



RESEARCH ARTICLE



Human Cytomegalovirus Influences Host circRNA Transcriptions during Productive Infection

Jingui Deng^{1,2} · Yujing Huang¹ · Qing Wang¹ · Jianming Li¹ · Yanping Ma¹ · Ying Qi¹ · Zhongyang Liu¹ · Yibo Li^{3,4} · Qiang Ruan¹

Received: 26 April 2020 / Accepted: 8 July 2020 / Published online: 5 August 2020
© Wuhan Institute of Virology, CAS 2020

Abstract

Human cytomegalovirus (HCMV) is a double-strand DNA virus widely infected in human. Circular RNAs (circRNAs) are non-coding RNAs with most functions of which keep unknown, and the effects of HCMV productive infection on host circRNA transcriptions remain unclear. In this study, we profiled 283 host circRNAs that significantly altered by HCMV productive infection in human embryonic lung fibroblasts (HELFL) by RNA deep sequencing and bioinformatics analysis. Among these, circSP100, circMAP3K1, circPLEKHM1, and circTRIO were validated for their transcriptions and sequences. Furthermore, characteristics of circSP100 were investigated by RT-qPCR and northern blot. It was implied that circSP100 was produced from the sense strand of the SP100 gene containing six exons. Kinetics of circSP100 and SP100 mRNA were significantly different after infection: circSP100 levels increased gradually along with infection, whereas SP100 mRNA levels increased in the beginning and dropped at 24 h post-infection (hpi). Meanwhile, a total number of 257 proteins, including 10 HCMV encoding proteins, were identified potentially binding to cytoplasmic circSP100 by RNA antisense purification (RAP) and mass spectrometry. Enrichment analysis showed these proteins were mainly involved in the spliceosome, protein processing, ribosome, and phagosome pathways, suggesting multiple functions of circSP100 during HCMV infection.

Keywords Human cytomegalovirus (HCMV) · Productive infection · Circular RNA (circRNA) · circSP100 · Transcription

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s12250-020-00275-6>) contains supplementary material, which is available to authorized users.

✉ Qiang Ruan
ruanq@sj-hospital.org

✉ Yujing Huang
huangyj@sj-hospital.org

- ¹ Virology Laboratory, The Affiliated Shengjing Hospital, China Medical University, Shenyang 110004, China
- ² Department of Laboratory, Central Hospital Affiliated to Shenyang Medical College, Shenyang 110024, China
- ³ Department of Gynecology and Obstetrics, The Affiliated Shengjing Hospital, China Medical University, Shenyang 110004, China
- ⁴ Department of Obstetrics, Central Hospital Affiliated to Shenyang Medical College, Shenyang 110024, China

Introduction

Human cytomegalovirus (HCMV) is a beta-herpesvirus that can infect many kinds of human cells, including fibroblasts, endothelial cells, neuronal cells, monocytes/macrophages, and dendritic cells (Sinzger *et al.* 2008). HCMV often causes asymptomatic infection in healthy individuals; however, HCMV is a life-threatening pathogen in congenitally infected newborns and immunocompromised populations (Pass *et al.* 1986; Khoshnevis and Tying 2002). To explore the pathogenesis of HCMV infection, many works that focused on the viral protein are done in the past. It's reported that HCMV IE1 protein promotes viral growth by impairing the antiviral function of SP100 proteins (Kim *et al.* 2011; Tavalai *et al.* 2011; Wagenknecht *et al.* 2015) and viral proteins UL97 and pUL83 (pp65) are beneficial to viral replication by disturbing the role of restriction factors include nuclear proteins IFN- γ inducible protein 16 (IFI16) (Landolfo *et al.*

2016). Recently, it is demonstrated that long non-coding RNA (lncRNA) RNA 4.9 participates in the establishment of HCMV latent infection in CD14(+) monocytes (Rossetto *et al.* 2013). However, the precise pathogenesis of HCMV infection is still unclear, and the relation between circRNAs and HCMV infection has rarely been reported until now.

Circular RNAs (CircRNAs) are a type of non-coding RNAs with a circular structure in which the 5' and 3' terminus are covalently linked together (Memczak *et al.* 2013). CircRNAs were firstly discovered in plant viruses in 1976 (Sanger *et al.* 1976). Nevertheless, circRNAs were regarded as a byproduct of the RNA splicing process concerned with their low abundance and unclear functions for a long time (Nigro *et al.* 1991; Capel *et al.* 1993; Cocquerelle *et al.* 1993). Gradually, more and more evidence revealed that circRNAs exist extensively and play crucial roles in many biological processes in organisms, including serving as miRNA or protein sponges (Salzman *et al.* 2012; Jeck *et al.* 2013; Memczak *et al.* 2013) and protein scaffolding (Du *et al.* 2016, 2017). Research supports that circRNAs regulate protein functions by serving as protein sponges or decoys. CircANRIL triggers nucleolar stress and activation of p53 in human vascular smooth muscle cells and macrophages by sponging pescadillo homologue 1 (PES1) (Holdt *et al.* 2016). Also, it's indicated that the synthase activity of cyclic GMP-AMP (cGAMP) synthase (cGAS) is blocked by circRNA ciacGAS, which decoys cGAS in dormant long-term hematopoietic stem cells (LT-HSCs) (Xia *et al.* 2018). In recent years, some studies are focused on the relationship between circRNAs and herpes virus infections. It was reported that Kaposi's sarcoma herpesvirus (KSHV) infection significantly changed host circRNA transcriptions in primary human umbilical vein endothelial cells (HUVECs) and MC116 (a KSHV-infected B cell line). Those influenced host circRNAs could repress viral gene expressions during infection (Tagawa *et al.* 2018). CircBARTs produced by Epstein-Barr virus, are highly expressed in infected tissue and cell lines with potential functions in viral oncogenesis (Toptan *et al.* 2018). Nonetheless, little is known about the relationship between host circRNAs and HCMV productive infection.

In our study, RNA sequencing (RNA-seq) was performed to investigate the circRNA profiles of mock-infected and HCMV-infected human embryonic lung fibroblasts (HELFs). Host circRNAs significantly influenced by HCMV productive infection were compared and analyzed. Transcriptions and sequences of noteworthy circRNAs (circSP100, circMAP3K1, circPLEKHM1, and circTRIO) were validated. Characteristics of circSP100 in HCMV-infected cells were further confirmed. Moreover, circSP100 potentially binding proteins were purified,

determined, and analyzed, indicating multiple functions of circSP100 during HCMV productive infection.

Materials and Methods

Cell Culture and HCMV Infection

The HELF cell line was acquired from the Shanghai Institute for Biological Sciences, Chinese Academy of Sciences (CAS). HELFs were cultured in Dulbecco's Modified Eagle's Medium (DMEM) containing 10% fetal bovine serum (FBS), 100 units/mL penicillin, and 100 units/mL streptomycin.

HCMV low-passage strain HAN (GenBank accession number: KJ426589.1) was isolated from a urine sample of a 5-month-old infant hospitalized in Shengjing Hospital of China Medical University (Zhao *et al.* 2016). 1×10^7 HELFs were prepared in 10-cm plates and infected with HAN at a multiplicity of infection (MOI) of 3.

RNA Extraction

HCMV infected HELF cells were harvested at 72 h post-infection (hpi). PBS-treated cells were harvested as control. Total RNAs were extracted using TRIzol Reagent (Thermo Fisher) according to the manufactures' instruction. The quality of the total RNAs was estimated using the NanoDrop ND-1000 spectrophotometer and Agilent Bioanalyzer 2100 (Agilent Technologies). Total RNAs were then subjected to ribosomal RNA depletion using Ribo-Zero rRNA Removal Kits (Epicentre). Linear RNAs were removed by RNase R treatment (Epicentre) according to the manufactures' instruction.

