Arch Biol Sci. 2023;75(2):199-210 https://doi.org/10.2298/ABS230407016K

Validation and characterization of murine gammaherpesvirus 68 antisense transcripts by northern blot analysis and quantitative reverse transcription-PCR

Mehmet Kara

Bursa Uludağ University, Faculty of Arts and Sciences, Department of Molecular Biology and Genetics, Nilüfer, Bursa, Türkiye

Author's email address: mehmetkara@uludag.edu.tr

Preprint: The manuscript is available as a preprint at the following web server address: https://www.biorxiv.org/content/10.1101/2023.01.18.524592v1.full, which received the following DOI: https://doi.org/10.1101/2023.01.18.524592

Received: April 7, 202; **Revised**: April 19, 2024; **Accepted**: April 21, 2023; **Published online**: May 3, 2023

Abstract: The transcription of mammalian genomes exhibits an intriguing complexity and numerous novel RNA molecules have been identified. Viruses with large DNA genomes, especially herpesviruses, generate many different RNA species, including long non-coding RNAs (lncRNAs). Dense viral genomes can generate multigenic transcripts in addition to commonly observed antisense transcripts. It is essential to study the biological roles of these transcripts aside from the protein-coding counterparts. Multiple antisense transcripts from the open reading frame (ORF) 63-64 locus in murine gammaherpesvirus 68 (MHV68) were detected by northern blotting. Expression analysis by quantitative reverse transcription PCR (qRT-PCR) did not detect different isoforms. Several alternative splicing isoforms exist during lytic replication; however, they are not detected during latency. To identify the roles of these new transcripts, qRT-PCR may not be enough and should be supported by an alternative method such as northern blotting. A more detailed transcriptional map of the locus of interest is useful to design experimental strategies and perform functional studies, especially when working with gene-dense viral genomes.

Keywords: antisense transcription, gammaherpesvirus, long non-coding RNA (lncRNA), northern blotting

Abbreviations: murine gammaherpesvirus 68 (MHV68); Kaposi sarcoma-associated herpesvirus (KSHV); Epstein-Barr virus (EBV); herpes simplex virus 1 (HSV-1); open reading frame (ORF); long non-coding RNA (lncRNA); single molecule real time (SMRT); transcriptome resolution through integration of multiplatform data (TRIMD); 5' cap analysis of gene expression (5' CAGE); integrative genome viewer (IGV)

INTRODUCTION

In the last decade, developments and the relatively inexpensive availability of new sequencing technologies have increased. These rapid advancements caused a paradigm shift in our understanding of how much of the mammalian genome is actively transcribed [1-3]. Thorough transcriptomic data analysis has identified an array of intergenic, intronic, antisense, alternatively spliced and alternative promoter-using transcripts. As opposed to protein-coding transcripts, these long, mostly 5' capped and polyadenylated transcripts, are not associated with the cellular translational machinery, and are named long non-coding RNAs (lncRNA) [4]. Most lncRNAs reside in the nucleus, which is consistent with the idea that these RNAs remain untranslated and function through sequence-specific interactions with DNA and/or secondary motifs of proteins [5]. Therefore, the characterization of these recently identified transcripts is important for functional studies and understanding how they perform specific key roles in cellular processes.

Viruses, with their limited capacity genomes, usually contain overlapping and bidirectional genes [6]. They have evolved to use similar transcription strategies as their host to produce many different coding and non-coding RNA molecules [7,8], which presents a challenge for characterizing and studying the function of these transcripts. Even though sequencing information from long RNA molecules by PacBio's single

molecule real-time (SMRT) technology combined with short Illumina-based sequencing data have provided high resolution annotation for the transcriptional complexity of the dense viral genomes [9], it is still important to validate the transcripts experimentally. A particular locus in viruses, as shown in this study, and to some extent in mammalian genomes, has the ability to generate multiple transcripts that are not available in standard repositories such as the National Center for Biotechnology Information (NCBI).

Herpesviruses have double-stranded (ds) DNA genomes ranging from 100 to 250 kilobases in size and can code for up to 235 proteins [10]. Recently, many viruses in this family were shown to generate putative lncRNAs during the lytic and/or latent stages of the life cycle [9,11–13]. The quantitative reverse transcription PCR (qRT-PCR) assay is used to study the function of these transcripts in the laboratory [14-16]. In general, the terms qPCR, real-time PCR, RT-PCR are used interchangeably to define the method of measuring gene expression, although they may refer to slightly different techniques. For example, qPCR can be used to quantify DNA molecules without the need for reverse transcription. It involves the amplification of a specific target sequence using PCR and measuring the amount of amplicon in real time. This technique is often used to detect the presence of a specific pathogen or gene in a sample and can also be used for quantification of gene copy number. In a standard qRT-PCR assay, an RNA sample is converted to complementary DNA (cDNA) by reverse transcriptase enzyme and the cDNA is PCR-amplified using specific primers, 100-200 nucleotides apart from each other, to quantify the RNA of interest. qRT-PCR is particularly useful for studying changes in gene expression under different conditions, such as disease or drug treatments. The cDNA conversion step is very often conducted by the random priming method and, in theory, all the RNA molecules are converted to DNA. Therefore, it becomes impossible to differentiate between the distinct isomeric RNA forms and the antisense transcripts in terms of quantification of the desired gene product. The second part of this problem can be addressed by utilizing gene-specific primers for the initial cDNA conversion. However, different isomeric forms can still be captured in the qRT-PCR assay. Therefore, it is useful to have a transcriptional map of a given locus by strand-specific northern blotting so that distinct transcripts can be studied. While examining overlapping regions in the mammalian genome, a similar problem may arise. Although short-read sequencing data provide an incredible amount of information and the regions of interest are evaluated with qRT-PCR assays, the concept of alternative transcript isoforms and antisense transcription should not be ignored.

Murine gammaherpesvirus 68 (MHV68) provides a valuable model system to study human or related gammaherpesvirus pathology. The transcriptional schematic of the virus has been studied [12,17,18] and the most recent annotation of the viral transcriptome is available [13] via the utilization of a bioinformatics pipeline named (Transcriptome Resolution through Integration of Multiplatform Data TRIMD). This pipeline uses PacBio long-read sequencing, Illumina sequencing and 5' cap sequencing (CAGE) data for RNA isoform identification. TRIMD has identified ~250 novel transcripts during lytic MHV68 infection.

In this work, the ORF63-64 region in the MHV68 genome was shown to contain several readily detectable transcript isoforms by northern blotting and qRT-PCR. The antisense transcripts in this paper are partially discussed as part of the previously published northern blot protocol [19]. Herein, further examination with different probes and primer sets to map the transcriptome is provided. The aim of the present study is to analyze the different transcriptional isoforms of a putative lncRNA expressed by the ORF 63-64 locus in the viral genome and to compare the results of qRT-PCR with northern blot analyses of the same region. The time course for the transcripts and the latency expression profiles were analyzed to assess whether the expression of the different transcript variants changed.

