

# Long-read Time-course Profiling of the Host Cell Response to Herpesvirus Infection Using Nanopore and Synthetic Long-Read Transcriptome Sequencing

**Zoltán Maróti**

University of Szeged <https://orcid.org/0000-0002-0515-117X>

**Norbert Moldován**

University of Szeged <https://orcid.org/0000-0003-4138-586X>

**Gábor Torma**

University of Szeged

**Victoria A. Jefferson**

Mississippi State University

**Zsolt Csabai**

University of Szeged <https://orcid.org/0000-0003-0031-0116>

**Gábor Gulyás**

University of Szeged

**Ákos Dörmő**

University of Szeged

**Miklós Boldogkői**

University of Szeged

**Tibor Kalmár**

University of Szeged <https://orcid.org/0000-0002-0419-2009>

**Florenca Meyer**

Mississippi State University <https://orcid.org/0000-0002-6788-3886>

**Dóra Tombácz**

University of Szeged <https://orcid.org/0000-0001-5520-2978>

**Zsolt Boldogkői** (✉ [boldogkoi.zsolt@med.u-szeged.hu](mailto:boldogkoi.zsolt@med.u-szeged.hu))

University of Szeged <https://orcid.org/0000-0003-1184-7293>

---

## Research Article

**Keywords:** herpesvirus, bovine alphaherpesvirus type 1, transcriptome, virus-host interactions, RNA sequencing, nanopore sequencing, MinION, LoopSeq, long-read sequencing

**Posted Date:** February 22nd, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-264666/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

Third-generation sequencing is able to read full-length transcripts and thus to efficiently identify RNA molecules and transcript isoforms, including transcript length and splice isoforms. In this study, we report the time-course profiling of the effect of bovine alphaherpesvirus type 1 on the gene expression of bovine epithelial cells using direct cDNA sequencing carried out on MinION device of Oxford Nanopore Technologies. These investigations revealed a substantial up- and down-regulatory effect of the virus on several gene networks of the host cells, including those that are associated with antiviral response, as well as with viral transcription and translation. Additionally, we report a large number of novel bovine transcripts identified by nanopore and synthetic long-read sequencing. This study demonstrates that viral infection does not lead to a change in the average distance between promoters and transcription start sites, and between polyadenylation signals and transcription end sites. However, it causes differential expression of transcript isoforms. We could not detect an increased rate of transcriptional readthroughs as described in another alphaherpesvirus. According to our knowledge, this is the first report on the use of LoopSeq for the analysis of eukaryotic transcriptomes. This is also the first report on the application of nanopore sequencing for the kinetic characterization of cellular transcriptomes. This study also demonstrates the utility of nanopore sequencing for the characterization of dynamic changes of transcriptomes in any organisms.

## Introduction

Bovine alphaherpesvirus type 1 (BoHV-1) is a large DNA virus belonging to the *Alphaherpesvirinae* subfamily. This virus infects cattle and causes the disease commonly known as bovine respiratory disease, which leads to severe economic losses annually worldwide (van Oirschot, 1995). Like other alphaherpesviruses, such as herpes simplex virus type 1 (HSV-1), or pseudorabies virus (PRV), BoHV-1 also enters a latent state most commonly in the trigeminal ganglia following primary infection (Jones, 1998). From this state, the virus can be reactivated by various types of stress and can re-establish an acute infection (Nataraj et al., 1997).

Short-read sequencing (SRS) technology has expanded the frontiers of genomic and transcriptomic research due to its capacity to collect vast quantities of sequencing data at a relatively low cost. However, the past decade has witnessed incredible advances in long-read sequencing (LRS) technology. Besides the Pacific Biosciences and Oxford Nanopore Technologies platforms, Loop Genomics has recently also developed an LRS technique based on single molecule synthetic long-read sequencing (LoopSeq). LRS approaches present a strategy that is able to elude the limitations of SRS, including its ineffectiveness in the identification of transcript isoforms and in distinguishing overlapping RNA molecules. Recently, LRS techniques have been widely applied for the transcriptome analysis of a variety of organisms (Byrne et al., 2017; Chen et al., 2017; Tombácz et al., 2018a; Boldogkői et al., 2019; Zhao et al., 2019), including herpesviruses (Balázs et al., 2017; Moldován et al., 2018; Prazsák et al., 2018; Tombácz et al., 2018b).