RNA Sequencing

RNA fragmentation was performed randomly, resulting in fragmented RNA ranging from approximately 200 nucleotides (nt) to 500 nt. Fragmented RNA was reversely transcribed into cDNA using SuperScriptTM II Reverse Transcriptase (ThermoFisher) according to its protocol. Libraries were constructed using the TruSeq RNA Sample Prep Kit (Illumina). Paired-end RNA sequencing was performed on IlluminaHiSeq 2500 at Biotechnology Co. Ltd., Shanghai, China. The RNA sequencing data have been deposited in the Gene Expression Omnibus database under accession number GSE138836.

circRNAs Prediction

Ribosome RNA reads and reads less than 25 nt or containing adaptor sequences were filtered using Seqtk

(<https://github.com/lh3/seqtk>). Clean reads were mapped to reference genomes (GRCH37/hg19) in the UCSC genome browser (<http://genome.ucsc.edu/>). As described previously (Li and Durbin 2010), the unmapped reads were used for further identification. Sequences from both ends of the unmapped reads were extracted and aligned independently to find unique anchor positions within spliced exons by CIRI (Gao *et al.* 2015). CircRNAs splicings were predicted by anchors that aligned in the reversed orientation (head-to-tail).

The predicted circRNAs were aligned with circBase (<http://circrna.org/>). Using ANNOVAR database (Wang *et al.* 2010), splicing ends of each circRNA candidate were mapped to the genomic regions, and then the results were compared to the RefSeq and UCSC databases to annotate the functional element of circRNA.

circRNAs Data Analysis

Junction reads for each circRNA candidate was normalized to spliced reads per billion mappings (SRPBM) to compute the circRNA expression (Li *et al.* 2015). Differential circRNA transcriptions were calculated according to SRPBM with edgeR (Robinson *et al.* 2010). The q-value (adjusted *P* value) was achieved by controlling the false discovery rate (FDR). Compared to circRNAs in mock-infected cells, circRNAs with > 2 (or < 0.5) fold changes and q-value < 0.05 were determined to be significantly differential circRNAs.

To assess the potential biological function of differential circRNAs, we performed a cluster analysis of the parental genes of differential circRNAs for Gene Ontology (GO, <http://www.geneontology.org/>) and KEGG (<https://www.kegg.jp/>).

RT-PCR and Sequencing

2 µg of total RNA was subjected to RNase R treatment (5 units/µg RNA) at 37 °C for 30 min and reversely transcribed using PrimeScriptTM RT-PCR Kit (TaKaRa). To verify the PCR specificity for circRNAs, we amplified cDNAs of linear SP100 and linear GAPDH as RNase R non-resistant controls. Divergent primers were designed to target circRNA junction sites, while convergent primers were designed to target linear SP100 (NM_001080391) and linear GAPDH (NM_002046.7). Primer information is listed in Supplementary Table S1. 100 ng of cDNA was amplified using the PrimeScriptTM RT-PCR Kit (TaKaRa) according to the manufacturer's protocol. PCR products were analyzed by electrophoresis on 1.5% agarose gel and visualized under UV-light. PCR products with expected sizes were purified using Wizard[®] Plus SV Minipreps DNA Purification System (Promega) and cloned into pCR 2.1

vector using TA CloningTM Kit (Thermo Fisher). Sequencing was performed using Sanger sequencing. Sequences were analyzed using DNAClub and BioEdit software.

Northern Blot

CircSP100 was detected with sense probes and antisense probes. The sense probes target antisense transcripts, while antisense probes target sense transcripts. The specific probes were synthesized and labeled *in vitro* with DIG-UTP (Roche) using DIG Northern Starter Kit (Roche) according to manufacturer's protocol. Primers for generating these probes are listed in Supplementary Table S1.

20 µg of RNA was separated by electrophoresis on agarose gel with 2% formaldehyde and transferred to IMMOBILON-NY+ nylon membranes (MILLIPORE) by capillary transfer. Subsequently, the membrane was rinsed in 2 × SSC for twice briefly and baked at 80 °C for 2 h. The membrane was hybridized with 2 µL of DIG-labelled probes at 60 °C overnight. Stringency wash in 2 × SSC and 0.1% SDS at room temperature was performed twice, followed by a rinse in 2 × SSC and 0.1% SDS at 68 °C twice. The membrane was then incubated with anti-DIG-AP. Specific detection was visualized with the addition of CDP-Star and imaged by Bio-Rad molecular imager chemiDox XRS with ImageLab software (Bio-Rad). Quantities of 28S and 18S rRNAs in each sample were used as the loading controls.

RT-Quantitative PCR (RT-qPCR)

2 µg of total RNA was subjected to RNase R treatment (5 units/µg RNA) at 37 °C for 30 min and reversely transcribed using the PrimeScriptTM RT reagent Kit with gDNA Eraser (TaKaRa). Quantitative PCR (qPCR) was performed on Applied Biosystems 7300 using TB Green Premix *Ex Taq*TMII (TliRNase H Plus) (TaKaRa) according to the manufacturer's protocol. Primer information for circSP100, linear SP100, and linear GAPDH amplification are listed in Supplementary Table S1. Each sample was performed in three replicates. For the kinetic of circSP100 and linear SP100, the result was shown using $2^{-\Delta\Delta CT}$ value.

Subcellular Fractionation

5×10^6 HCMV-infected HELF cells (72hpi) were collected for subcellular fractionations. Nuclear and cytoplasmic fractionations were performed using the Cytoplasmic & Nuclear RNA Purification Kit (NORGEN) according to the manufacturer's protocol. Semiquantitative PCR of circSP100 was performed in cytoplasmic and

nuclear RNA, respectively. Human U2 small nuclear 1 (RNU2-1, NR_002716) and GAPDH were amplified as the nuclear and cytoplasmic reference genes, respectively. The primers are listed in Supplementary Table S1.

RNA Probe Preparation

RNA probe library containing 11 probes antisense to sequence around the circSP100 junction site was used for RNA antisense purification (RAP). RNA probes were designed to be 120 nt in length, imbricating the junction site (Supplementary Fig. S1A). T7 promoter sequence was added to the 5' terminus. DNA sequences of probes before *in vitro* transcription were amplified by RT-PCR, and the primers are listed in Supplementary Table S2. Biotin-labeled RNA probes were produced using HiScribe T7 High Yield RNA Synthesis Kit (NEB) and Biotin-16-UTP (Roche) according to the manufacturer's instruction.

RNA Antisense Purification

Batches of 0.5×10^8 HCMV-infected HELFs (72hpi) were washed once with iced PBS and subjected to crosslink on ice using Spectrolinker UV Crosslinker (8000 J/cm^2 of UV at 254 nm). Cells were collected by centrifugation at $2000g$ for 5 min at 4°C . Meantime, HCMV-infected HELFs without crosslinking were collected as the control for the non-specific binding.

The total cell lysate preparation and RAP were performed as previously described (McHugh *et al.* 2015) with modification. The change included in our study was as followed: ultrasonication was omitted to maintain the intact circRNA structure in cell lysate preparation procedure. Biotinylated RNA probes were used for the hybridization, and 100 μL of Streptavidin DynabeadsMyOne C1 magnetic beads (Thermo Fisher) were used for each sample. Incubation of probes/lysate mixtures and beads/probes/lysate complexes were performed at 45°C with shaking in a shaking bath. At the end of the elution, samples were stored at -80°C for further analysis.

RNA of the elution sample and RNA input sample were magnetically separated and incubated with 1 mg/mL Proteinase K (Beyotime) at 55°C for 1 h. Then, RNA was recovered using SILANE beads (ThermoFisher) according to the manufacturer's instruction. RT-PCR was performed as previously mentioned. Amplification of circSP100 sequence, including junction site, indicated the specific purification of circSP100 and amplification of linear GAPDH was used to monitor the non-specific purification. Here, we also amplified the cDNA transcribed from the circSP100 RAP RNA probes, which referred to the negative control for the RT-PCR.

Silver Staining

SDS-PAGE gels were prepared using the SDS-PAGE reagent kit (Beyotime). Each of 8 μL of RNA elution sample (non-crosslinked and UV crosslink groups) was mixed with 2 μL of $5 \times$ SDS-PAGE Sample Loading Buffer (Beyotime) and heated at 100°C for 10 min, followed by ice incubation for 10 min. Electrophoresis was performed at 50 V for concentration and ran at 100 V for 1.5 h during the separation. Then, silver staining was performed using Protein Stains K Kit (Sangon Biotech) according to the manufacturer's protocols.