MATERIALS AND METHODS

Cell culture and virus infection

For cell culture and infections, the mouse fibroblast cell line NIH 3T12 was used. Cells were maintained in Dulbecco's modified Eagle medium (DMEM) with 10% fetal calf serum with 100 U/mL of penicillin and streptomycin and 2 mM L-glutamine. MHV68. OR73Bla, a recombinant virus that expresses beta-lactamase as a fusion to the LANA protein encoded

by ORF73 [20], was used for the viral infections. The A20 mouse B cell lymphoma cell line and its derivative A20-HE.2.1 [21] that carries the viral genome under antibiotic selection, was used for the latency model. B cells were maintained in Roswell Park Memorial Institute (RPMI) medium under the same conditions as above in addition to 50 μM ß-mercaptoethanol. The A20-HE.2.1 cell line was kept under 300 μg/mL hygromycin B for viral genome persistence.

RNA extraction and northern blot analysis

RNA extraction and the northern blot protocol were based on the previously published protocol [19]. Briefly, fibroblasts were infected with 5 plaque-forming units of MHV68 virus per cell in a 10-cm cell culture plate. At the designated time points post-infection, the cells were washed with phosphate-buffered saline (PBS) and collected in 1 mL TRIzol. Chloroform (0.2 mL) was added to the TRIzol sample, the upper phase was collected, precipitated with isopropanol, and washed with ice-cold 75% ethanol. Total RNA was resuspended in diethylpyrocarbonate (DEPC)-treated deionized water and kept at -80°C.

For the northern blot protocol, 3-5 μg total RNA from the samples were loaded onto a 6% formaldehyde-containing 1% agarose gel in a fume hood with the RNA ladder (Millenium Marker (Ambion, Thermo Fisher Scientific, Waltham, MA, USA)). The gel was run at ~80 V for 3-4 h in 3-(N-morpholino) propanesulfonic acid (MOPS) buffer, then blotted onto a Hybond XL nylon membrane (Life Technologies, Thermo Fisher Scientific, USA) overnight with a Turbolotter kit in 20× standard saline citrate (SSC) buffer. The membrane was washed, the RNA was crosslinked to the membrane by exposure to UV light, and the membrane was stained with 0.02% methylene blue for visualization of ribosomal RNA integrity and markers. The crosslinked membrane was prehybridized at 62°C for 2-4 h in ULTRAhyb (Ambion) buffer, and the labeled probe was added for overnight incubation. The next day, the membrane was washed once with each of the following buffers: 2×SSC, 1×SSC and 0.1×SSC with 0.1% sodium dodecyl sulfate (SDS) and exposed to a film at 80°C for the appropriate time.

The radioactively labeled riboprobes were prepared with Maxiscript T7/Sp6 (Thermo Fisher). Briefly, the probe template was PCR amplified with primers containing T7 and Sp6 promoters on each side (Supplementary Table S1). The template was labeled with the kit by adding 10 μCi α-CTP (Perkin Elmer Inc. Waltham, MA, USA) to the reaction for 4 h at 37°C. The DNA template was digested with DN*ase* for 20 min, and the reaction was stopped by adding ethylenediaminetetraacetic acid (EDTA). The probe was used without purification.

Quantitative reverse transcription (qRT) PCR

Two μg of total RNA was converted to cDNA by an initial DN*ase* digestion with Turbo DNase (Ambion) following either random priming or gene-specific primers using the iScript cDNA synthesis kit (BioRad Laboratories Inc. Hercules, CA, USA). Twenty μL of the cDNA reaction was diluted 5-fold before the quantitative PCR reaction. Two μL of the cDNA sample was used for each triplicate of the PCR reaction. A 2×master mix (Fermentas Inc. Thermo Fisher Scientific) with the primers listed in Supplementary Table S1 was used for qPCR amplification in a CFX96 BioRad Real-Time PCR machine. The data were analyzed with Graphpad Prism software.

Integrative genomics viewer (IGV) data visualization

Original sequencing data for the MHV68 transcriptome is published and available at GSE117651 from NCBI. The TRIMD analysis package can be obtained https://github.com/flemingtonlab/TRIMD and the IGV is downloadable from https://software.broadinstitute.org/software/igv/. The data sets can be downloaded and analyzed free of charge. Briefly, using the MHV68 genome as reference, the '.bed' files are displayed in IGV and the 14 kb region spanning the ORF63-64 antisense transcript is visualized.

RESULTS

Visualization of the viral transcripts in the IGV

A transcriptionally complex region from the MHV68 genome was selected for further analysis. A 14 kb region harboring the antisense transcripts, which stem from

the ORF63-64 locus, is shown in Fig. 1. This region contains 17 total transcripts validated with TRIMD analysis. Most of the transcripts have scores lower than 10. The score can be roughly interpreted as lower expression and detection with the method. Some of the transcripts have splice junctions and none of the spliced transcripts have scores higher than 7. One interesting observation is that, even though ORF64 is annotated in NCBI, the RNA product of this gene was not detected by the sequencing analysis. ORF64 mRNA should be a rightward transcript as opposed to the antisense transcripts synthesized from a negative strand and going leftwards (Supplementary Fig. S1). The locations of the primers and the probes used in this study are presented in Fig. 1, and five major transcripts (with scores higher than 50) can be expected from the locus.

Northern blot analysis

RNA samples from mock or infected cells were used to detect the transcriptional complexity obtained from the region. Three sets of radiolabeled strand-specific probes were generated and used as shown in Fig. 1. According to the TRIMD analysis, the major transcripts should be about 1.9 kb, 3 kb, 3.5 kb, 4 kb and 4.6 kb in length. Probe 1 detected five major transcripts which are about 2 kb, 3 kb, 4 kb, 4.5 kb and an RNA molecule larger than 9 kb, but its size could not be determined correctly using the RNA ladder (Fig. 2). In general, the transcripts are in a very close size range for the sequencing data identified transcripts. The large RNA was not detected by TRIMD analysis most likely because a very long RNA molecule was probably not converted to cDNA as a full-length molecule, and the sequencing capacity for molecules over 10 kb is quite low. As expected, probes 2 and 3 did not detect shorter isoforms but larger bands were visible. Antisense probe 2 should detect ORF64 if expressed. However, during these conditions, the signal for ORF64 mRNA could not be detected and this is in accordance with the TRIMD analysis. The 18S and 28S rRNAs served as the loading control and as well as an estimate of the integrity of the isolated total RNA. Note that different probes work with different efficiencies for northern blotting, and the thickness of the bands should not be interpreted as a correlation of the expression levels.