These studies have uncovered a far more complex transcriptional landscape of the examined species than previously thought. Genome-wide sequencing assays have annotated the global transcriptome of BoHV-1 (Moldován et al., 2020), including microRNAs (Glazov et al., 2010). The effect of herpesvirus infection on host cell transcription using SRS (Illumina HiSeq) has been characterized by (Hu et al., 2016). In this paper, the authors described alternative splicing and polyadenylation in human skin fibroblast cells due to the infection by HSV-1.

In this work, we carried out a time-lapse assay for the examination of the effect of BoHV-1 infection on host [bovine (*Bos taurus*)] cell gene expression. The transcriptome analysis was performed using MinION sequencing from Oxford Nanopore Technologies (ONT) and Illumina-based LoopSeq from Loop Genomics.

## Materials And Methods

### Cells and viruses

Madin Darby Bovine Kidney (MDBK) cells were infected with the Cooper isolate of Bovine Herpesvirus 1.1. (GenBank Accession JX898220.1). Cells were incubated at 37°C in a humidified incubator with 5% CO<sub>2</sub>, and were cultured with Dulbecco's modified Eagle's medium (DMEM) supplemented with 5% (v/v) fetal bovine serum, 100 µg/mL streptomycin and 100 U/mL penicillin. Cells were either mock-infected or infected with Cooper isolate of BoHV-1 at a multiplicity of infection (MOI) of 5 plaque-forming units (PFU)/cell, incubated at 4°C for one hour for synchronization of infection, and then placed in an incubator at 37°C and 5% CO<sub>2</sub>. Infected cells were collected at 1, 2, 4, 6, 8, and 12 hours post infection (HPI). Each time-point and mock infection consisted of three replicates (n=3). Cells were washed with phosphate buffered saline (PBS), scraped from the culture plate and centrifuged at 300 RPM for 5 minutes at 4°C.

### RNA isolation

RNA from infected and uninfected cells (MDBK cells) was extracted using the NucleoSpin RNA kit (Machery-Nagel, Bethlehem, PA, USA), with the lysis step augmented by the addition of proteinase K (final concentration 0.37 mg/mL).

### Poly(A) RNA selection and rRNA depletion

For the analysis of the polyadenylated RNAs, the RNA fraction was enriched using Oligotex mRNA Mini Kit (Qiagen). To obtain potential non-polyadenylated transcripts, rRNA depletion was performed using Ribo-Zero Magnetic Kit H/M/R (Epicentre/Illumina).

### ONT non-amplified cDNA sequencing

Direct cDNA libraries were prepared from the mock and six BoHV-1 p.i samples in three replicates using the ONT's Direct cDNA Sequencing Kit (SQK-DCS109) according to the manufacturer's instructions. The first cDNA strand synthesis was performed using Maxima H Minus Reverse Transcriptase (Thermo Fisher

Scientific) with SSP and VN primers (supplied in the kit) and 100 ng of poly(A)+ RNA for each sample. This was followed by the removal of potential RNA contamination using RNase Cocktail Enzyme Mix (Thermo Fisher Scientific), and second strand synthesis using LongAmp Taq Master Mix (New England Biolabs). Double stranded cDNA ends were repaired using NEBNext End repair /dA-tailing Module (New England Biolabs). This was followed by ligation of sequencing adapter employing the NEB Blunt /TA Ligase Master Mix (New England Biolabs). Libraries were barcoded using Native Barcoding (12) Kit (ONT) according to the manufacturer's instructions.

### **RT with oligo(dT) primers**

50 ng of poly(A)+ RNA was reverse transcribed using SuperScript IV Reverse Transcriptase and oligo(dT) primers (supplied in the kit). The cDNA samples were subjected to PCR using KAPA HiFi DNA Polymerase (Kapa Biosystems) and Ligation Sequencing Kit Primer Mix. End repair and sequencing adapter ligation were carried out as described for the dcDNA-Seq library preparation.