Mass Spectrometry

Liquid chromatography-tandem mass spectrometry (LC-MS/MS) was conducted for each of 20 μL of RAP final products (non-crosslinked and UV crosslink groups) as previously described (Kalli and Hess 2012; McHugh *et al.* 2015). Concisely, liquid chromatography was performed on an EASY-nLC 1000 UPLC system. The separation was performed at a constant flow rate of 350 nL/min. The gradient was set as follows: 6%–23% solvent B (16 min), 23%–35% solvent B (8 min), 40%–80% solvent B (4 min), and 80% solvent B (4 min). The peptides were then subjected to nanospray ionization (NSI) at 2.0 kV. Tandem mass spectrometry (MS/MS) was operated in data-dependent acquisition (DDA) mode to alternate between a full scan ($m/z = 350\text{--}1800$) in an Orbitrap Q Exactive Plus mass spectrometer (Thermo). Higher energy collision dissociation (HCD) tandem mass spectrometry scans (resolution = 17,500) were performed after the full scan (resolution = 70,000) on top 20 most abundant ions using automatic gain control (AGC) of 1×10^5 ions. The maximum ion fill-time for MS/MS scans was 100 ms, the HCD-normalized collision energy (NCE) was 28, and the dynamic exclusion time was 15 s.

The resulting data were processed using Proteome Discoverer 1.3, and tandem mass spectra were searched against homo sapiens and human cytomegalovirus entries database (20,590 entries in total). Trypsin was specified as a cleavage enzyme allowing up to 2 missing cleavages. The mass error was set to 10 p.p.m for precursor ions and 0.02 Da for fragment ions. Carbamidomethyl on Cys was specified as a fixed modification, and oxidation on Met was specified as variable modification. Peptide confidence was set at high, and peptide ion score was set > 20 .

Statistic Analysis

All statistics were calculated using Prism Version 8.0 (GraphPad, San Diego, CA 92108, USA). The data in RT-qPCR are presented as mean \pm standard deviation (SD).

Results

Host circRNA Profiles in Mock-Infected and HCMV-Infected HELFs

In our study, a total number of 27,409 and 35,515 distinct host circRNAs were identified in mock-infected and HCMV-infected HELFs, respectively (Fig. 1A). The host circRNA profiles in mock-infected and HCMV-infected HELF cells shared a similar pattern in the distribution of backspliced read accounts. More than 96% of the backspliced read accounts of the identified host circRNAs were less than 100 (Fig. 1A and Supplementary Table S3). Referring to the issues where the circRNAs originate from, it was found that more than 82% of host circRNAs were derived from exons; the others were obtained from introns or intergenic regions.

Generally, host circRNAs were transcribed abundantly and diversely in HELF cells. By analyzing the genomic locations of the parental genes of circRNAs, it was found that the host circRNAs were transcribed extensively throughout the genomes, and no difference of the distribution was found between mock-infected and HCMV-infected cells (Fig. 1B). Alternative splicing often occurs during mRNA transcription. In our results, alternative splicing was also observed for circRNA transcription. Excluding intergenic circRNAs, about 70% of circRNAs have more than two various isoforms of circRNAs (Fig. 1C). In total, 40,274 circRNA variants were produced from 8,138 host genes due to alternative splicing. By comparison, approximately 37% of circRNAs identified in this study (15,798 of 43,200) were annotated in circBase (Fig. 1D). The remainder was novel circRNAs.

Alterations of circRNA Transcriptions by HCMV Infection in HELFs

CircRNAs profiles of mock-infected and HCMV-infected HELFs were compared to address the effects of HCMV infection on circRNA transcriptions. Generally, 19,724 host circRNAs were found in both mock-infected and HCMV-infected cells (Fig. 2A), 15,791 host circRNAs newly appeared after infection, and 7685 circRNAs were found only in mock-infected cells. By selecting circRNAs of which fold-change > 2 (or < 0.5) and q -value < 0.05 ,

283 host circRNAs were identified to be significantly changed by HCMV infection (Fig. 2B). Detailed information is listed in Supplementary Table S3. Among them, transcriptions of 184 host circRNAs were up-regulated significantly after infection, whereas transcriptions of 99 circRNAs were down-regulated (Fig. 2B).

To investigate the potential functions of the differentially expressed circRNAs associated with HCMV infection, the parental genes of these changed circRNAs were assigned to GO term analysis with Gene Ontology. Consequently, the parental genes were enriched among 44 biological processes, 16 molecular functions, and nine cellular components, all of which were filtered by the threshold value of q -value < 0.05 . The top five enrichment of each GO term analysis is presented in Fig. 2C. The parental genes were mainly enriched in cellular processes involved in the release of calcium ion into the cytosol, response to type I interferon, protein ubiquitination, cellular protein modification, and protein phosphorylation. Meanwhile, KEGG enrichment was also performed for parental genes of differential circRNAs. The parental genes were indicated to be enriched in pathways involved in ECM-receptor interaction, focal adhesion, amoebiasis, and PI3K-Akt signaling pathway (Fig. 2D). The results suggested that circRNAs significantly changed by HCMV infection might play potential roles in the regulation of viral entry, apoptosis, and cell proliferation.

Validation of Candidate circRNA Transcriptions Changed by HCMV Productive Infection

Transcription levels of host circRNAs significantly changed by HCMV infection were shown in the heat map (Fig. 3A). Candidate circRNAs with significant differences (circPLEKHM1 and circTRIO) or attractive known functions for their parental genes (circSP100 and circMAP3K1) were selected for further validation by RT-PCR using RNA post-RNase R treatment. To verify the PCR specificity for circRNAs, we amplified cDNAs of linear SP100 and linear GAPDH as RNase R non-resistant controls. As shown in Fig. 3B, since linear RNAs are unresistant to RNase R digestion, products of linear SP100 and GAPDH markedly decreased after RNase R treatment. No decrease was observed for circRNAs products after RNase R digestion.

PCR products were cloned and sequenced then. As indicated in Fig. 3C, the sequences of circRNA junction sites were consistent with our previous RNA-seq results and data in CircInteractome (Dudekula *et al.* 2016). The full sequence information of circSP100, circMAP3K1, circPLEKHM1, and circTRIO was achieved by employing the “rolling” replication strategy during the PCR process (Barrett *et al.* 2015; You *et al.* 2015). By alignment with their transcripts, it was confirmed that the validated

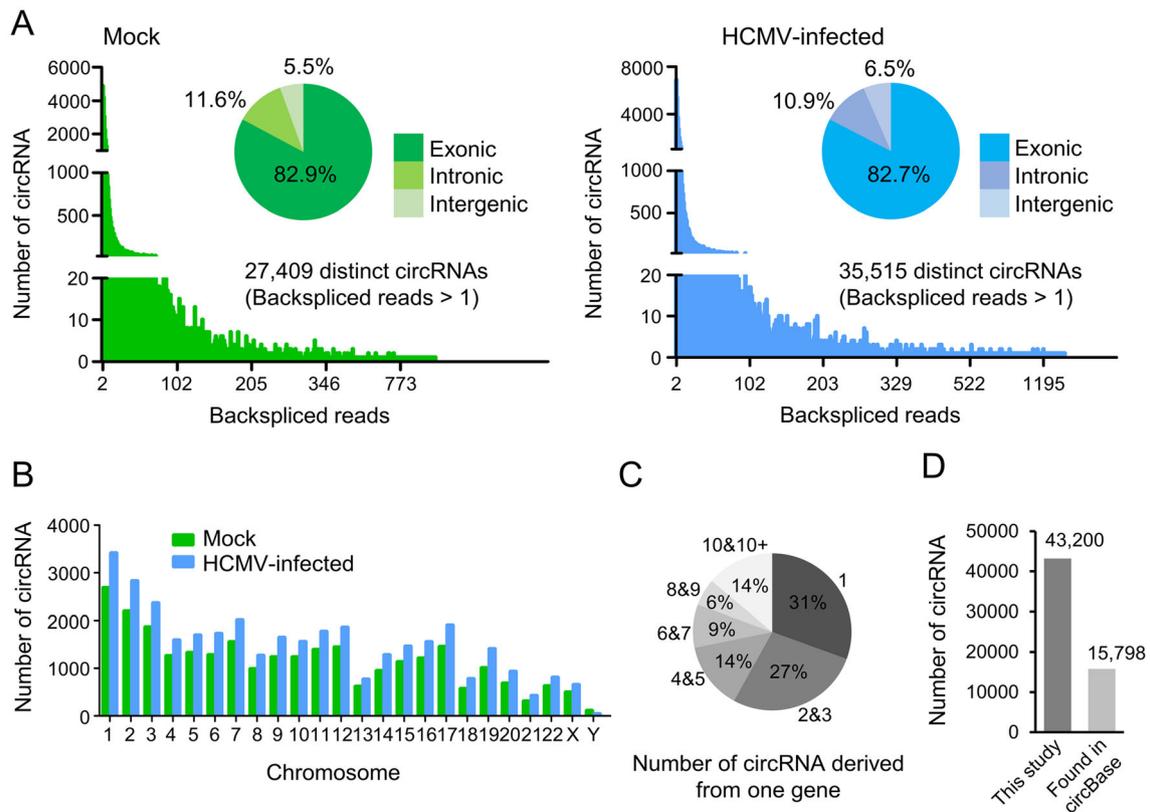


Fig. 1 Host circRNAs profiles in mock-infected and HCMV-infected HELFs. **A** circRNAs expressed in mock-infected and HCMV-infected HELFs; **B** The distribution of circRNAs in the human chromosome between mock-infected and HCMV-infected HELFs; **C** The number of circRNAs derived from one gene. 8,138 host genes produced