Quantitative reverse transcription PCR analysis

To quantify gene expression, three sets of primers with identical viral sequences to the northern blotting probes were used to quantify antisense RNA transcription. The viral genome containing bacterial artificial chromosome (BAC) DNA was isolated from *E. coli* cells and the DNA was quantified by Nanodrop. Using the DNA copy number calculation formula (DNA amount in ng×Avogadro number/length of DNA×10^9×600), serial dilutions for 10^4 to 10^7 copy numbers were prepared and used as a template for qPCR primer efficiency. At about 10^6 copy number, the primers give Ct values lower than 20 suggesting good quality amplification (Fig. 3). Primers were also used for the cDNA samples from mock or infected cells with or without reverse transcriptase enzyme as a negative control. On average, the Ct value ranged from 18.9 for set 2 to 21.1 for set 1. Primer efficiency depends on the amplicon generation ability of the primer set and it is conceivable that each set would result in different Ct values. Usually, the best primer set is selected for further analysis. However, the best primer set selected here (Set 2), depending on the Ct value, will not detect some of the transcripts from the region (Fig. 2). Therefore, it is essential to have a transcriptome map if quantification for a given lncRNA is studied.

Reverse transcription (RT)-PCR analysis

A splice junction is visible in some of the transcripts in Fig. 1. To validate the splicing, RT PCRs are set for the mock or infected samples with set 2 forward and 3 reverse primers. The full-length amplicon of about 800 bp (should be 778 bp) is visible from the positive control BAC DNA and the infected RT+ samples. The spliced forms generate the lower molecular weight amplicon around 400 bp (should be 427 bp), which is consistent with the 351-nucleotide intron size. Even though the spliced isoforms are represented in low numbers with TRIMD analysis and are not detected by northern blotting, they possess a fainter band in the RT-PCR samples (Fig. 4A). To grasp the full complexity of the transcripts from the locus, cDNA is prepared with the gene-specific primer (Supplementary Table S1) and PCR was performed with set 1 forward and set 3 reverse primers. The resulting amplicons were run on a 10% polyacrylamide gel (Fig. 4B). This clearly indicated that multiple variants were detected

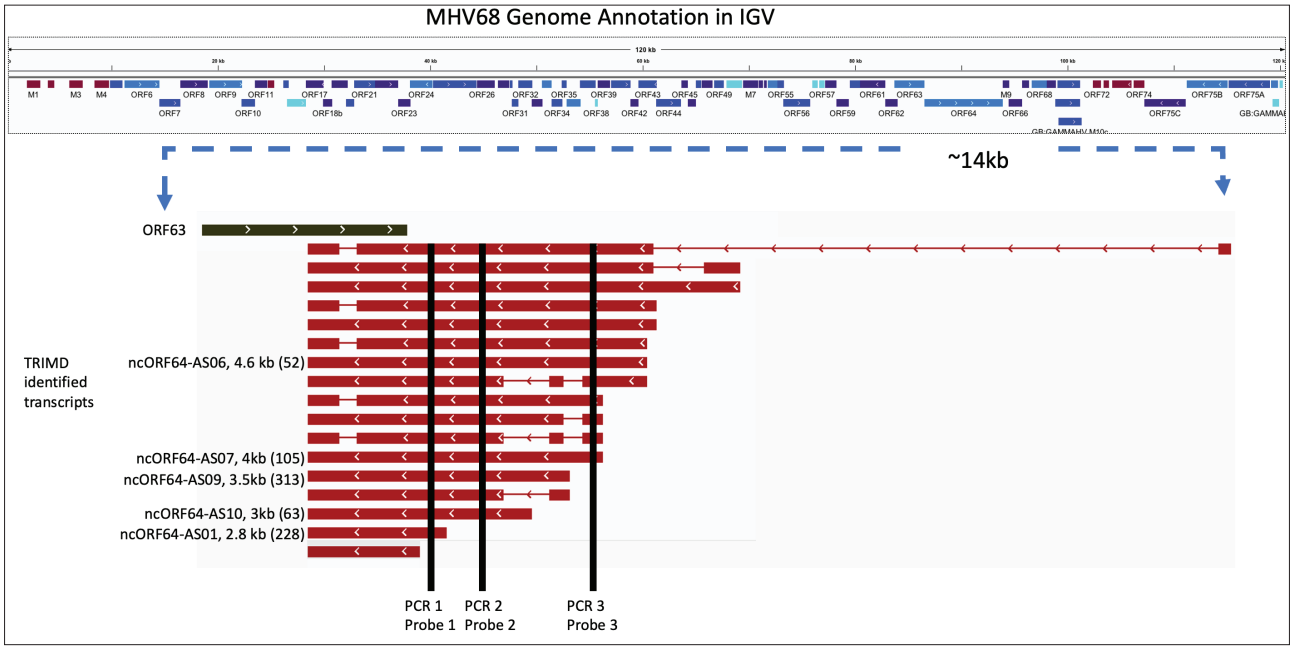


Fig. 1. Integrative Genome Viewer (IGV) visualization of ORF6364 antisense transcripts in the MHV68 genome and locations of the qRT-PCR regions and probes for northern blots. The names and lengths are given next to the transcript line for abundant transcripts. Arrows indicate the transcriptional direction. The numbers in parentheses are TRIMD scores.

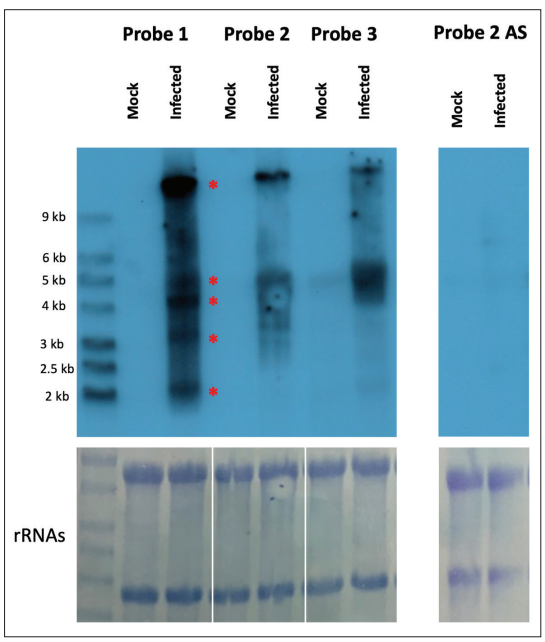


Fig. 2. Northern blot of the ORF6364 antisense (AS) transcript with four different probes. The RNA ladder is shown and marked with the corresponding sizes. Kb – kilobase. Red * indicates abundant transcripts with high TRIMD scores. Probe 2 AS is the complementary strand probe.

