Identification and characterization of a novel group of natural anti-sense transcripts from RNA1.2 gene locus of human cytomegalovirus

Bing-Yang Liu[1,2], Yan-Ping Ma[2], Ying Qi[2], Yu-Jing Huang[2], Zhong-Yang Liu[2], Meng-Qing Lu[2], Qiang Ruan[2]

[1]Department of Endocrinology, Affiliated Shengjing Hospital, China Medical University, Shenyang, Liaoning 110004, China;
[2]Virus Laboratory, Affiliated Shengjing Hospital, China Medical University, Shenyang, Liaoning 110004, China.

Abstract

Background: Natural anti-sense transcripts (NATs), which are transcribed from the complementary DNA strand of annotated genes, exert regulatory function of gene expression. Increasing studies recognized anti-sense transcription widespread throughout human cytomegalovirus (HCMV) genome, whereas the anti-sense transcription of RNA1.2 gene locus has never been investigated. In this study, the transcription of the RNA1.2 anti-sense strand was investigated in clinically isolated HCMV strain.

Methods: Strand-specific high-through RNA-sequencing (RNA-seq) was performed to find possible anti-sense transcripts (ASTs). For analyzing and visualization of RNA-seq data sets, Integrative Genomics Viewer software was applied. To confirm these possibilities, Northern blotting and rapid amplification of cDNA ends (RACE) were used.

Results: Transcription of the opposite strand of RNA1.2 gene locus was detected by RNA-sequencing using RNAs extracted from human embryonic lung fibroblasts infected with HCMV clinical isolate HAN. At least three HCMV NATs, named RNA1.2 AST 1, RNA1.2 AST2, and RNA1.2 AST3, were characterized by Northern blotting and RACE analyses. These RNA1.2 ASTs orientated from the complementary strand of RNA1.2 locus during the late phase of HCMV infection. The 5′- and 3′-termini of these transcripts were located within the opposite sequence of the predicted RNA1.2 gene.

Conclusion: A cluster of novel NATs was transcribed from the opposite sequence of the HCMV RNA1.2 gene region.

Keywords: Human cytomegalovirus; RNA1.2; Natural anti-sense transcript

Introduction

Natural anti-sense transcripts (NATs) were first observed more than 30 years ago.[1,2] Since then, huge amount of NATs have been discovered by microarrays and high-throughput sequencing through different species, for example, virus,[3,4] human,[5,6] plants,[7,8] and other eukaryotes.[9] Several NATs are characterized and proved to be important modulators during pre-transcriptional, transcriptional, and post-transcriptional regulation of gene expression.[10,11]

Human cytomegalovirus (HCMV) infection is generally asymptomatic in healthy adults but causes life-threatening infections in immunocompromised patients and neonates.[12] Congenital cytomegalovirus (cCMV) infection is a major cause of neonate sensorineural hearing loss and central nervous system impairments that affect vision, language development, cognition, and vestibular function.[13] Infants with cCMV may suffer from intrauterine growth restriction, thrombocytopenia, splenomegaly, hepatomegaly, hepatitis.[14] Several studies suggested that cytomegalovirus infection is associated with an increased risk of cardiovascular disease.[15,16] As a member of beta-herpesvirus family, the genome of HCMV is 230 to 240 kb in length[17,18] and encodes RNAs of multi-function.[19] Viral non-coding RNAs (ncRNAs), which have been assumed to be "junk" previously,[20] account for a large part of these multi-functional RNAs. Recently, more and more studies confirmed that ncRNAs play critical roles in adaptive immune responses, gene expression, and viral life cycles.[21-24] However, there remain many herpesvirus ncRNAs to be exposed and functionally characterized.

HCMV genome has been first mapped since 1990.[25] Subsequently, with development of second- and third-generation sequencing technologies, the complexity of HCMV transcriptome is revealed.[4,26] But only partial of its transcripts has been fully annotated.[27] Four major long non-coding RNAs (lncRNA) of HCMV, RNA1.2, RNA2.7, RNA4.9, and RNA5.0, account for over 65% of polyadenylated (polyA) viral transcripts.[4] Current finding about RNA1.2 indicates that it may also encode a polypeptide.[18]

Access this article online

Quick Response Code:

Website: www.cmj.org

DOI: 10.1097/CM9.0000000000000299

Correspondence to: Dr. Qiang Ruan, Virus Laboratory, Affiliated Shengjing Hospital, China Medical University, Shenyang, Liaoning 110004, China
E-Mail: ruanq@sj-hospital.org

Copyright © 2019 The Chinese Medical Association, produced by Wolters Kluwer, Inc. under the CC-BY-NC-ND license. This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Chinese Medical Journal 2019;132(13)

Received: 12-03-2019 **Edited by:** Xin Chen

In the present study, transcription from anti-sense orientation of HCMV RNA1.2 locus was first identified in HCMV-infected cells by RNA-sequencing (RNA-seq). The structures of a cluster of RNA1.2 ASTs were confirmed by rapid amplification of cDNA ends (RACE) and Northern blotting analyses. The findings may offer experimental basis for further research on biologic functions of this S-AS transcript pair in HCMV infection.

Methods

Cell cultures and virus

Human embryonic lung fibroblasts (HELFs) were a generous gift from Professor Min-Hua Luo at Wuhan Institute of Virology, Chinese Academy of Sciences. HELFs were cultured in Dulbecco modified eagle medium supplemented with 10% fetal bovine serum and penicillin-streptomycin (100 IU/mL and 100 mg/mL) at 37°C in a humidified atmosphere containing 5% CO_2.

HCMV clinical isolate HAN (GenBank accession number: KJ426589.1) used in our study was propagated in HELFs and titrated by plaque-forming assay. The bacterial artificial chromosome (BAC) containing HAN genome with an SV40-GFP tag was constructed previously.[28] Virus was rescued by electroporation of HAN-BAC DNA into HFLFs according to a standard procedure established in our laboratory. Virus stocks, named as vHAN-BAC, were prepared by ultracentrifugation at 55,000×*g* for 1 h in SW32 Ti rotor. Virus titers were determined by $TCID_{50}$ assay as follows: after scoring GFP-positive or negative wells in a 96-well dish using a fluorescent microscope, the scores were matched to the TCID50 chart. HCMV isolate HAN and vHAN-BAC were stored at −80°C for subsequent assays.