40,274 kinds of circRNA in this study (intergenic circRNAs are excluded); **D** The number of circRNAs in this study which were annotated in circBase. 15,798 of 43,200 circRNAs overlapped to circBase.

circRNAs were exonic circRNAs and consisted of multiple exons except for circPLEKHM1 (Fig. 3C).

Characteristics and Potential Functions of circSP100 during HCMV Productive Infection

Northern blot showed that circSP100 was a sense transcript as predicted and increased after infection (Fig. 4A). Sub-cellular fractionation and RT-PCR indicated that circSP100 was preferentially localized in the cytoplasm (Fig. 4B).

As known, productions of linear RNA and circRNA maintain homeostasis at transcriptional levels (Ashwal-Fluss *et al.* 2014; Liang *et al.* 2017). To investigate the relationship between transcriptions of linear SP100 and circSP100, we analyzed the kinetics of SP100 mRNA, and circSP100 at different time points post-infection. It was found that circSP100 levels were increased continuously post-infection, while SP100 mRNA levels were increased before 24 hpi and then dropped gradually (Fig. 4C).

CircRNAs often serves as a scaffold or protein sponge (Holdt *et al.* 2016). For the preliminary investigation of circSP100 potential functions, RAP was conducted to pull

down circSP100 binding proteins (Fig. 5A). The capture of circSP100 was confirmed by RT-PCR (Supplementary Fig. S1B). Proteins pulled down by RAP were detected by silver staining (Fig. 5B). Protein samples were analyzed by mass spectrometry.

In general, 291 proteins were recognized in samples from UV-crosslinked cells and 42 host proteins in samples from the control cells (Fig. 5C and Supplementary Table S4). Excluding the 34 proteins both in UV-crosslinked cells and control cells, 257 proteins were identified to be circSP100-binding proteins, among which 10 were HCMV encoding proteins. About 59% of the circSP100-binding proteins were determined to be cytosolic proteins, whereas 37% were nucleoproteins (Fig. 5D).

Captured proteins with top unique peptide values were listed in Fig. 5E. A total of 257 circSP100 binding proteins were analyzed for KEGG enrichment. The results suggested these circSP100-binding proteins were mainly involved in the spliceosome, protein processing in ER, ribosome, and phagosome pathways (Fig. 5F), indicating the multifunctional characteristic of circSP100 during infection. Notably, 18 of the circSP100-interacting proteins

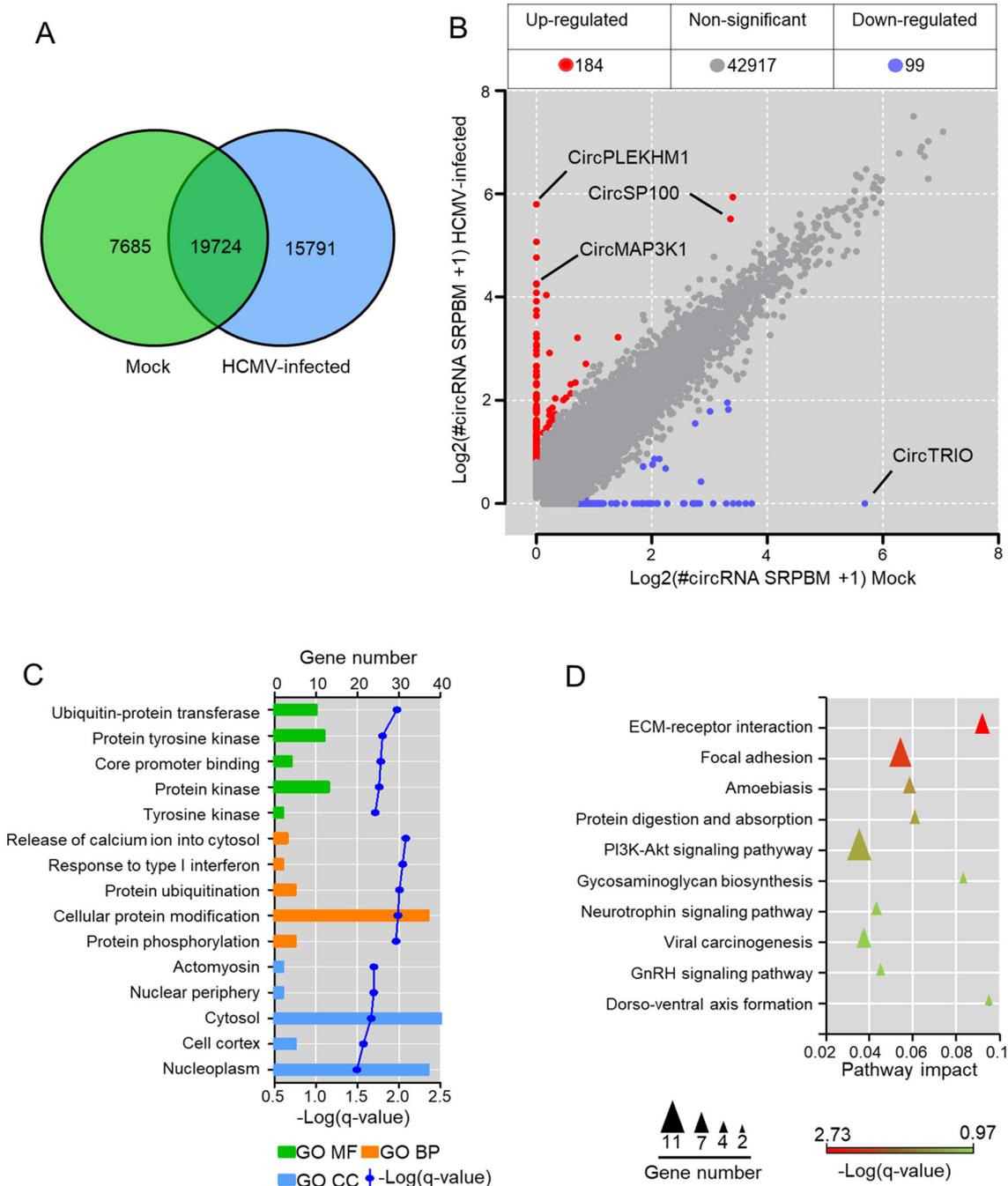


Fig. 2 Alterations of circRNA transcriptions by HCMV infection in HELFs. **A** Venn diagrams of differentially expressed circRNAs in RNA-seq; **B** circRNAs significant differential based on standards of > 2 (or < 0.5) fold-change and q-value < 0.05 between mock-infected and HCMV infected HELFs; **C** GO enrichment of host genes

of the significant change circRNAs. Only the top five of each category (molecular function, biological process, and cellular component), which were arranged by q-value, are presented; **D** KEGG enrichment of host genes of the significant change circRNAs. The top ten of the pathway arranged by q-value is exhibited.