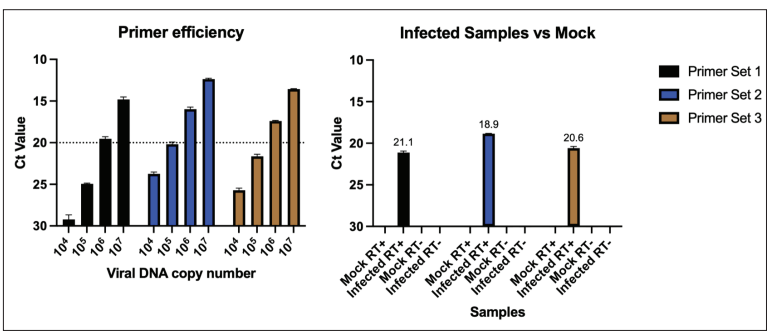


Fig. 3. qRT-PCR analysis of the ORF6364 antisense transcripts. Primer efficiency was assessed by serial dilutions of the viral genome. Mock vs wild type (WT) virus infected fibroblasts were compared with different primer sets. The cycle threshold value (ct) is given in reverse order on the y-axis.

by RT-PCR and that viruses display much more complex transcriptional activity. Additionally, some of the splice variants generated smaller RNA species which can be visualized even though they are not detected by northern blotting.

Northern blot analysis for the time course and latency

Viruses utilize different transcriptional programs during infection depending on the time and state of the infection (lytic or latency). MHV68 proceeds into the

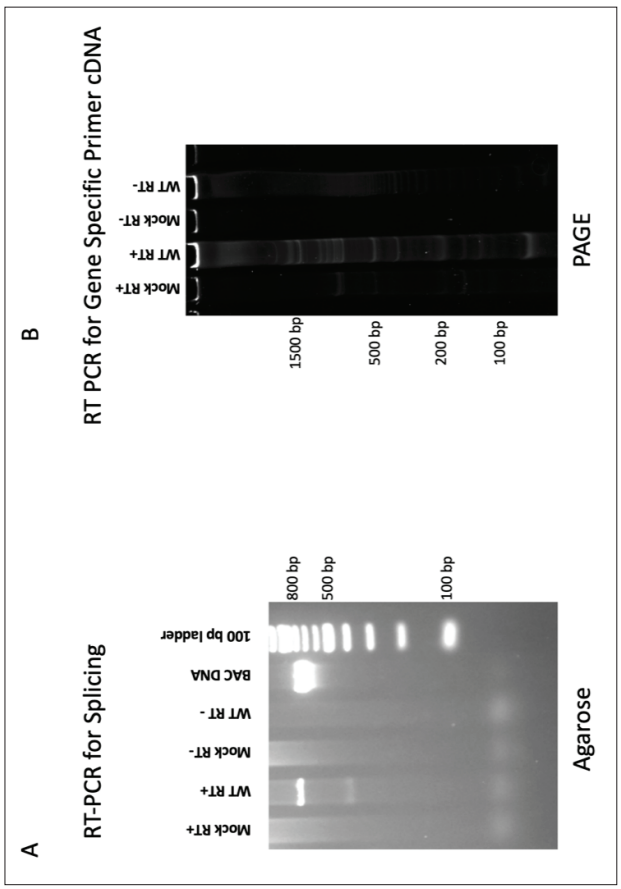


Fig. 4. Reverse Transcription (RT)-PCR for the antisense transcripts in ORF63-64 locus. **A** – Set 2 forward and set 3 reverse primers were used and the PCR products and run on a 3% agarose gel. **B** – Set 1 forward and set 3 reverse primers were used, and the PCR products were loaded onto 10% polyacrylamide gels.

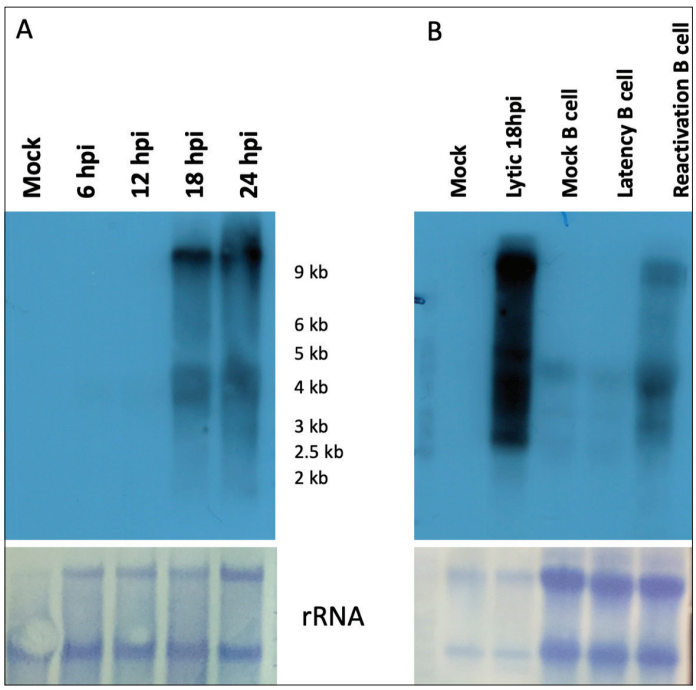


Fig. 5. Northern blot for ORF6463AS transcripts. **A** – the time course during lytic infection in 3T12 fibroblasts; **B** – latency and reactivation in B cells. Probe 1 was used in both blots.

lytic infection cycle in fibroblast cell culture models and most viral genes are expressed in this state. However, during latency, very limited gene expression (if any) was observed in B cells. To test whether there was differential expression of transcript isoforms in either viral state, lytic replication in 3T12 fibroblasts was used first. To determine whether transcripts differed during infection, samples were collected at different time points after infection and northern blotting was performed. The earliest time point that any of the transcripts were detected was at 18 h post infection and no apparent distinction regarding the isoforms was observed (Fig. 5A). In addition, the latency phase of the viral infection in B cells was tested. During latency much less viral transcripts are produced, and the amount of RNA loaded on the gel from latently infected samples increased to 15 µg while 3 µg of RNA was loaded from the lytic sample as the control (Fig. 5B). During latency, no specific transcripts were observed, however, upon induction of the reactivation with 12-O-tetradecanoylphorbol-13-acetate (TPA), a commonly used reagent to induce lytic replication for gammaherpesvirus, the transcripts became visible even though to a lesser extent when compared to the lytic infection.

