitis E virus ORF3-expressing HEK 293 cells. Virus Genes. 2013;47(1):49–55.

40. Ouyang Z, Tan T, Zhang X, Wan J, Zhou Y, Jiang G, Yang D, Guo X, Liu T. CircRNA hsa_circ_0074834 promotes the osteogenesis-angiogenesis coupling process in bone mesenchymal stem cells (BMSCs) by acting as a ceRNA for miR-942-5p. Cell Death Dis. 2019;10(12):932.

41. Xiang Z, Xu C, Wu G, Liu B, Wu D. CircRNA-UCK2 increased TET1 inhibits proliferation and invasion of prostate cancer cells via sponge MiRNA-767-5p. Open Med (Wars). 2019;14:833–842.

42. Wang X, Cao X, Dong D, Shen X, Cheng J, Jiang R, Yang Z, Peng S, Huang Y, Lan X, Elnour IE, et al. Circular RNA TTN acts as a miR-432 sponge to facilitate proliferation and differentiation of myoblasts via the IGF2/PI3K/AKT signaling pathway. Mol Ther Nucleic Acids. 2019;18: 966–980.

43. Yang G, Zhang Y, Yang J. Identification of potentially functional CircRNA-miRNA-mRNA regulatory network in gastric carcinoma using bioinformatics analysis. Med Sci Monit. 2019;25:8777–8796.

44. Lu X, Yu Y, Liao F, Tan S. Homo sapiens circular RNA 0079993 (hsa_circ_0079993) of the POLR2J4 gene acts as an oncogene in colorectal cancer through the microRNA-203a-3p.1 and CREB1 Axis. Med Sci Monit. 2019;25: 6872–6883.

45. Lu X, Liu Y, Xuan W, Ye J, Yao H, Huang C, Li J. Circ_1639 induces cells inflammation responses by sponging miR-122 and regulating TNFRSF13C expression in alcoholic liver disease. Toxicol Lett. 2019;314:89–97.

46. Mao W, Huang X, Wang L, Zhang Z, Liu M, Li Y, Luo M, Yao X, Fan J, Geng J. Circular RNA hsa_circ_0068871 regulates FGFR3 expression and activates STAT3 by targeting miR-181a-5p to promote bladder cancer progression. J Exp Clin Cancer Res. 2019;38(1):169.

47. Zhu K, Zhan H, Peng Y, Yang L, Gao Q, Jia H, Dai Z, Tang Z, Fan J, Zhou J. Plasma hsa_circ_0027089 is a diagnostic biomarker for hepatitis B virus-related hepatocellular carcinoma. Carcinogenesis. 2020;41(3):296–302.

48. Huang XY, Huang ZL, Zhang PB, Huang XY, Huang J, Wang HC, Xu B, Zhou J, Tang ZY. CircRNA-100338 is associated with mTOR signaling pathway and poor prognosis in hepatocellular carcinoma. Front Oncol. 2019;9:392.

49. Huang XY, Huang ZL, Xu YH, Zheng Q, Chen Z, Song W, Zhou J, Tang ZY, Huang XY. Comprehensive circular RNA profiling reveals the regulatory role of the circRNA-100338/miR-141-3p pathway in hepatitis B-related hepatocellular carcinoma. Sci Rep. 2017;7(1):5428.

50. Generoso G, Janovsky C, Bittencourt MS. Triglycerides and triglyceride-rich lipoproteins in the development and progression of atherosclerosis. Curr Opin Endocrinol Diabetes Obes. 2019;26(2):109–116.

51. Dron JS, Hegele RA. Genetics of triglycerides and the risk of atherosclerosis. Curr Atheroscler Rep. 2017;19(7):31.

52. Behloul N, Baha S, Liu Z, Wei W, Zhu Y, Rao Y, Shi R, Meng J. Design and development of a chimeric vaccine candidate against zoonotic hepatitis E and foot-and-mouth disease. Microb Cell Fact. 2020;19(1):137.

53. Sooryanarain H, Meng XJ. Swine hepatitis E virus: Cross-species infection, pork safety and chronic infection. Virus Res. 2020;284:197985.

54. Bian X, Gao W, Wang Y, Yao Z, Xu Q, Guo C, Li B. Riboflavin deficiency affects lipid metabolism partly by reducing apolipoprotein B100 synthesis in rats. J Nutr Biochem. 2019; 70:75–81.

55. Duerden JM, Bates CJ. Effect of riboflavin deficiency on lipid metabolism of liver and brown adipose tissue of sucking rat pups. Br J Nutr. 1985;53(1):107–115.

56. Olpin SE, Bates CJ. Lipid metabolism in riboflavin-deficient rats. 1. effect of dietary lipids on riboflavin status and fatty acid profiles. Br J Nutr. 1982;47(3):577–596.

57. Manthey KC, Chew YC, Zempleni J. Riboflavin deficiency impairs oxidative folding and secretion of apolipoprotein B-100 in HepG2 cells, triggering stress response systems. J Nutr. 2005;135(5):978–982.