were involved in the spliceosome process (including HNRNPM, HNRNPK, DDX5, DHX15, HNRNPA1, HNRNPC, HNRNPU, PCBP1, SRSF1, SRSF6, SNRPB, SNRPD1, SNRPF, SART1, SNRNP200, SF3B3, TRA2A, and HNRNPA3). It is suggested that circSP100 might involve in spliceosome associated pathway by working as

protein sponges for one or more of these 18 kinds of proteins. Moreover, PRKDC (also known as DNA-PKcs) and XRCC6 (also known as Ku70) were also identified as circSP100 interacting proteins with high unique peptide values in the RAP assay. PRKDC and XRCC6 are role components of DNA-dependent protein kinase (DNA-PK)

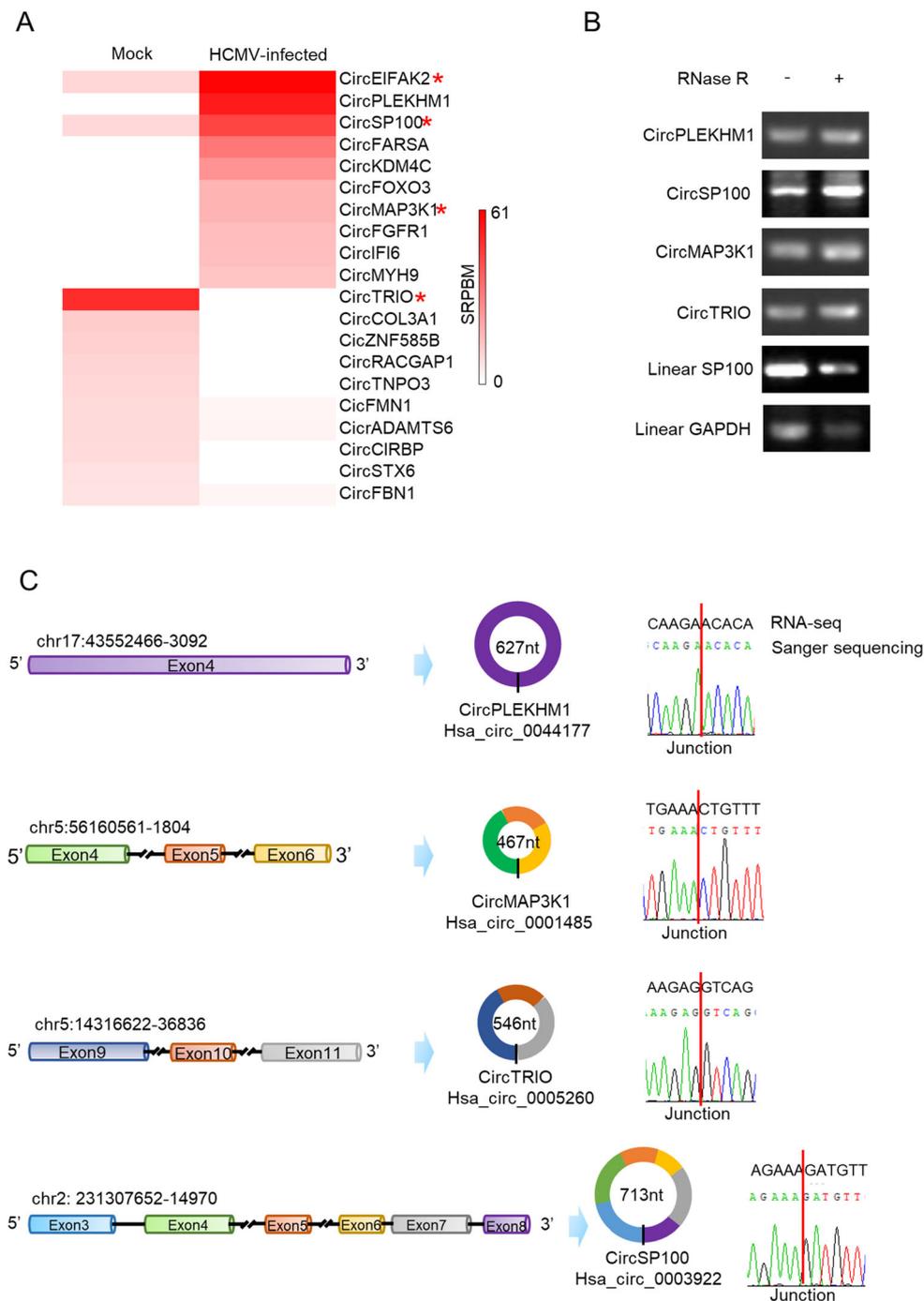


Fig. 3 Validation of candidate circRNA transcriptions changed by HCMV productive infection. **A** Heatmap of the significant up-regulated and down-regulated circRNAs. Red star symbols mark the circRNAs selected for further verification; **B** Verification of cyclization for the selected circRNAs by RNase R treatment. RNase R

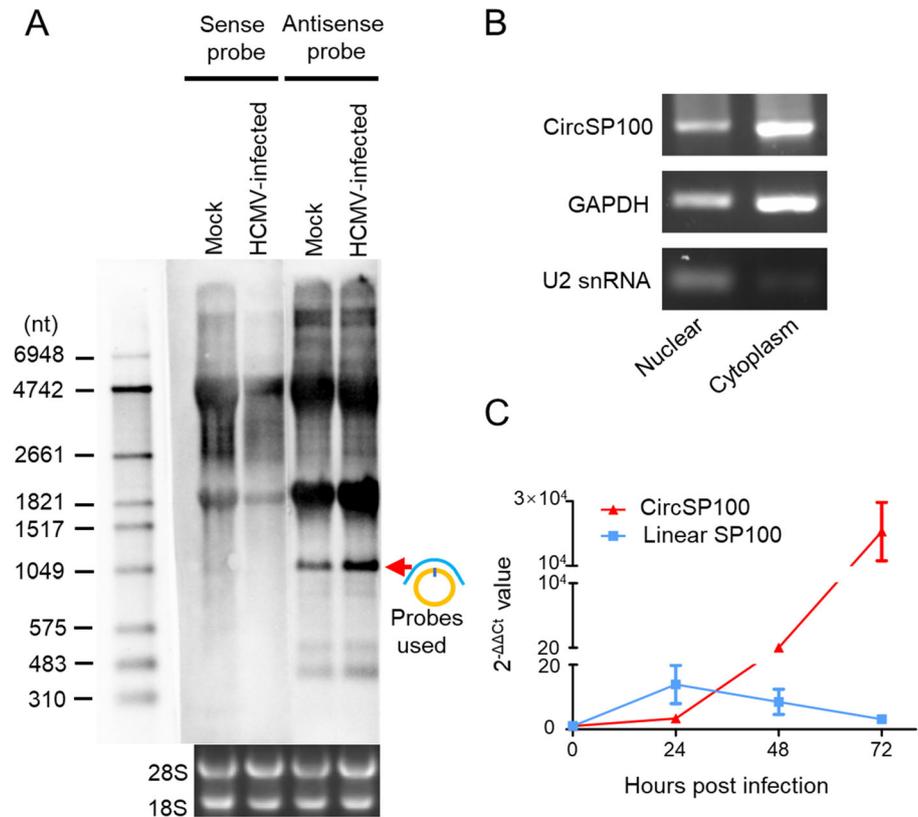
treatment indicates noticeable depletion of linear RNA (SP100 and GAPDH) but not affect the circRNAs (circPLEKHM1, circSP100, circMAP3K1, and circTRIO); **C** Verification of the selected circRNAs by Sanger sequencing. Sanger sequencing indicated all of the chosen circRNAs were consistent with RNA-seq results.

complex that serves as a DNA sensor and induces innate immune response during DNA virus infection (Nussenzweig *et al.* 1996; Gu *et al.* 1997). It was speculated that circSP100 might participate in the formation of

DNA-PK as a scaffold, or inhibit the formation of DNA-PK by consuming PRKDC and XRCC6 as a protein sponge during HCMV infection.

Fig. 4 Characteristics of circSP100 during HCMV productive infection.

A Northern blot of circSP100. The red arrow indicates the circSP100 RNA transcript; **B** Subcellular fractionation of circSP100. GAPDH and U2 snRNA were used as the markers for cytoplasm and nuclear fractionation, respectively; **C** RT-qPCR for the transcript level of circSP100 and linear SP100 mRNA (NM_001080391) in HELFs infected with HCMV at the suggested time points. The data represent the mean \pm SD from three replicates.