RNA preparation

HELFs were inoculated with HCMV HAN or vHAN-BAC at a multiplicity of infection of 3. After incubation for 3 h, the supernatant was removed and the cells were washed with phosphate-buffered saline once. Subsequently, the cells were incubated in fresh culture medium for 24, 48, and 72 h post-infection (hpi). The uninfected cells were harvested in parallel as controls. Total RNA was isolated from infected or uninfected cells using Trizol reagent (Invitrogen, Carlsbad, CA, USA) and treated with TURBO DNA-free kit (Ambion, Austin, TX, USA) to remove possibly contaminated DNA. Quantity and purity of the RNA preparations were estimated by optical density with Nanodrop 1000 spectrophotometer (Life Technology, Gaithersburg, MD, USA). The integrity and size of the isolated RNA was analyzed by formaldehyde agarose gel electrophoresis. All RNA preparations were stored at –80°C till use.

Strand-specific RNA sequencing

Ribosomal RNA was removed from RNA preparations using the NEBNext rRNA depletion kit (NEB, Ipswich, MA, USA). Following purification, both the poly(A)+ and poly (A)−RNA fractions were fragmented into small pieces. The cDNA libraries were generated with the protocol for the mRNA-Seq sample preparation kit (NEB). Paired-end sequencing was performed with Illumina Hiseq 4000 at Amogene Biotech Co., Ltd (Xiamen, China) following the vendor's recommended protocol. Based on the clean data obtained, sequencing reads were aligned to HCMV HAN genome by STAR V20201. Fragments per kilobase per million mapped reads (FPKM) of each gene and its anti-sense transcript (AST) were calculated to estimate their expression levels using the RNA-seq data sets, which include three biologic replicates for each time point. For visualization of RNA-seq data sets, alignment files were normalized by the number of mapped reads, and, after conversion to wiggle [.wig] files, displayed using Integrative Genomics Viewer (IGV) software (http://software.broadinstitute.org/software/igv/).

Northern blotting

To identify RNA1.2 ASTs, RNA preparations from cells infected with vHAN-BAC at 24, 48, and 72 hpi were analyzed by Northern blotting with the RNA of uninfected cells as control. Northern blotting was performed as described previously[29] according to recommended protocol of Digoxigenin (DIG) Northern Starter kit (Roche, Indianapolis, IN, USA). The riboprobes for detection of RNA1.2 and RNA1.2 ASTs were synthesized and labeled *in vitro* with Dig-UTP (Roche) according to the manufacturer's protocol provided by HiScribe™ T7 High Yield RNA synthesis kit (NEB). Specific primers for generating these riboprobes are listed in Table 1. For analyzing RNA1.2 ASTs and RNA1.2, 30 and 5 μg of total RNAs were respectively separated by electrophoresis on 2% formaldehyde containing agarose gel and transferred onto a positively charged nylon membrane by capillary transfer. After baking at 80°C for 2 h, the membrane was hybridized with riboprobes for 16 h at 66°C. Following washing process, the membrane was incubated with anti-DIG antibody conjugated to alkaline phosphatase. Specific blots were visualized with addition of the substrate of CDP-Star (Roche) and analyzed by Bio-Rad molecular imager chemiDoc XRS with ImageLab software (Bio-Rad, Hercules, CA, USA). To ensure equal RNA loading, quantities of 28S and 18S rRNAs in each RNA preparation were used as the loading controls.

Rapid amplification of cDNA ends

To determine 5′- and 3′-ends of RNA1.2 ASTs, the 5′- and 3′-RACE were preformed according to the manuals of SMARTer™ RACE 5′/3′ kit (Takara Bio, Mountain View, CA, USA). The RNA preparation of vHAN-BAC-infected cells at 72 hpi was used to produce first-strand cDNAs. The cDNAs for 5′- and 3′-RACE were synthesized with SMARTer transcriptase using oligo (dT)-adaptor primers. A positive control cDNA was also synthesized using the Mouse Heart Total RNA provided by the kit. Gene-specific primers (gsp) for the 5′- and 3′-RACE reactions were designed based on sequence data of RNA-seq data sets using Primer Premier 5.0. For 3′-RACE, the nested polymerase chain reaction (PCR) amplifications were performed using universal primers supplied in the kit

Table 1: Primers used in the present study.

Experiments	Primers	Primer sequence (5′ → 3′)	Position at the 5′*
Northern Blotting	RNA1.2-forward	TGCGAACTTTTAGGAACCAG	7441
	RNA1.2-reverse	TAATACGACTCACTATAGGG†GTTTATTTTGGCTGCTGCTA	7117
	RNA1.2 AST-forward	GTGAAGATTTTATGACGTTGGT	7453
	RNA1.2 AST-reverse	TAATACGACTCACTATAGGG†CGCTTAATTTGTTTAGAGCC	7818
RACE	UPM	CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGT	
	UPM short	CTAATACGACTCACTATAGGGC	
	5′-RNA1.2 AST-gsp1	*GATTACGCCAAGCTT*‡CCCGCCAACGAGAAAACCGAAAAGTA	7669
	5′-RNA1.2 AST-gsp2	*GATTACGCCAAGCTT*‡TGTTTAGAGCCGCACGCTTGACAACG	7809
	3′-RNA1.2 AST-gsp1	*GATTACGCCAAGCTT*‡GTGATGGTTATTCGCTGGTTCGTTCT	7504
	3′-RNA1.2 AST-gsp2	*GATTACGCCAAGCTT*‡TTTTCTCGTTGGCGGGTCAGAGGTAG	7654