DISCUSSION

The work here investigates the ability of techniques used for lncRNA studies to comprehend transcriptional complexity from a given locus in a gene-dense gammaherpesvirus genome. Genomes are pervasively transcribed, and a vast pool of RNA products is generated [1,22]. Usually, Illumina-based RNA sequencing methods in combination with long-read sequencing techniques have been developed to decipher complex transcriptomes. Multiple datasets from these platforms are integrated to identify isomeric forms of the transcripts. The analysis of alternative mRNA forms and lncRNAs in the genomes of viruses such as EBV, MHV68, KSHV and HSV-1 was performed using similar strategies [23]. Transcriptomic analyses of other virus families, such as herpesviruses,

retroviruses, baculoviruses and poxviruses, were also conducted and the key findings from these analyses demonstrate that antisense transcription, multi-gene transcripts, transcripts passing over a known polyA signal sequences called "read-through", and overlapping transcription are widely used by the viruses [24]. An important method (if not the only one) to study the novel transcripts identified by the recently developed sequencing platforms is the qRT-PCR assay, and the majority of published papers use this method to quantify lncRNAs and mRNAs [25]. This powerful method of quantification has certain drawbacks which should not be overlooked, especially while working with gene dense large DNA viruses.

To understand the function of these newly identified classes of RNA molecules, a detailed transcriptomic map of the genome proved to be useful to design primers and generate mutants. For example, the large >9 kb transcript is not found in sequencing data sets and the NCBI repository. It is an antisense transcript to several genes in the region and the function of the transcript is not known. Therefore, it is important to consider the effect of this transcript while designing experimental strategies. If one qRT-PCR assay is employed for gene expression analysis of the neighboring "sense" strand genes, the presence of this transcript may cause misinterpretation of the results. In the ORF6463 locus example here, the Ct values according to the qRT-PCR suggest that if alternative transcript isoforms are not known, primer set 2 would be preferable simply because of the lower Ct value. However, one of the abundant transcripts will not be captured by this primer set. Reverse transcription PCR suggested that many more alternatively spliced transcripts were produced. Although we did not detect a different expression pattern for these variants in the time-course experiment and latency conditions, it is still worth considering that these variant RNA molecules might have different functions depending on the context of the viral state.

The locus discussed in this work was studied in several studies [26–28]. It encodes for ORF64 protein, which is a large tegument protein with a deubiquitinating protease activity. The work presented in these studies predates the most recent transcriptome analyses that revealed antisense transcription. In some of these studies, the data from the qRT-PCR assays were utilized to assess the expression level of ORF64. This proposes a question how much of the expression can be attributed to the ORF64 mRNA and how much of the expression is because of the antisense transcripts. Similar scenarios are observed for other herpesviruses in different loci. KSHV, EBV and HSV-1 encode for large transcripts that are antisense to important latency genes in their genomes [29,30]. As an example, HSV-1 produces a non-coding transcript named the latency associated transcript (LAT) which is associated with numerous cellular pathways that block apoptosis, influence the survival of neurons, affect virulence, establishment of latency, reactivation from latency and so on [31,32]. These recently identified antisense transcripts and alternative isoforms have gone largely unnoticed in previous studies, and it is important to determine how complex transcription may affect experimental designs, such as generation of mutants and antisense oligo-targeted gene silencing.

It can be concluded that a more detailed investigation of the region of interest by alternative RNA detection methods such as northern blotting and qRT-PCR is essential, especially when working with viral transcripts. When generating viral mutants and/or inserting gene segments for recombinant protein production, it is important to understand the transcriptional complexity of the region using methods such as northern blotting.

Funding: The author received no specific funding for this work.

Acknowledgments: I would like to thank my PhD advisor Prof. Scott Tibbetts PhD and Tibbett Lab members. The laboratory work was performed at the University of Florida during my postdoctoral research at the Tibbett Lab.

Conflict of interest disclosure: The author declares no conflict of interest.

Data availability: The data underlying the findings presented in this study are available in the Supplementary Material.

REFERENCES

1. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A, Knowles DG, Lagarde J, Veeravalli L, Ruan X, Ruan Y, Lassmann T, Carninci P, Brown JB, Lipovich L, Gonzalez JM, Thomas M, Davis CA, Shiekhattar R, Gingeras TR, Hubbard TJ, Notredame C, Harrow J, Guigó R. The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. Genome Res. 2012;22(9):1775-89. https://doi.org/10.1101/gr.132159.111

2. Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, Tanzer A, Lagarde J, Lin W, Schlesinger F, Xue C, Marinov GK, Khatun J, Williams BA, Zaleski C, Rozowsky J, Röder M, Kokocinski F, Abdelhamid RF, Alioto T, Antoshechkin I, Baer MT, Bar NS, Batut P, Bell K, Bell I, Chakrabortty S, Chen X, Chrast J, Curado J, Derrien T, Drenkow J, Dumais E, Dumais J, Duttagupta R, Falconnet E, Fastuca M, Fejes-Toth K, Ferreira P, Foissac S, Fullwood MJ, Gao H, Gonzalez D, Gordon A, Gunawardena H, Howald C, Jha S, Johnson R, Kapranov P, King B, Kingswood C, Luo OJ, Park E, Persaud K, Preall JB, Ribeca P, Risk B, Robyr D, Sammeth M, Schaffer L, See L-H, Shahab A, Skancke J, Suzuki AM, Takahashi H, Tilgner H, Trout D, Walters N, Wang H, Wrobel J, Yu Y, Ruan X, Hayashizaki Y, Harrow J, Gerstein M, Hubbard T, Reymond A, Antonarakis SE, Hannon G, Giddings MC, Ruan Y, Wold B, Carninci P, Guigó R, Gingeras TR. Landscape of transcription in human cells. Nature. 2012;489(7414):101-8. https://doi.org/10.1038/nature11233
3. Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Morales DR, Thomas K, Presser A, Bernstein BE, Oudenaarden A van, Regev A, Lander ES, Rinn JL. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc Natl Acad Sci U S A. 2009;106(28):11667-72. https://doi.org/10.1073/pnas.0904715106
4. Kamel M, El-Sayed A. Utilization of herpesviridae as recombinant viral vectors in vaccine development against animal pathogens. Virus Res. 2019;270:197648. https://doi.org/10.1016/j.virusres.2019.197648
5. Guh C-Y, Hsieh Y-H, Chu H-P. Functions and properties of nuclear lncRNAs-from systematically mapping the interactomes of lncRNAs. J Biomed Sci. 2020;27(1):44. https://doi.org/10.1186/s12929-020-00640-3
6. Schlub TE, Holmes EC. Properties and abundance of overlapping genes in viruses. Virus Evol. 2020;6(1):veaa009. https://doi.org/10.1093/ve/veaa009
7. Boldogkői Z, Moldován N, Balázs Z, Snyder M, Tombácz D. Long-Read Sequencing - A Powerful Tool in Viral Transcriptome Research. Trends Microbiol. 2019;27(7):578-92 https://doi.org/10.1016/j.tim.2019.01.010
8. Tombácz D, Csabai Z, Szűcs A, Balázs Z, Moldován N, Sharon D, Snyder M, Boldogkői Z. Long-Read Isoform Sequencing Reveals a Hidden Complexity of the Transcriptional Landscape of Herpes Simplex Virus Type 1. Front Microbiol. 2017;8:1079. https://doi.org/10.3389/fmicb.2017.01079
9. O'Grady T, Wang X, Höner Zu Bentrup K, Baddoo M, Concha M, Flemington EK. Global transcript structure resolution of high gene density genomes through multi-platform data integration. Nucleic Acids Res. 2016;44(18):e145. https://doi.org/10.1093/nar/gkw629
10. Arvin A, Campadelli-Fiume G, Mocarski E, Moore PS, Roizman B, Whitley R, Yamanishi K, editors. Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis. Cambridge: Cambridge University Press; 2007. https://doi.org/10.1017/CBO9780511545313
11. Arias C, Weisburd B, Stern-Ginossar N, Mercier A, Madrid AS, Bellare P, Holdorf M, Weissman JS, Ganem D. KSHV 2.0: A Comprehensive Annotation of the Kaposi's Sarcoma-Associated Herpesvirus Genome Using Next-Generation Sequencing Reveals Novel Genomic and Functional Features. PLOS Pathog. 2014;10(1):e1003847. https://doi.org/10.1371/journal.ppat.1003847
12. Canny SP, Reese TA, Johnson LS, Zhang X, Kambal A, Duan E, Liu CY, Virgin HW. Pervasive Transcription of a Herpesvirus Genome Generates Functionally Important RNAs. mBio. 2014;5(2):e01033-13. https://doi.org/10.1128/mBio.01033-13
13. O'Grady T, Feswick A, Hoffman BA, Wang Y, Medina EM, Kara M, van Dyk LF, Flemington EK, Tibbetts SA. Genome-wide Transcript Structure Resolution Reveals Abundant Alternate Isoform Usage from Murine Gammaherpesvirus 68. Cell Rep. 2019;27(13):3988. https://doi.org/10.1016/j.celrep.2019.05.086
14. Elfman J, Li H. Detection and Measurement of Chimeric RNAs by RT-PCR. In: Li H, Elfman J, editors. Chimeric RNA: Methods and Protocols. New York, NY: Springer US; 2020. p. 83-94. (Methods in Molecular Biology). https://doi.org/10.1007/978-1-4939-9904-0_6
15. Shi X, Qin F, Li H. Confirmation of Transcriptional Read-Through Events by RT-PCR. Methods Mol Biol. 2020;2079:177-86. https://doi.org/10.1007/978-1-4939-9904-0_14
16. Su W-C, Lai MMC. Quantitative RT-PCR Analysis of Influenza Virus Endocytic Escape. Methods Mol Biol. 2018;1836:185-94. https://doi.org/10.1007/978-1-4939-8678-1_9
17. Cheng BYH, Zhi J, Santana A, Khan S, Salinas E, Forrest JC, Zheng Y, Jaggi S, Leatherwood J, Krug LT. Tiled microarray identification of novel viral transcript structures and distinct transcriptional profiles during two modes of productive murine gammaherpesvirus 68 infection. J Virol. 2012;86(8):4340-57. https://doi.org/10.1128/JVI.05892-11
18. Virgin HW, Presti RM, Li XY, Liu C, Speck SH. Three distinct regions of the murine gammaherpesvirus 68 genome are transcriptionally active in latently infected mice. J Virol. 1999;73(3):2321-32. https://doi.org/10.1128/JVI.73.3.2321-2332.1999
19. Kara M, Tibbetts SA. Empirical Validation of Overlapping Virus lncRNAs and Coding Transcripts by Northern Blot. In: Navarro A, editor. Long Non-Coding RNAs in Cancer. New York, NY: Springer US; 2021. p. 243-53. (Methods in Molecular Biology). https://doi.org/10.1007/978-1-0716-1581-2_17
20. Nealy MS, Coleman CB, Li H, Tibbetts SA. Use of a virus-encoded enzymatic marker reveals that a stable fraction of memory B cells expresses latency-associated nuclear antigen throughout chronic gammaherpesvirus infection. J Virol. 2010;84(15):7523-34. https://doi.org/10.1128/JVI.02572-09
21. Forrest JC, Speck SH. Establishment of B-Cell Lines Latently Infected with Reactivation-Competent Murine Gammaherpesvirus 68 Provides Evidence for Viral Alteration of a DNA Damage-Signaling Cascade. J Virol. 2008;82(15):7688-99. https://doi.org/10.1128/JVI.02689-07
22. Hangauer MJ, Vaughn IW, McManus MT. Pervasive Transcription of the Human Genome Produces Thousands of Previously Unidentified Long Intergenic Noncoding RNAs. PLOS Genet. 2013;9(6):e1003569. https://doi.org/10.1371/journal.pgen.1003569