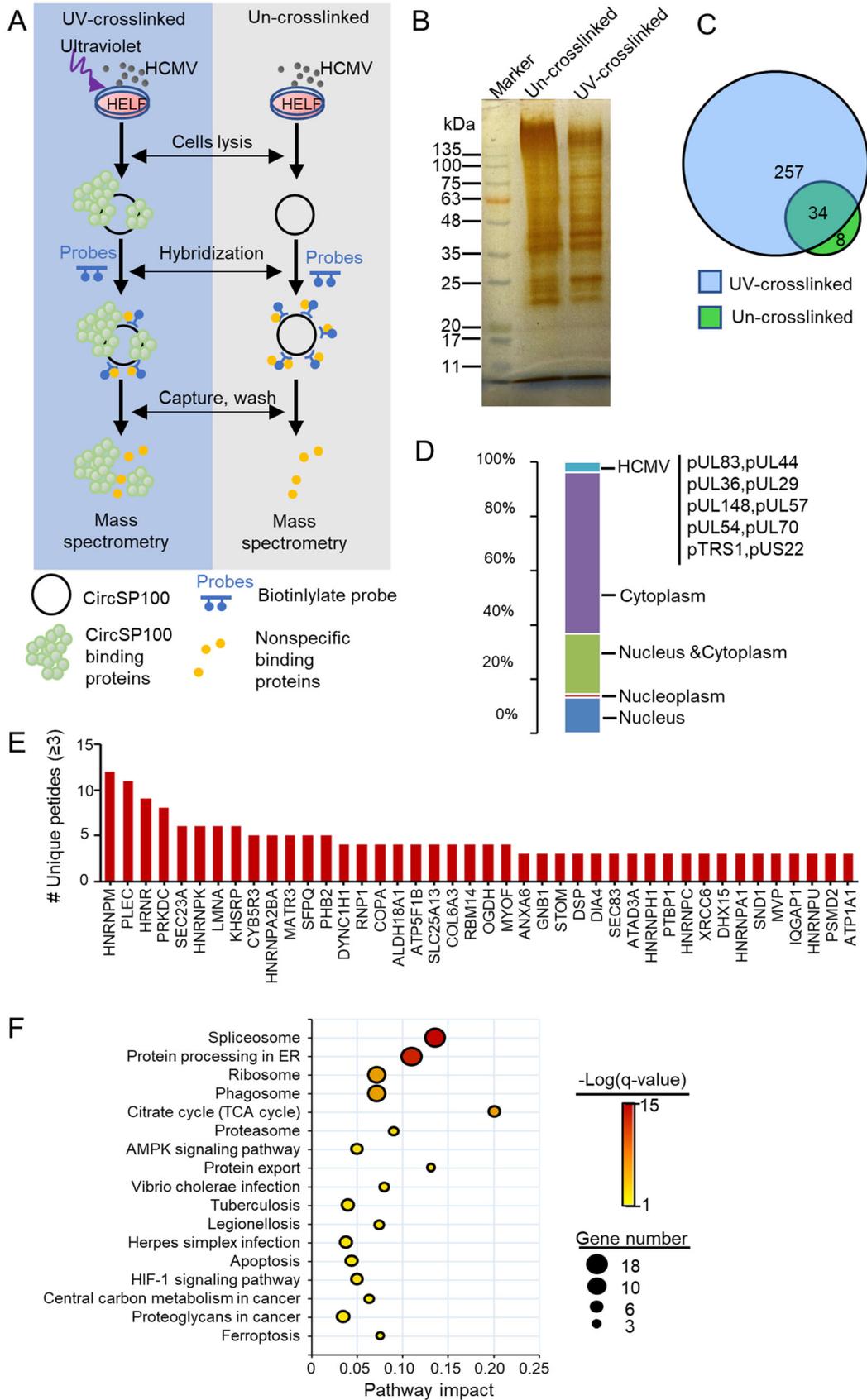


Discussion

HCMV is apt to cause abortive infection in healthy humans. It is estimated that the global seroprevalence of HCMV IgG antibodies is 83% in the general population, 86% in women of reproductive age, and 86% in donors of blood or organs (Zuhair *et al.* 2019). However, HCMV infection is fatal for the immunocompromised population and newborns (Pass *et al.* 1986; Khoshnevis and Tyring 2002). The pathogenesis of HCMV infection remains poorly understood. Many non-coding RNAs have been recently confirmed to be involved in the pathogenesis of HCMV infection, such as microRNAs (Guo *et al.* 2015; Jiang *et al.* 2017) and long non-coding RNAs (lncRNAs) (Mohammad *et al.* 2017). Nevertheless, the relationship between host circRNAs and HCMV infection is unclear.

It has been found that infection of herpesviruses could change host circRNA profiles. Host circRNAs changed by KSHV infection might serve as antiviral factors (Tagawa *et al.* 2018). Recently, it is found that HCMV latent infection could alter host circRNA transcriptions in THP-1 cells (Lou *et al.* 2019). In our study, circRNA profiles were analyzed in HELF cells with or without HCMV productive infection. 184 host circRNAs were up-regulated and 99 down-regulated post HCMV infection (Fig. 2B). The host circRNA profiles of mock-infected and HCMV-infected

cells shared similar characteristics, such as the distribution of backspliced read accounts, the original of host circRNA, and the genomic location of the parental gene of circRNAs. KEGG enrichment showed that the host genes of differential circRNAs were related to ECM-receptor interaction, focal adhesion, amoebiasis, and PI3K-Akt signaling pathway (Fig. 2C and 2D). As known, the PI3K-Akt signaling pathway is activated by many types of cellular stimulation and increases metabolism, growth, proliferation, and cell survival, all of which are helpful to virus replication (Plas and Thompson 2005; Buchkovich *et al.* 2008). During HCMV infection, transient and consistent activation of PI3K-Akt was activated by the attachment between HCMV and viral receptors (Johnson *et al.* 2001), and expression of IE1 and IE2 proteins, respectively (Johnson *et al.* 2001; Buchkovich *et al.* 2008). Consequently, HCMV lytic replication occurs due to the activation of the PI3K-Akt pathway (Johnson *et al.* 2001; Buchkovich *et al.* 2008). ECM-receptor and focal adhesion are reported to be associated with virus entry, spread, and viral persistence. Integrins, the most critical ECM-receptors in cells, were hijacked during viral infection (Hamidi and Ivaska 2018). It's reported that integrin $\alpha 3\beta 1$ promoted KSHV entry into target cells by interacting with KSHV envelope glycoprotein B and inducing the integrin-dependent activation of focal adhesion kinase (Akula *et al.* 2002). Also, it's



◀**Fig. 5** Analysis of circSP100 binding proteins captured by RAP. **A** The strategy of RAP mass spectrometry; **B** Proteins purified with RAP from UV-crosslinked or un-crosslinked lysis extracted from HCMV infected HELF cells; **C** Venn diagrams of proteins purified in RAP; **D** Subcellular location of the proteins that interacted directly with circSP100; **E** The circSP100 binding proteins. Only the proteins of unique peptides ≥ 3 were indicated; **F** KEGG enrichment of the proteins that interacted with circSP100.

suggested that HCMV uses epidermal growth factor receptor (EGFR)- and integrin-dependent signaling pathways to activate signal transducer and activator of transcription (STAT) 1, which promotes the motility, differentiation, and polarization of the infected cell (Collins-McMillen *et al.* 2017). Furthermore, it's pointed out that focal adhesion damage facilitates HCMV infection in HCMV-infected cells (Stanton *et al.* 2007). Collectively, it's suggested that the host gene of the significantly changed circRNAs may be associated to HCMV infection in viral entry, replication, spread, and viral persistence.

Four host circRNAs with significant differences or attractive functions for parental genes were further focused: circPLEKHM1 was only found in HCMV-infected cells at the high transcriptional level, while circTRIO was oppositely found only in HELF cells and eliminated by HCMV productive infection; MAP3K1 is involved in signal transduction events, including the ERK/JNK kinase pathway (Xia *et al.* 1998; Ye *et al.* 2007) and NF-kappa-B pathway (Xia *et al.* 1998); SP100 plays vital roles in tumorigenesis (Held-Feindt *et al.* 2011), gene regulation (Yordy *et al.* 2004, 2005), and antiviral responses to HCMV infection (Kim *et al.* 2011; Tavalai *et al.* 2011; Wagenknecht *et al.* 2015; Landolfo *et al.* 2016). The transcriptions and sequences of these host circRNAs were then validated consistently to those results of RNA sequencing. It was confirmed that HCMV productive infection influences the transcription of host circRNAs. It was speculated that the circRNAs changed by HCMV productive infection might affect the expressions or functions of their parental genes associated with infection.