* Genomic positions were referred to that of HCMV clinical isolate HAN (GenBank accession number: KJ426589.1). † Nucleotides shown by underlined letters are promoter sequence of T7 RNA polymerase for promoting synthesis of RNA probes *in vitro*. ‡ Nucleotides shown by italic letters are overlaps with the pRACE vector to facilitate cloning of RACE PCR products. AST: Anti-sense transcript; HCMV: Human cytomegalovirus; PCR: Polymerase chain reaction; RACE: Rapid amplification of cDNA ends; UPM: Universal *Primer* A mix.

and 3′-RNA1.2 AST-gsp1/2. For 5′-RACE, nested PCR was carried out with universal primers, which is complementary to the sequence of adaptor involved in the 5′-end of cDNAs, in combination with 5′-RNA1.2 AST-gsp1/2. The primary reaction was performed with the cycling parameters: five cycles of 94°C for 30 s, 68°C for 2 min; five cycles of 94°C for 30 s, 65°C for 30 s, and 72°C for 2 min; 25 cycles of 94°C for 30 s, 63°C for 30 s, and 72°C for 2 min. The second-round PCR using following program: 94°C for 4 min, 25 cycles of 94°C for 30 s, 63°C for 30 s, and 72°C for 2 min, with a final extension at 72°C for 10 min. Primer sequences are provided in Table 1.

Cloning and sequencing

PCR products were electrophoresed on agarose gel, and all of the separated bands were purified respectively using Wizard SV gel and PCR clean-up system (Promega, Madison, WI, USA). The purified PCR products were cloned into pRACE vector supplied in SMARTer™ RACE 5′/3′ kit. The plasmids were transformed into Stellar competent cells. Clones containing the inserts in the recombinant vectors were briefly identified by PCR. The inserted sequences of 8 to 32 randomly selected clones were sequenced using M13 primers (forward primer 5′-TGTAAAACGACGGCCAGT-3′, reverse primer 5′-CAGGAAACAGCTATGACC-3′) by Sanger sequencing method.

BLAST searches and sequence analysis

Standard nucleotide-nucleotide BLAST was performed on the NCBI website. The nucleotide positions mentioned in this study are in reference to the sequence of the HCMV clinical isolate HAN. The open reading frames (ORFs) of identified transcripts were predicted with Editseq in the DNAStar package.

Results

RNA1.2 AST detected in HCMV-infected cells by RNA-seq

To investigate ASTs in HCMV-infected cells, directional RNA-seq was performed in rRNA-depleted total RNAs from cells collected at different time points post-infection (24, 48, and 72 hpi). All RNA-seq reads were aligned to genome of HCMV clinical isolate HAN and calculated FPKM using Cuffdiff with the following library-type fr-firststrand and fr-secondstrand. By this method, reads aligned to the complementary strand of an annotated gene indicated anti-sense transcription. When analyzing virus transcripts, widespread anti-sense transcription (data not shown) was observed. ASTs from RNA1.2 gene region were found in RNAs from cells collected at the three time points and named as RNA1.2 AST. The mapped reads of RNA1.2 and RNA1.2 AST were displayed using IGV software [Figure 1]. Importantly, RNA1.2 AST persisted in infected cells from 24 to 72 hpi, with FPKM increasing from 763.9 to 1481.7.

RNA1.2 ASTs verified by Northern blotting

To verify RNA1.2 ASTs, Northern blotting was performed with 24, 48, and 72 hpi RNA preparations from vHAN-BAC-infected HELFs and RNA from uninfected cells. As shown in Figure 2A, three bands with lengths of ~1100, ~1000, and ~600 nt, respectively, were detected in the 72 hpi RNA using the specific RNA probe of RNA1.2 AST-P. No band was found in 24 hpi, 48 hpi, or uninfected RNAs. The result suggests that a cluster of natural anti-sense RNA is transcribed from the complementary strand of RNA1.2 locus. In addition, a transcript of ~1.2 kb was detected with the probe of RNA1.2-P, which confirmed the transcription of RNA1.2 [Figure 2B].

Ends of RNA1.2 ASTs determined by RACE

To determine the ends of RNA1.2 ASTs, 3′-RACE and 5′-RACE analyses were carried out with the 72 hpi RNA of the HCMV HAN-infected HELFs. As shown in Figure 3, two bands of about 400 and 250 bp were seen in the products of 3′-RACE. Sequencing results of the recovered fragments showed that the 3′-ends of the RNA1.2 ASTs were located at nt 8062 to 8093 and nt 7843 to 7855, downstream of a consensus polyA signal (AATAAA) at nt 7843 to 7848. Three predominant bands with about 500, 350, and 250 bp were found in 5′-RACE amplification of