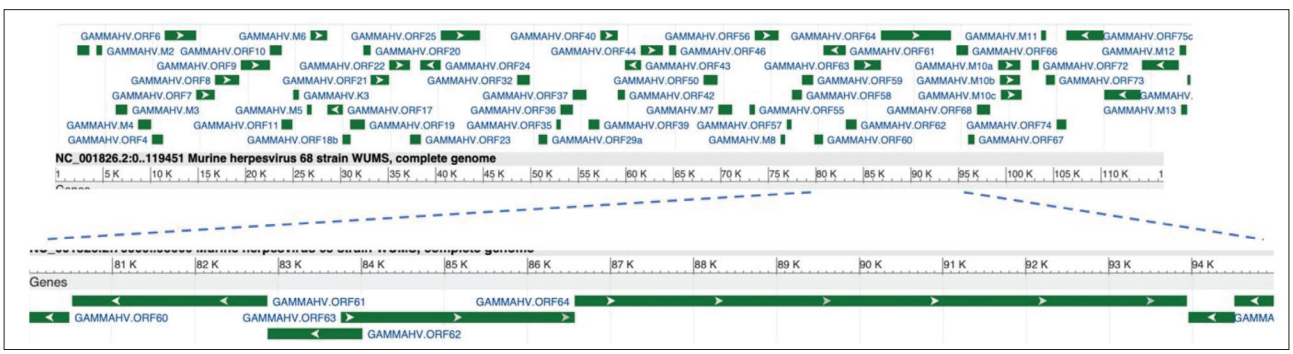
23. Tombácz D, Balázs Z, Csabai Z, Snyder M, Boldogkői Z. Long-Read Sequencing Revealed an Extensive Transcript Complexity in Herpesviruses. Front Genet. 2018;9:259. https://doi.org/10.3389/fgene.2018.00259
24. Tombácz D, Csabai Z, Oláh P, Balázs Z, Likó I, Zsigmond L, Sharon D, Snyder M, Boldogkői Z. Full-Length Isoform Sequencing Reveals Novel Transcripts and Substantial Transcriptional Overlaps in a Herpesvirus. PLoS One. 2016;11(9):e0162868. https://doi.org/10.1371/journal.pone.0162868
25. Kolenda T, Ryś M, Guglas K, Teresiak A, Bliźniak R, Mackiewicz J, Lamperska K. Quantification of long noncoding RNAs using qRT-PCR: comparison of different cDNA synthesis methods and RNA stability. Arch Med Sci. 2021;17(4):1006-15. https://doi.org/10.5114/aoms.2019.82639
26. Gredmark S, Schlieker C, Quesada V, Spooner E, Ploegh HL. A Functional Ubiquitin-Specific Protease Embedded in the Large Tegument Protein (ORF64) of Murine Gammaherpesvirus 68 Is Active during the Course of Infection. J Virol. 2007;81(19):10300-9. https://doi.org/10.1128/JVI.01149-07
27. Sun C, Schattgen SA, Pisitkun P, Jorgensen JP, Hilterbrand AT, Wang LJ, West JA, Hansen K, Horan KA, Jakobsen MR, O'Hare P, Adler H, Sun R, Ploegh HL, Damania B, Upton JW, Fitzgerald KA, Paludan SR. Evasion of innate cytosolic DNA sensing by a gammaherpesvirus facilitates establishment of latent infection. J Immunol. 2015;194(4):1819-31. https://doi.org/10.4049/jimmunol.1402495
28. Xuan Y, Gong D, Qi J, Han C, Deng H, Gao G. ZAP Inhibits Murine Gammaherpesvirus 68 ORF64 Expression and Is Antagonized by RTA. J Virol. 2013;87(5):2735-43. https://doi.org/10.1128/JVI.03015-12
29. Kara M, O'Grady T, Feldman ER, Feswick A, Wang Y, Flemington EK, Tibbetts SA. Gammaherpesvirus Readthrough Transcription Generates a Long Non-Coding RNA That Is Regulated by Antisense miRNAs and Correlates with Enhanced Lytic Replication In Vivo. Noncoding RNA. 2019;5(1):6. https://doi.org/10.3390/ncrna5010006
30. Schifano JM, Corcoran K, Kelkar H, Dittmer DP. Expression of the Antisense-to-Latency Transcript Long Noncoding RNA in Kaposi's Sarcoma-Associated Herpesvirus. J Virol. 2017;91(4):e01698-16. https://doi.org/10.1128/JVI.01698-16
31. Perng G-C, Jones C, Ciacci-Zanella J, Stone M, Henderson G, Yukht A, Slanina SM, Hofman FM, Ghiasi H, Nesburn AB, Wechsler SL. Virus-Induced Neuronal Apoptosis Blocked by the Herpes Simplex Virus Latency-Associated Transcript. Science (1979). 2000;287(5457):1500-3. https://doi.org/10.1126/science.287.5457.1500
32. Cohen JI. Herpesvirus latency. J Clin Invest. 130(7):3361-9. https://doi.org/10.1172/JCI136225.

SUPPLEMENTARY MATERIAL

Supplementary Table S1. Primer list for qRT-PCR and northern blots.

Primer	Sequence
Set 1 FWD	AAATTGGGCCTCTTATCTCCTGG
Set 1 REV	GTTTGAATAATTGGCCGTCAGC
Set 2 FWD	ATATGACCTACCAACCACCAA
Set 2 REV	GGGTGCGCCCAGATTCTTTC
Set 3 FWD	AGTCATCAACGCCTGTGTGG
Set 3 REV	TTTGAGGTCGTTTAAGTGGCTGG
Gene specific primer	GGCATTGTCCTCATCACCCAG
T7 promoter sequence	TAATACGACTCACTATA**G**GG
Sp6 promoter sequence	ATTTAGGTGACACTATA**G**AA

T7 or Sp6 promoter sequence-containing primers are ordered with the same sequence for Sets 1, 2, 3. The underlined bold **G** nucleotide in the promoter sequence is the first nucleotide of the RNA probe. FWD and REV indicate the forward and reverse primers, respectively.



Supplementary Fig. S1. Murine gammaherpesvirus 68 WUMS strain reference genome [https://www.ncbi.nlm.nih.gov/nuccore/NC_001826.2?report=graph]. The ORF6463AS-containing region of 80kb-95 kb as shown in the NCBI browser.

DATA SET

Raw Table S1. Ct values for primer efficiency.

DNA Copy Number	Primer Set 1			Primer Set 2			Primer Set 3		
10^4	29.85	28.97	28.55	23.85	23.89	23.49	25.58	26.01	25.55
10^5	24.86	25.04	24.94	20.03	20.51	20.03	21.93	21.56	21.45
10^6	19.77	19.27	19.53	16.15	15.68	16.09	17.35	17.37	17.47
10^7	15.08	14.86	14.49	12.43	NA	12.29	13.61	13.52	13.55

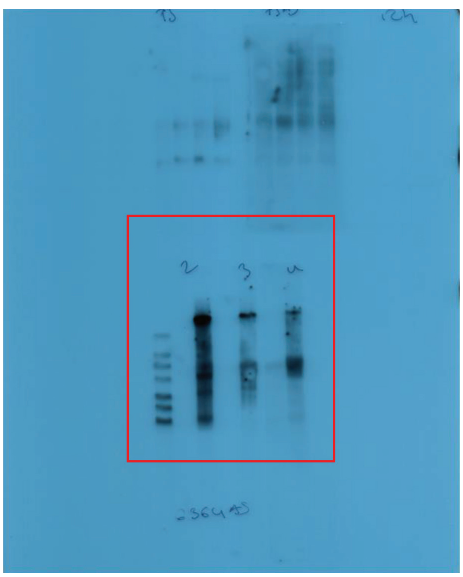
Raw Table S2. Ct Values for qPCR.