In our results, circSP100 is 713 nt in length and contains six exons (back-spliced at 5' terminus of exon 3 and 3' terminus of exon 8) (Fig. 3C). By northern blotting with sense and antisense probes, it was confirmed that circSP100 was transcribed from the antisense strand of the genome. It seemed that circSP100 might be produced from the same pre-RNA with SP100 mRNA. To address the relationship between circSP100 and its parental gene, the transcriptions of circSP100 and SP100 mRNA were compared at different time points post-infection. The results showed that circSP100 transcriptions increased along with HCMV infection, whereas SP100 mRNA decreased after 24hpi. It was speculated that since circSP100 and SP100

mRNA were produced from the same pre-RNA strand, the homeostasis between circSP100 and SP100 production might be broken upon HCMV infection. The more circSP100 was created, the less SP100 mRNA was transcribed. It would lead to a reduction of SP100 protein expression, benefiting HCMV survival.

CircSP100 was found mainly located in the cytoplasm, indicating it might play roles by binding with cytoplasmic molecules. Hence, we performed RAP to capture potential binding proteins of circSP100. Consequently, 257 proteins were detected binding to circSP100 (Fig. 5C). These proteins involved in multiple bioprocesses, such as spliceosome, protein processing in ER, ribosome, and phagosome. It has been demonstrated that inhibition of canonical mRNA splicing by depleting core spliceosomal components could increase circRNA transcriptions (Liang *et al.* 2017). Eighteen proteins involved in the spliceosome process were found in our RAP results. HCMV infection may induce the interaction between circSP100 and one or more of these eighteen kinds of proteins, resulting in the dysfunction of the spliceosome. Subsequently, the balance of transcription shifted from SP100 mRNA to circSP100, result in the reduction of SP100 proteins and the promotion of HCMV infection, as mentioned above.

Exhilaratingly, DNA-PKcs accompanied by X-ray repair cross-complementing protein 6 (XRCC6), splicing factor, proline-and glutamine-rich (SFPQ), RNA-binding protein 14 (RBM 14), Matrin-3, and Non-POU domain-containing octamer-binding protein (NONO) was found binding to circSP100 in our study. These proteins were involved in the activations of the cGAS-STING-IRF3 pathway against DNA virus infection (Morchikh *et al.* 2017). Therefore, it was speculated that circSP100 might serve as a sponge to DNA-PK, which would block the combination of DNA-PK and HCMV DNA, resulting in inhibition of innate immunity against HCMV infection (Supplementary Fig. S2). In conclusion, it was demonstrated in our study that HCMV productive infection influenced host circRNA transcriptions in HELFs. Also, circSP100 transcribes differently from SP100 mRNA and potentially binds to 257 proteins. Nevertheless, the combination between circSP100 and the purified proteins needs to be proved, and the role of circRNAs during HCMV infection requires further investigation.

Acknowledgements This work was supported by the National Natural Science Foundation of China (81672028 and 81371788).

Author Contributions JD, YH and QR conceived and designed the experiments. JD, QW, JL, YL and ZL performed the experiments. YM and YQ analyzed the data. JD and YH wrote the manuscript and prepared the figures. YM and YQ provided helpful suggestions about the study. QR checked and finalized the manuscript. All authors read and approved the final manuscript.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

Animal and Human Rights This article does not contain any studies with human or animal subjects performed by any of the authors.

References

- Akula SM, Pramod NP, Wang FZ, Chandran B (2002) Integrin alpha3beta1 (CD 49c/29) is a cellular receptor for Kaposi's sarcoma-associated herpesvirus (KSHV/HHV-8) entry into the target cells. *Cell* 108:407–419
- Ashwal-Fluss R, Meyer M, Pamudurti NR, Ivanov A, Bartok O, Hanan M, Evantal N, Memczak S, Rajewsky N, Kadener S (2014) circRNA biogenesis competes with pre-mRNA splicing. *Mol Cell* 56:55–66
- Barrett SP, Wang PL, Salzman J (2015) Circular RNA biogenesis can proceed through an exon-containing lariat precursor. *Elife* 4:e07540
- Buchkovich NJ, Yu Y, Zampieri CA, Alwine JC (2008) The TORrid affairs of viruses: effects of mammalian DNA viruses on the PI3K-Akt-mTOR signalling pathway. *Nat Rev Microbiol* 6:266–275
- Capel B, Swain A, Nicolis S, Hacker A, Walter M, Koopman P, Goodfellow P, Lovell-Badge R (1993) Circular transcripts of the testis-determining gene *Sry* in adult mouse testis. *Cell* 73:1019–1030
- Cocquerelle C, Mascrez B, Hetuin D, Bailleul B (1993) Mis-splicing yields circular RNA molecules. *Faseb j* 7:155–160
- Collins-McMillen D, Stevenson EV, Kim JH, Lee BJ, Cieply SJ, Nogalski MT, Chan GC, Frost RW 3rd, Spohn CR, Yurochko AD (2017) Human cytomegalovirus utilizes a nontraditional signal transducer and activator of transcription 1 activation cascade via signaling through epidermal growth factor receptor and integrins to efficiently promote the motility, differentiation, and polarization of infected monocytes. *J Virol* 91:e00622–17
- Du WW, Yang W, Liu E, Yang Z, Dhaliwal P, Yang BB (2016) Foxo3 circular RNA retards cell cycle progression via forming ternary complexes with p21 and CDK2. *Nucl Acids Res* 44:2846–2858
- Du WW, Fang L, Yang W, Wu N, Awan FM, Yang Z, Yang BB (2017) Induction of tumor apoptosis through a circular RNA enhancing Foxo3 activity. *Cell Death Differ* 24:357–370
- Dudekula DB, Panda AC, Grammatikakis I, De S, Abdelmohsen K, Gorospe M (2016) CirInteractome: a web tool for exploring circular RNAs and their interacting proteins and microRNAs. *RNA Biol* 13:34–42
- Gao Y, Wang J, Zhao F (2015) CIRI: an efficient and unbiased algorithm for de novo circular RNA identification. *Genome Biol* 16:4
- Gu Y, Jin S, Gao Y, Weaver DT, Alt FW (1997) Ku70-deficient embryonic stem cells have increased ionizing radiosensitivity, defective DNA end-binding activity, and inability to support V(D)J recombination. *Proc Natl Acad Sci USA* 94:8076–8081
- Guo X, Huang Y, Qi Y, Liu Z, Ma Y, Shao Y, Jiang S, Sun Z, Ruan Q (2015) Human cytomegalovirus miR-UL36-5p inhibits apoptosis via downregulation of adenine nucleotide translocator 3 in cultured cells. *Arch Virol* 160:2483–2490
- Hamidi H, Ivaska J (2018) Every step of the way: integrins in cancer progression and metastasis. *Nat Rev Cancer* 18:533–548
- Held-Feindt J, Hattermann K, Knerlich-Lukoschus F, Mehdorn HM, Mentlein R (2011) SP100 reduces malignancy of human glioma cells. *Int J Oncol* 38:1023–1030
- Holdt LM, Stahringer A, Sass K, Pichler G, Kulak NA, Wilfert W, Kohlmaier A, Herbst A, Northoff BH, Nicolaou A, Gabel G, Beutner F, Scholz M, Thiery J, Musunuru K, Krohn K, Mann M, Teupser D (2016) Circular non-coding RNA ANRIL modulates ribosomal RNA maturation and atherosclerosis in humans. *Nat Commun* 7:12429
- Jeck WR, Sorrentino JA, Wang K, Slevin MK, Burd CE, Liu J, Marzluff WF, Sharpless NE (2013) Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA* 19:141–157
- Jiang S, Huang Y, Qi Y, He R, Liu Z, Ma Y, Guo X, Shao Y, Sun Z, Ruan Q (2017) Human cytomegalovirus miR-US5-1 inhibits viral replication by targeting Geminin mRNA. *Virol Sin* 32:431–439
- Johnson RA, Wang X, Ma XL, Huang SM, Huang ES (2001) Human cytomegalovirus up-regulates the phosphatidylinositol 3-kinase (PI3-K) pathway: inhibition of PI3-K activity inhibits viral replication and virus-induced signaling. *J Virol* 75:6022–6032
- Kalli A, Hess S (2012) Effect of mass spectrometric parameters on peptide and protein identification rates for shotgun proteomic experiments on an LTQ-orbitrap mass analyzer. *Proteomics* 12:21–31
- Khoshnevis M, Tyring SK (2002) Cytomegalovirus infections. *Dermatol Clin* 20(291–299):vii
- Kim YE, Lee JH, Kim ET, Shin HJ, Gu SY, Seol HS, Ling PD, Lee CH, Ahn JH (2011) Human cytomegalovirus infection causes degradation of Sp100 proteins that suppress viral gene expression. *J Virol* 85:11928–11937
- Landolfo S, De Andrea M, Dell'Oste V, Gugliesi F (2016) Intrinsic host restriction factors of human cytomegalovirus replication and mechanisms of viral escape. *World J Virol* 5:87–96
- Li H, Durbin R (2010) Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* 26:589–595
- Li Y, Zheng Q, Bao C, Li S, Guo W, Zhao J, Chen D, Gu J, He X, Huang S (2015) Circular RNA is enriched and stable in exosomes: a promising biomarker for cancer diagnosis. *Cell Res* 25:981–984
- Liang D, Tatomer DC, Luo Z, Wu H, Yang L, Chen LL, Cherry S, Wilusz JE (2017) The output of protein-coding genes shifts to circular RNAs when the pre-mRNA processing machinery is limiting. *Mol Cell* 68(940–954):e943
- Lou YY, Wang QD, Lu YT, Tu MY, Xu X, Xia Y, Peng Y, Lai MM, Zheng XQ (2019) Differential circRNA expression profiles in latent human cytomegalovirus infection and validation using clinical samples. *Physiol Genomics* 51:51–58
- McHugh CA, Chen CK, Chow A, Surka CF, Tran C, McDonel P, Pandya-Jones A, Blanco M, Burghard C, Moradian A, Sweredoski MJ, Shishkin AA, Su J, Lander ES, Hess S, Plath K, Guttman M (2015) The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature* 521:232–236
- Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, Maier L, Mackowiak SD, Gregersen LH, Munschauer M, Loewer A, Ziebold U, Landthaler M, Kocks C, le Noble F, Rajewsky N (2013) Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* 495:333–338
- Mohammad AA, Costa H, Landazuri N, Lui WO, Hultenby K, Rahbar A, Yaiw KC, Soderberg-Naucler C (2017) Human cytomegalovirus microRNAs are carried by virions and dense bodies and are delivered to target cells. *J Gen Virol* 98:1058–1072
- Morchikh M, Cribier A, Raffel R, Amraoui S, Cau J, Severac D, Dubois E, Schwartz O, Bennasser Y, Benkirane M (2017) HEXIM1 and NEAT1 long non-coding RNA form a multi-