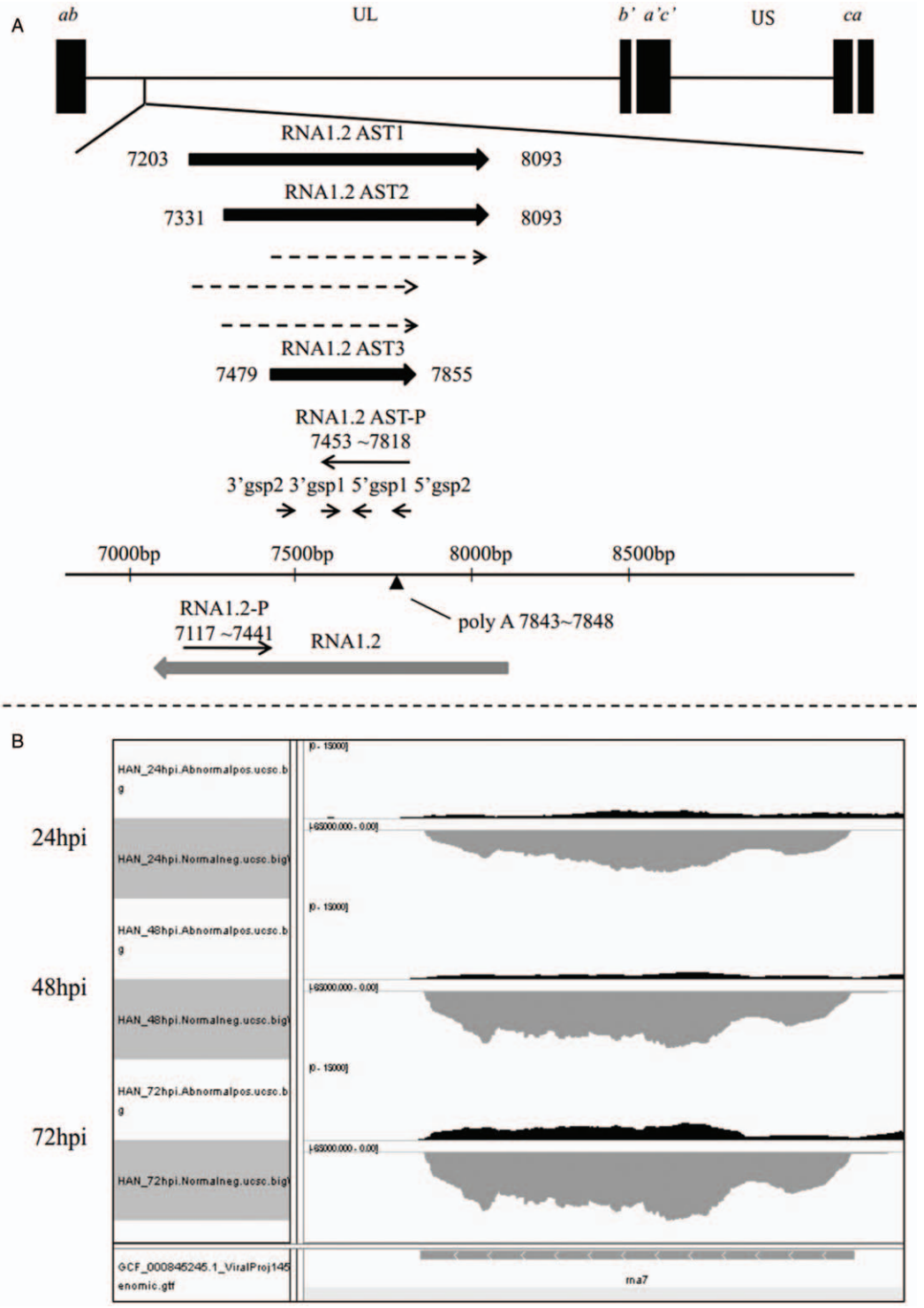

Figure 1: Characterization of RNA1.2 ASTs. (A) Graphic representation of the HCMV genome and RNA1.2 ASTs. The genome is shown in five sections, UL, US, and the inverted repeats (ab/b′a′ and a′c′/ca) shaded black. Sense gene is depicted as gray arrow running from right to left, ASTs as black arrows running from left to right. Arrows shown by dotted line represent indeterminate transcripts of RNA1.2 ASTs. The relative positions of the primers and probes used in this study are indicated. The relative positions are referenced to HCMV clinical isolate HAN (GenBank: KJ426589.1). (B) RNA1.2 gene region is visualized using IGV software at different time points post-infection. The reads in black above the horizon represent RNA1.2 ASTs, whereas the reads in gray below the horizon represent RNA1.2. AST: Anti-sense transcript; HCMV: Human cytomegalovirus.

RNA1.2AST by nested PCR [Figure 4]. Sequencing results of the DNA fragments showed that the RNA1.2 ASTs initiated at nt 7479, nt 7331, and nt 7203 of HCMV HAN genome. These loci are downstream of the putative "TATA" box, initiating from nt 7074.

Analysis of the RNA1.2 ASTs sequences

The locations of the RNA1.2 ASTs, the primers and riboprobes in HCMV genome are shown in Figure 1A. Based on the RACE sequencing results, that is, three

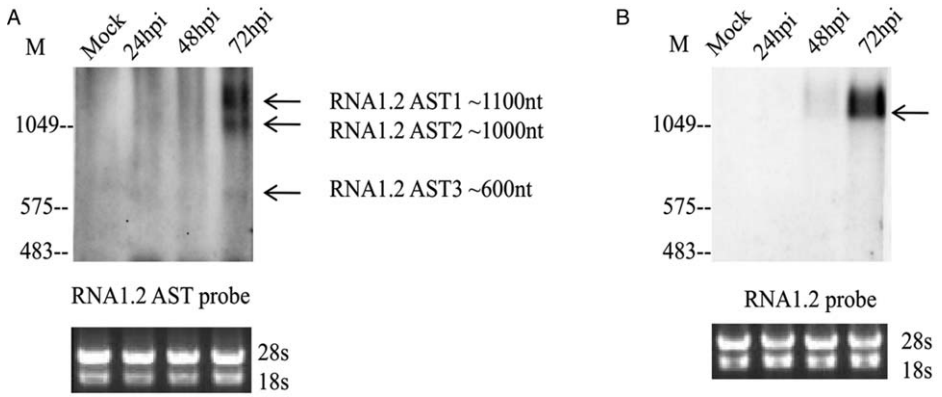

Figure 2: Northern blotting analyses of the RNA1.2 ASTs from HCMV vHAN-BAC-infected HELF. (A) Representative result showing RNA1.2 ASTs. Northern blotting was performed with 30 μg RNA per lane. Arrows represent the transcripts detected. (B) Representative result showing RNA1.2. Northern blotting performed with 5 μg RNA per lane. A 1.2-kb transcript is detected at 48 and 72 hpi. AST: Anti-sense transcript; BAC: Bacterial artificial chromosome; HCMV: Human cytomegalovirus; HELF: Human embryonic lung fibroblasts; M: Marker; Mock: Mock-infected HELF RNA; 24, 48, and 72 hpi: Represent RNAs from cells collected at different time points post-infection.