	Primer Set 1			Primer Set 2			Primer Set 3		
Mock RT+			31.05	32.37	32.05	32.07		32.95	31.50
Infected RT+	21.23	21.22	20.92	18.84	18.81	18.90	20.37	20.79	20.64
Mock RT-				31.36	33.02	33.08	33.01	32.89	33.32
Infected RT-						30.16	32.78	32.57	32.15

Raw Image S1.
Raw northern blot.
The red area was used in Fig. 1.
Loading order:
1 RNA Millennium Marker 0.5-9kb
2 Mock infected fibroblasts
3 Wild type virus-infected fibroblasts
4 Mock-infected fibroblasts
5 Wild type virus-infected fibroblasts
6 Mock infected fibroblasts
7 Wild-type virus-infected fibroblasts

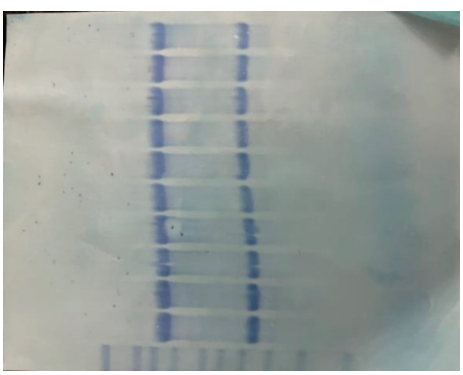
The membrane was cut into small pieces and hybridized with different probes (probe Sets 1, 2, 3), and the membrane was exposed.

Originally, more than 3 probe sets were ordered but the 1st probe was discarded; that is why the 1st two lanes are labelled as 2, and the next ones 3 and 4.



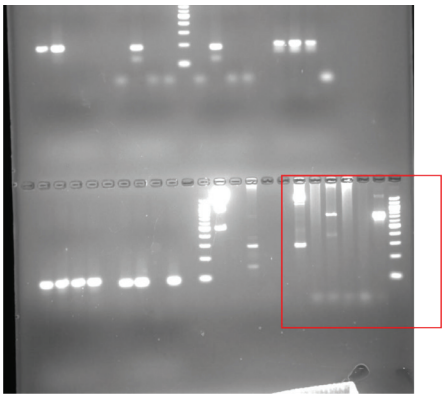
Raw Image S2.
Photograph of the membrane used for the Northern blot image above. The membrane was cut into small pieces. Only the 1st 7 lanes were used in this publication.

Methylene blue staining renders rRNA visible.
Only the 1st 6 lanes of this membrane were used for the publication.

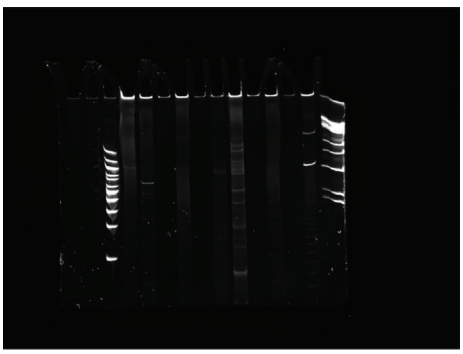


Raw Image S3.
Raw images of the gel after PCR Set 2 forward and set 3 reverse primers were used and PCR products were run on a 3% agarose gel.
Loading order:
1. PCR-positive control
2. Mock reverse transcriptase+
3. WT virus-infected reverse transcriptase+
4. Mock reverse transcriptase–
5. WT virus-infected reverse transcriptase–
6. 100 bp ladder NEB

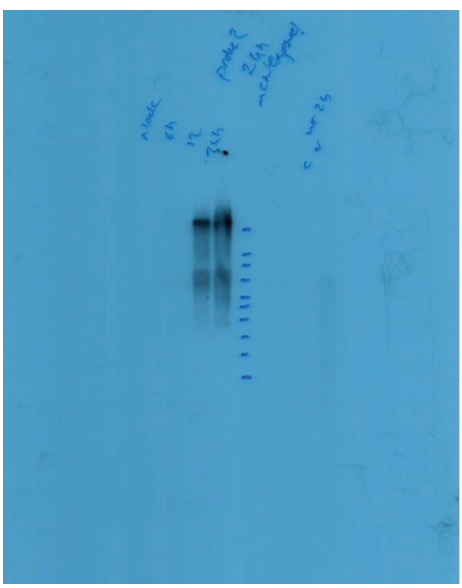
Only the red area was used for the publication.



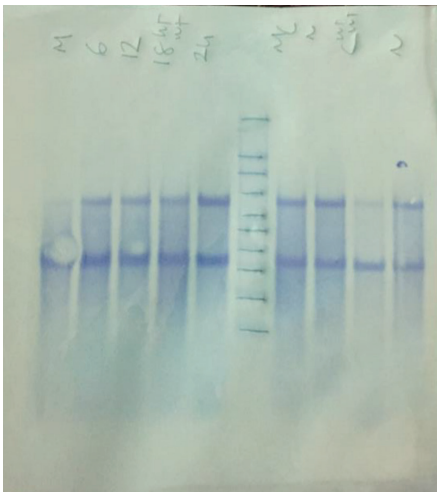
Raw Image S4.
Raw images of the gel after RT-PCR Set 1 forward and set 3 reverse primers were used and the PCR products were loaded onto a 10% polyacrylamide gel.
Loading order:
1. 100 bp NEB ladder
2-5. is the loading order below but with a different primer set;
6. Mock reverse transcriptase+
7. WT virus-infected reverse transcriptase+
8. Mock reverse transcriptase–
9. WT virus-infected reverse transcriptase–
10. PCR-positive control
11. 1 kb ladder NEB



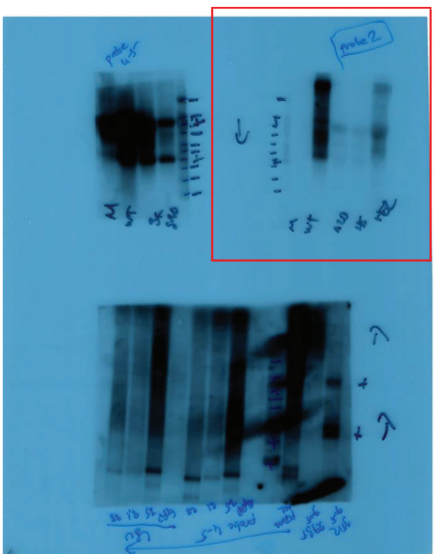
Raw image S5A.
Raw northern blot film image for Fig. 5.
Loading order:
1. Mock
2. WT-infected 6 h post infection
3. WT-infected 12 h post infection
4. WT-infected 18 h post infection
5. WT-infected 24 h post infection
6. Millenium RNA marker



Raw image S5B.
Membrane image of the blot above.
Gel loading as above.



Raw image S5C.
Raw northern blot film image for figure 5.
Only the top right corner in the red box image was used for this publication.
Gel loading:
1. Millenium RNA marker
2. Mock
3. WT-infected fibroblasts
4. A20 B cell RNA
5. Latently infected HE2.1 A20 B cell RNA
6. Latently infected HE2.1 A20 B cell RNA after reactivation for 24 h.



Raw image S5D.
The membrane image of the blot above.
Only the right side of this membrane was used for this publication.
Gel loading as above.