- subunit complex that regulates DNA-mediated innate immune response. *Mol Cell* 67:387–399.e385
- Nigro JM, Cho KR, Fearon ER, Kern SE, Ruppert JM, Oliner JD, Kinzler KW, Vogelstein B (1991) Scrambled exons. *Cell* 64:607–613
- Nussenzweig A, Chen C, da Costa Soares V, Sanchez M, Sokol K, Nussenzweig MC, Li GC (1996) Requirement for Ku80 in growth and immunoglobulin V(D)J recombination. *Nature* 382:551–555
- Pass RF, Hutto C, Stagno S, Britt WJ, Alford CA (1986) Congenital cytomegalovirus infection: prospects for prevention. *Ann N Y Acad Sci* 477:123–127
- Plas DR, Thompson CB (2005) Akt-dependent transformation: there is more to growth than just surviving. *Oncogene* 24:7435–7442
- Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26:139–140
- Rossetto CC, Tarrant-Elorza M, Pari GS (2013) Cis and trans acting factors involved in human cytomegalovirus experimental and natural latent infection of CD14 (+) monocytes and CD34 (+) cells. *PLoS Pathog* 9:e1003366
- Salzman J, Gawad C, Wang PL, Lacayo N, Brown PO (2012) Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. *PLoS ONE* 7:e30733
- Sanger HL, Klotz G, Riesner D, Gross HJ, Kleinschmidt AK (1976) Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures. *Proc Natl Acad Sci USA* 73:3852–3856
- Sinzger C, Digel M, Jahn G (2008) Cytomegalovirus cell tropism. *Curr Top Microbiol Immunol* 325:63–83
- Stanton RJ, McSharry BP, Rickards CR, Wang EC, Tomasec P, Wilkinson GW (2007) Cytomegalovirus destruction of focal adhesions revealed in a high-throughput Western blot analysis of cellular protein expression. *J Virol* 81:7860–7872
- Tagawa T, Gao S, Koparde VN, Gonzalez M, Spouge JL, Serquina AP, Lurain K, Ramaswami R, Uldrick TS, Yarchoan R, Ziegelbauer JM (2018) Discovery of Kaposi's sarcoma herpesvirus-encoded circular RNAs and a human antiviral circular RNA. *Proc Natl Acad Sci USA* 115:12805–12810
- Tavalai N, Adler M, Scherer M, Riedl Y, Stamminger T (2011) Evidence for a dual antiviral role of the major nuclear domain 10 component Sp100 during the immediate-early and late phases of the human cytomegalovirus replication cycle. *J Virol* 85:9447–9458
- Toptan T, Abere B, Nalesnik MA, Swerdlow SH, Ranganathan S, Lee N, Shair KH, Moore PS, Chang Y (2018) Circular DNA tumor viruses make circular RNAs. *Proc Natl Acad Sci USA* 115:E8737–E8745
- Wagenknecht N, Reuter N, Scherer M, Reichel A, Muller R, Stamminger T (2015) Contribution of the major ND10 proteins PML, hDaxx and Sp100 to the regulation of human cytomegalovirus latency and lytic replication in the monocytic cell line THP-1. *Viruses* 7:2884–2907
- Wang K, Li M, Hakonarson H (2010) ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucl Acids Res* 38:e164
- Xia Y, Wu Z, Su B, Murray B, Karin M (1998) JNKK1 organizes a MAP kinase module through specific and sequential interactions with upstream and downstream components mediated by its amino-terminal extension. *Genes Dev* 12:3369–3381
- Xia P, Wang S, Ye B, Du Y, Li C, Xiong Z, Qu Y, Fan Z (2018) A circular RNA protects dormant hematopoietic stem cells from DNA sensor cGAS-mediated exhaustion. *Immunity* 48(688–701):e687
- Ye B, Yu WP, Thomas GM, Haganir RL (2007) GRASP-1 is a neuronal scaffold protein for the JNK signaling pathway. *FEBS Lett* 581:4403–4410
- Yordy JS, Li R, Sementchenko VI, Pei H, Muise-Helmericks RC, Watson DK (2004) SP100 expression modulates ETS1 transcriptional activity and inhibits cell invasion. *Oncogene* 23:6654–6665
- Yordy JS, Moussa O, Pei H, Chaussabel D, Li R, Watson DK (2005) SP100 inhibits ETS1 activity in primary endothelial cells. *Oncogene* 24:916–931
- You X, Vlatkovic I, Babic A, Will T, Epstein I, Tushev G, Akbalik G, Wang M, Glock C, Quedenau C, Wang X, Hou J, Liu H, Sun W, Sambandan S, Chen T, Schuman EM, Chen W (2015) Neural circular RNAs are derived from synaptic genes and regulated by development and plasticity. *Nat Neurosci* 18:603–610
- Zhao F, Shen ZZ, Liu ZY, Zeng WB, Cheng S, Ma YP, Rayner S, Yang B, Qiao GH, Jiang HF, Gao S, Zhu H, Xu FQ, Ruan Q, Luo MH (2016) Identification and BAC construction of Han, the first characterized HCMV clinical strain in China. *J Med Virol* 88:859–870
- Zuhair M, Smit GSA, Wallis G, Jabbar F, Smith C, Devleeschauwer B, Griffiths P (2019) Estimation of the worldwide seroprevalence of cytomegalovirus: a systematic review and meta-analysis. *Rev Med Virol* 29:e2034