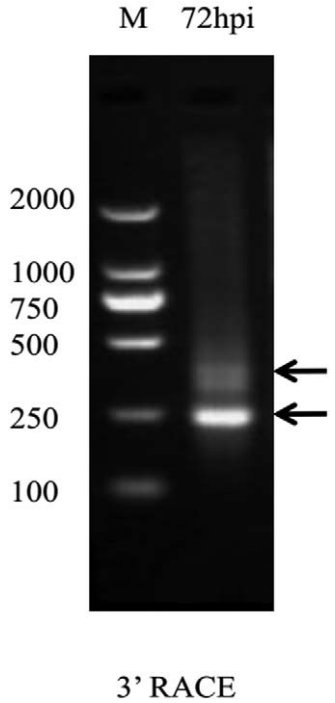

Figure 3: The 3′-RACE result for RNA1.2 ASTs. Nested PCR were performed using the 3′-RACE adaptor primers together with gene-specific primers of 3′-RNA1.2 AST gsp1/3′-RNA1.2 AST gsp2. The predominant bands are about 400 and 250 bp. Products in all the indicated bands (shown by arrows) were recovered, cloned and sequenced. AST: Anti-sense transcript; BAC: Bacterial artificial chromosome; HCMV: Human cytomegalovirus; HELF: Human embryonic lung fibroblasts; PCR: Polymerase chain reaction; RACE: Rapid amplification of cDNA ends; 72 hpi: Experiment performed using 72 hpi RNA from HCMV vHAN-BAC-infected HELF.

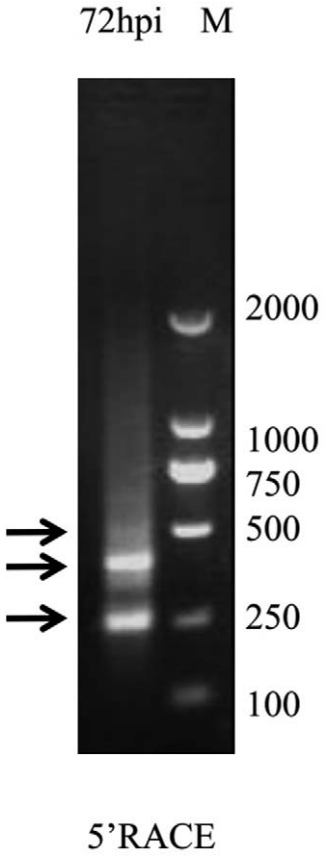

Figure 4: The 5′-RACE result for RNA1.2 ASTs. Nested PCR were performed using the 5′-RACE adaptor primers together with gene-specific primers of 5′-RNA1.2 AST gsp1/5′-RNA1.2 AST gsp2. The predominant bands are about 500, 350, and 250 bp. Products of all the predominant bands indicated by black arrows were recovered, cloned, and sequenced. AST: Anti-sense transcript; BAC: Bacterial artificial chromosome; HCMV: Human cytomegalovirus; HELF: Human embryonic lung fibroblasts; PCR: Polymerase chain reaction; RACE: Rapid amplification of cDNA ends; 72 hpi: Experiment performed using 72 hpi RNA from HCMV vHAN-BAC-infected HELF.

5′-termini and two 3′-ends, there should be six transcripts of RNA1.2 ASTs in theory. However, only three kinds of RNA1.2 ASTs of ~1100, ~1000, and ~600 nt were detected by Northern blotting. So, it is ascertained that at least three transcripts are present, namely RNA1.2 AST1, 2, and 3, as shown in Table 2 and Figure 1. According to RNA-seq result, there is a conventional potential promoter element (TATA box) present 130 bp upstream the RNA initiation site of RNA1.2 AST1. No canonical ORFs are predicted by Editseq analyses in the RNA1.2 ASTs.

Discussion

NATs are transcripts from the complementary DNA strand of annotated genes and lack protein-coding potential.[30] NATs have been indicated to exist in various

Table 2: Analysis of RNA1.2 ASTs sequences.

5′-termini		3′-termini	Length prior to polyadenylation	Length of bands detected by Northern blotting	Name of transcripts
A nt 7203	a	nt 8062–8093	859–890 nt	~1100 nt	RNA1.2 AST1
	b	nt 7843–7855	640–652 nt	Not detected	–
B nt 7331	a	nt 8062–8093	731–762 nt	~1000 nt	RNA1.2 AST2
	b	nt 7843–7855	512–524 nt	Not detected	–
C nt 7479	a	nt 8062–8093	583–614 nt	Not detected	–
	b	nt 7843–7855	364–376 nt	~600 nt	RNA1.2 AST3

There are three 5′-termini (A, B, and C) and two 3′-ends (a and b) based on the RACE sequencing results. Theoretically there should be six transcripts of RNA1.2 ASTs, while there are only three bands detected by Northern blotting analyses. Considering the lengths of these transcripts and bands detected by Northern blotting, there are only three possible transcripts of RNA1.2 ASTs, namely RNA1.2 AST1, AST2, and AST3. The other transcripts are indeterminate and not detected by Northern blotting. AST: Anti-sense transcript; RACE: Rapid amplification of cDNA ends.

organisms by cDNA library screening, microarrays and RNA-seq.[3,4,6,9] NATs, which were thought to be transcriptional noise arising from "junk DNA," are recently recognized as important gene modulators.[31] Multiple examples showed that NATs exert their function through different stages of gene expression process, such as transcriptional initiation, transcript processes, and post-transcriptional processes.[32-36] Increasing evidence confirmed that transcript pairs encoded by both strands of gene loci exist extensively through HCMV genome. Zhang *et al*[37] sequenced cDNA libraries of immediate-early, early, and late infection phases and found that over 35% of annotated transcripts in HCMV have ASTs. Others, who carried out RNA-seq of HCMV late-phase RNAs, reported that ncRNAs transcribed from anti-sense strand to protein-coding loci account for 8.7% of transcripts from these regions.[4] Ma *et al*[38] established first cDNA library of HCMV clinical strain HAN in 2011. By Northern blotting and RACE, three NATs, namely UL31 AST, UL83 AST, and UL87 AST, are described and characterized.[39-41] Recently, Balazs *et al*[26] discovered eight novel NATs by long-read sequencing of HCMV transcriptome. In the present study, at least three differentially regulated transcripts were found to originate from the anti-sense strand of the HCMV RNA1.2 gene. According to their positions with respect to RNA1.2 gene locus, RNA1.2 ASTs are internal NATs as they are fully covered by the sense transcript.

RNA1.2 is one of the most abundant viral transcripts, which counts for 7.9% of viral polyA RNA transcription.[4] Although RNA1.2 is considered as an lncRNA, it is confirmed by mass spectrometry to encode a 99aa protein.[18] What is more, the coding potential is highly conserved across different HCMV strains. However, functions of RNA1.2 have not been studied extensively so far.[19] Some NATs are atypical transcripts, while others are mRNA like with polyA tails. The later are localized in the cytoplasm, where they tend to interact with their overlapping sense RNAs and regulate the sense RNAs.[42] RNA1.2 ASTs are polyadenylated RNA confirmed by sequences of 3′-RACE products. Considering the locus specificities and sequence characteristics of RNA1.2 ASTs and RNA1.2, it is reasonable to hypothesize that one role for RNA1.2 ASTs is to regulate the sense strand expression.

It is found that the abundance of RNA1.2 ASTs is much lower than that of RNA1.2 in our study. Generally, NATs are 10-time lower in abundance than sense transcripts.[43,44] Recent evidences indicated that even very low abundance NATs can exert biologic effects.[45,46] Interestingly, Wang *et al*[47] reported very low abundance non-coding RNAs, as low as four copies per cell, can regulate target genes. Moreover, some ASTs alter the abundance of sense transcripts, while other NATs exert biologic function rather than affecting the amount of sense mRNA. For example, the synthesis of zinc-finger E-box-binding homeobox 2 (*ZEB2*) protein is controlled by Zeb2-NAT.[48] This regulation increases *ZEB2* mRNA translation efficiency without changing the abundance of it. It is known that maturation of RNA plays a crucial role in cellular accumulation and function of an lncRNA. Zong *et al*[49] found that a NAT from metastasis-associated lung adenocarcinoma transcript 1 (MALAT1) locus directly promotes 3′-end cleavage of MALAT1 lncRNA. The interaction of this S-AS pair may facilitate the recognition and cleavage by RNase P. Functions of the RNA1.2 ASTs are being investigated.

Usually, a cluster of NATs terminate at the same 3′-terminus,[3,40,41] but there are exceptions considering the complex transcription patterns in HCMV genome.[27] Other studies showed that during herpes simplex virus-1 infection, about 50% of the transcribed host cell genes experience transcriptional read-through.[5,50] In the above studies, putative NATs of host genome, which transcribed 1 kb downstream from polyA signal, were defined as read-through transcripts. RNA1.2 ASTs are a cluster of transcripts with two different 3′-termini. The 3′-terminus of RNA1.2 AST3 is located at nt 7843 to 7855 downstream a canonical polyA signal at nt 7843 to 7848, while the RNA1.2 AST1 and RNA1.2 AST2 terminate at over 250 nt downstream of the same polyA signal. The different transcription termination sites of RNA1.2 ASTs may due to unknown mechanism. Besides, it is inferred that RNA1.2 AST1 and RNA1.2 AST2 are possible read-through transcripts as those of NATs found in herpes simplex virus-1-infected cells.[5,50] It would be worthwhile investigating the different regulatory mechanisms involved in termination of RNA1.2 AST1/2 *vs*. RNA1.2 AST3 in future.

Last but not least, RNA1.2 ASTs are transcripts with different initiation sites at nt 7479, nt 7331, and nt 7203 of HCMV HAN genome. A putative "TATA" box initiates from nt 7074 upstream the transcription start sites (TSSs) of RNA1.2 ASTs. Some genes have classical TATA box at about 30 bp upstream TSS. Other genes are regulated by distant promoters with activity influenced by epigenetic context. High-CG promoters, which are characterized by their overlap with CpG islands, are often associated with multiple TSSs.[51] In the case of RNA1.2 ASTs, high CG content is observed upstream TSS of nt 7203 but not TSS of nt 7479 or nt 7331. Accordingly, it is possible that different regulatory mechanisms or signals may be involved in transcription initiation of RNA1.2 ASTs.

To summarize, it is conceivable that three forms of HCMV RNA1.2 ASTs are transcribed during the HCMV late infection phase. All these transcripts are different at their 3′- and 5′-termini, which indicate they are differentially regulated by unknown mechanism. No canonical ORFs are predicted in RNA1.2 ASTs. RNA1.2 and RNA1.2 ASTs form a new S-AS transcript pair in HCMV genome. Since this is an exploratory study, the role of RNA1.2 ASTs in regulating the expression of RNA1.2 gene warrants further investigation.

Funding

This work was supported by grants from the National Natural Science Foundation of China (Nos. 81672028 and 81371788) and the Outstanding Scientific Fund of Shengjing Hospital.

Conflicts of interest

None.

References

1. Itoh T, Tomizawa J. Formation of an RNA primer for initiation of replication of ColE1 DNA by ribonuclease H. Proc Natl Acad Sci U S A 1980;77:2450–2454. doi: 10.1073/pnas.77.5.2450.
2. Krinke L, Mahoney M, Wulff DL. The role of the OOP antisense RNA in coliphage lambda development. Mol Microbiol 1991;5:1265–1272. doi: 10.1111/j.1365-2958.1991.tb01900.x.
3. Majerciak V, Yang W, Zheng J, Zhu J, Zheng ZM. A genome-wide Epstein-Barr virus polyadenylation map and its antisense RNA to EBNA. J Virol 2019;93:e01593-18. doi: 10.1128/JVI.01593-18.
4. Gatherer D, Seirafian S, Cunningham C, Holton M, Dargan DJ, Baluchova K, *et al*. High-resolution human cytomegalovirus transcriptome. Proc Natl Acad Sci U S A 2011;108:19755–19760. doi: 10.1073/pnas.1115861108.
5. Wyler E, Menegatti J, Franke V, Kocks C, Boltengagen A, Hennig T, *et al*. Widespread activation of antisense transcription of the host genome during herpes simplex virus 1 infection. Genome Biol 2017;18:209. doi: 10.1186/s13059-017-1329-5.
6. Naemura M, Kuroki M, Tsunoda T, Arikawa N, Sawata Y, Shirasawa S, *et al*. The long noncoding RNA OIP5-AS1 is involved in the regulation of cell proliferation. Anticancer Res 2018;38:77–81. doi: 10.21873/anticanres.12194.
7. Yuan Y, Chen S. Widespread antisense transcription of Populus genome under drought. Mol Genet Genomics 2018;293:1017–1033. doi: 10.1007/s00438-018-1456-z.
8. Zhao X, Li J, Lian B, Gu H, Li Y, Qi Y. Global identification of Arabidopsis lncRNAs reveals the regulation of MAF4 by a natural antisense RNA. Nat Commun 2018;9:5056. doi: 10.1038/s41467-018-07500-7.
9. Wang Z, Xu H, Li T, Wu J, An L, Zhao Z, *et al*. Chicken GHR antisense transcript regulates its sense transcript in hepatocytes. Gene 2019;682:101–110. doi: 10.1016/j.gene.2018.10.001.
10. Sun Y, Li D, Zhang R, Peng S, Zhang G, Yang T, *et al*. Strategies to identify natural antisense transcripts. Biochimie 2017;132:131–151. doi: 10.1016/j.biochi.2016.11.006.
11. Georg J, Hess WR. Widespread antisense transcription in Prokaryotes. Microbiol Spectr 2018;6. doi: 10.1128/microbiolspec.RWR-0029-2018.
12. Britt W. Manifestations of human cytomegalovirus infection: proposed mechanisms of acute and chronic disease. Curr Top Microbiol Immunol 2008;325:417–470. doi: 10.1007/978-3-540-77349-8_23.
13. Rawlinson WD, Boppana SB, Fowler KB, Kimberlin DW, Lazzarotto T, Alain S, *et al*. Congenital cytomegalovirus infection in pregnancy and the neonate: consensus recommendations for prevention, diagnosis, and therapy. Lancet Infect Dis 2017;17:e177–e188. doi: 10.1016/s1473-3099(17)30143-3.
14. Schleiss MR. Congenital cytomegalovirus: impact on child health. Contemp Pediatr 2018;35:16–24.
15. Wang H, Peng G, Bai J, He B, Huang K, Hu X, *et al*. Cytomegalovirus infection and relative risk of cardiovascular disease (ischemic heart disease, stroke, and cardiovascular death): a meta-analysis of prospective studies up to 2016. J Am Heart Assoc 2017;6:e005025. doi: 10.1161/jaha.116.005025.
16. Lv YL, Han FF, Gong LL, Liu H, Ma J, Yu WY, *et al*. Human cytomegalovirus infection and vascular disease risk: a meta-analysis. Virus Res 2017;227:124–134. doi: 10.1016/j.virusres.2016.09.010.
17. Davison AJ, Dolan A, Akter P, Addison C, Dargan DJ, Alcendor DJ, *et al*. The human cytomegalovirus genome revisited: comparison with the chimpanzee cytomegalovirus genome. J Gen Virol 2003;84:17–28. doi: 10.1099/vir.0.18606-0.
18. Stern-Ginossar N, Weisburd B, Michalski A, Le VT, Hein MY, Huang SX, *et al*. Decoding human cytomegalovirus. Science 2012;338:1088–1093. doi: 10.1126/science.1227919.
19. Hancock MH, Skalsky RL. Roles of non-coding RNAs during herpesvirus infection. Curr Top Microbiol Immunol 2018;419:243–280. doi: 10.1007/82_2017_31.
20. Doolittle WF. Is junk DNA bunk? A critique of ENCODE. Proc Natl Acad Sci U S A 2013;110:5294–5300. doi: 10.1073/pnas.1221376110.
21. Canny SP, Reese TA, Johnson LS, Zhang X, Kambal A, Duan E, *et al*. Pervasive transcription of a herpesvirus genome generates functionally important RNAs. MBio 2014;5:e01033-13. doi: 10.1128/mBio.01033-13.
22. Rossetto CC, Pari GS. PAN's Labyrinth: molecular biology of Kaposi's sarcoma-associated herpesvirus (KSHV) PAN RNA, a multifunctional long noncoding RNA. Viruses 2014;6:4212–4226. doi: 10.3390/v6114212.
23. Campbell M, Kim KY, Chang PC, Huerta S, Shevchenko B, Wang DH, *et al*. A lytic viral long noncoding RNA modulates the function of a latent protein. J Virol 2014;88:1843–1848. doi: 10.1128/jvi.03251-13.
24. Hennig T, Michalski M, Rutkowski AJ, Djakovic L, Whisnant AW, Friedl MS, *et al*. HSV-1-induced disruption of transcription termination resembles a cellular stress response but selectively increases chromatin accessibility downstream of genes. PLoS Pathog 2018;14:e1006954. doi: 10.1371/journal.ppat.1006954.
25. Chee MS, Bankier AT, Beck S, Bohni R, Brown CM, Cerny R, *et al*. Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD169. Curr Top Microbiol Immunol 1990;154:125–169. doi: 10.1007/978-3-642-74980-3_6.
26. Balazs Z, Tombacz D, Szucs A, Csabai Z, Megyeri K, Petrov AN, *et al*. Long-read sequencing of human cytomegalovirus transcriptome reveals RNA isoforms carrying distinct coding potentials. Sci Rep 2017;7:15989. doi: 10.1038/s41598-017-16262-z.
27. Ma Y, Wang N, Li M, Gao S, Wang L, Zheng B, *et al*. Human CMV transcripts: an overview. Future Microbiol 2012;7:577–593. doi: 10.2217/fmb.12.32.
28. Zhao F, Shen ZZ, Liu ZY, Zeng WB, Cheng S, Ma YP, *et al*. Identification and BAC construction of Han, the first characterized HCMV clinical strain in China. J Med Virol 2016;88:859–870. doi: 10.1002/jmv.24396.
29. Lu Y, Ma Y, Liu Z, Han L, Gao S, Zheng B, *et al*. A cluster of 3′ coterminal transcripts from US12-US17 locus of human cytomegalovirus. Virus Genes 2016;52:334–345. doi: 10.1007/s11262-016-1308-z.

30. Khorkova O, Myers AJ, Hsiao J, Wahlestedt C. Natural antisense transcripts. Hum Mol Genet 2014;23:R54–R63. doi: 10.1093/hmg/ddu207.
31. Pelechano V, Steinmetz LM. Gene regulation by antisense transcription. Nat Rev Genet 2013;14:880–893. doi: 10.1038/nrg3594.
32. Halley P, Kadakkuzha BM, Faghihi MA, Magistri M, Zeier Z, Khorkova O, *et al*. Regulation of the apolipoprotein gene cluster by a long noncoding RNA. Cell Rep 2014;6:222–230. doi: 10.1016/j.celrep.2013.12.015.
33. Yu W, Gius D, Onyango P, Muldoon-Jacobs K, Karp J, Feinberg AP, *et al*. Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA. Nature 2008;451:202–206. doi: 10.1038/nature06468.
34. Li Q, Su Z, Xu X, Liu G, Song X, Wang R, *et al*. AS1DHRS4, a head-to-head natural antisense transcript, silences the DHRS4 gene cluster in cis and trans. Proc Natl Acad Sci U S A 2012;109:14110–14115. doi: 10.1073/pnas.1116597109.
35. Xu Z, Wei W, Gagneur J, Clauder-Munster S, Smolik M, Huber W, *et al*. Antisense expression increases gene expression variability and locus interdependency. Mol Syst Biol 2011;7:468. doi: 10.1038/msb.2011.1.
36. Carrieri C, Cimatti L, Biagioli M, Beugnet A, Zucchelli S, Fedele S, *et al*. Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. Nature 2012;491:454–457. doi: 10.1038/nature11508.
37. Zhang G, Raghavan B, Kotur M, Cheatham J, Sedmak D, Cook C, *et al*. Antisense transcription in the human cytomegalovirus transcriptome. J Virol 2007;81:11267–11281. doi: 10.1128/JVI.00007-07.
38. Ma YP, Ruan Q, Ji YH, Wang N, Li ML, Qi Y, *et al*. Novel transcripts of human cytomegalovirus clinical strain found by cDNA library screening. Genet Mol Res 2011;10:566–575. doi: 10.4238/vol10-2gmr1059.
39. Ma Y, Wang N, Li M, Gao S, Wang L, Ji Y, *et al*. An antisense transcript in the human cytomegalovirus UL87 gene region. Virology J 2011;8:515. doi: 10.1186/1743-422X-8-515.
40. Zheng B, Li M, Gao S, Wang L, Qi Y, Ma Y, *et al*. Characterization of a novel group of antisense transcripts in human cytomegalovirus UL83 gene region. J Med Virol 2014;86:2033–2041. doi: 10.1002/jmv.23887.
41. Ma Y, Gao S, Wang L, Wang N, Li M, Zheng B, *et al*. Analysis and mapping of a 3′ coterminal transcription unit derived from human cytomegalovirus open reading frames UL30-UL32. Virol J 2013;10:65. doi: 10.1186/1743-422x-10-65.
42. Castillo AF, Fan J, Papadopoulos V, Podesta EJ. Hormone-dependent expression of a steroidogenic acute regulatory protein natural antisense transcript in MA-10 mouse tumor Leydig cells. PLoS One 2011;6:e22822. doi: 10.1371/journal.pone.0022822.
43. Ozsolak F, Kapranov P, Foissac S, Kim SW, Fishilevich E, Monaghan AP, *et al*. Comprehensive polyadenylation site maps in yeast and human reveal pervasive alternative polyadenylation. Cell 2010;143:1018–1029. doi: 10.1016/j.cell.2010.11.020.
44. He Y, Vogelstein B, Velculescu VE, Papadopoulos N, Kinzler KW. The antisense transcriptomes of human cells. Science 2008;322:1855–1857. doi: 10.1126/science.1163853.
45. Fan J, Papadopoulos V. Transcriptional regulation of translocator protein (Tspo) via a SINE B2-mediated natural antisense transcript in MA-10 Leydig cells. Biol Reprod 2012;86 147:141–115. doi: 10.1095/biolreprod.111.097535.
46. Yoshigai E, Hara T, Inaba H, Hashimoto I, Tanaka Y, Kaibori M, *et al*. Interleukin-1beta induces tumor necrosis factor-alpha secretion from rat hepatocytes. Hepatol Res 2014;44:571–583. doi: 10.1111/hepr.12157.
47. Wang X, Arai S, Song X, Reichart D, Du K, Pascual G, *et al*. Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. Nature 2008;454:126–130. doi: 10.1038/nature06992.
48. Beltran M, Puig I, Pena C, Garcia JM, Alvarez AB, Pena R, *et al*. A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition. Genes Dev 2008;22:756–769. doi: 10.1101/gad.455708.
49. Zong X, Nakagawa S, Freier SM, Fei J, Ha T, Prasanth SG, *et al*. Natural antisense RNA promotes 3′ end processing and maturation of MALAT1 lncRNA. Nucleic Acids Res 2016;44:2898–2908. doi: 10.1093/nar/gkw047.
50. Rutkowski AJ, Erhard F, L'Hernault A, Bonfert T, Schilhabel M, Crump C, *et al*. Widespread disruption of host transcription termination in HSV-1 infection. Nat Commun 2015;6:7126. doi: 10.1038/ncomms8126.
51. Lenhard B, Sandelin A, Carninci P. Metazoan promoters: emerging characteristics and insights into transcriptional regulation. Nat Rev Genet 2012;13:233–245. doi: 10.1038/nrg3163.

How to cite this article: Liu BY, Ma YP, Qi Y, Huang YJ, Liu ZY, Lu MQ, Ruan Q. Identification and characterization of a novel group of natural antisense transcripts from RNA1.2 gene locus of human cytomegalovirus. Chin Med J 2019;132:1591–1598. doi: 10.1097/CM9.0000000000000299