### **RT with random oligonucleotide primers**

Fifty ng of ribodepleted RNA was reverse transcribed using SuperScript IV Reverse Transcriptase. Additionally, we used a custom-made primer mix composed of a random hexamer sequence and another which is complementary to the Ligation Sequencing Kit Primer (supplied in the kit). PCR, end repair and sequencing adapter ligation were identical to the oligo(dT)-primed RT library. The obtained libraries were barcoded using the 1D PCR Barcoding (96) Kit from Oxford Nanopore Technologies, according to the manufacturer's instructions. The random oligonucleotide-based cDNA sequencing was primarily used for the identification TESs of the host transcripts.

### **LoopSeq single-molecule synthetic long-read sequencing**

LoopSeq libraries were prepared from multiplexed 2 h and 12 h post-infection samples in three replicates using the LoopSeq™ Transcriptome 3 × 8-plex Kit. Phasing mRNA protocol was performed as recommended by the manufacturer.

### **Purification of libraries**

Libraries were purified after each enzymatic step using Agencourt AMPure XP magnetic beads or in the case of dRNA-Seq, RNAClean XP beads (both from Beckman Coulter). Qubit RNA BR and HS Assay Kits and Qubit DNA HS Assay Kit (Thermo Fisher Scientific) were used to measure the total RNA, poly(A)+ RNA, and cDNA concentrations, respectively.

### **Sequencers**

Sequencing of the ONT dcDNA libraries was performed on R9.4.1 SpotON Flow Cells (ONT). To avoid barcode cross-talk from later time points, mock-infected, 1h and 2h p.i. samples were sequenced

separately from other samples. The LoopSeq library was sequenced on a v2 300 flow cell on the Illumina MiSeq system.

## Pre-processing and data analysis

The MinION data was base called using the Guppy base caller v. 3.4.1. with `-qscore_filtering`. Reads with a Q-score greater than 7 were mapped to the host genome [*Bos taurus* Gene Bank accession GCF\_002263795.1] using the Minimap2 aligner (Li, 2018).

## Analysis of host cell gene expression

In order to assess the effect of the infection on host gene expression, we excluded MAPQ=0, secondary and supplementary alignments from all downstream analysis. The reads aligned to the host genome were associated to host genes according to the GCF\_002263795.1\_ARS-UCD1.2\_genomic.gff genome coordinates. Only reads matching the exon structure of the host reference genes (using a +/- 5 base pair window for matching exon start and end positions) were counted. We used edgeR\_3.24.3 (McCarthy et al., 2012) with R version 3.5.1 for differential expression (DE) analysis, and filtered out host genes with less than ten reads in any of the three biological replicates. Since we had mock, 1h, 2h, 4h, 6h, 8h, 12h measurements, we used the GLM model (`robust=True`) and the TMM normalization method in the edgeR analysis. In our model, we tested for DE against mock expression for each time point using data from three biological replicates. To detect genes with significantly changed expression levels, we applied a 0.01 false discovery rate (FDR) threshold, with p-values adjusted by the Benjamini & Hochberg procedure.

Medians of normalized pseudo-counts of DE genes were exported from edgeR (**Supplementary table S1**). Gene expression levels were normalized to maximal expression levels and were then compared to each other to reveal which genes had similar expression kinetics during viral infection. Genes were clustered by their relative expression profile using the `amap_0.8-16` R package Kmeans function with the Euclidean distance method. Based on the Calinski criteria, our dataset had an optimal cluster number of 6. Using the identified subset of genes, we performed overrepresentation analysis for each cluster using the number of expressed genes as reference via the PANTHER (version 14.1 using the 2018\_04 dataset release) (Mi et al., 2013) software tool. We summarized the results of our over-representation analysis (FDR<0.05) using the Gene Ontology (GO) biological processes and GO molecular functions annotation datasets.

Schematic representation of the workflow is shown in **Supplementary Figure S1**.

# Results

## Annotation of *Bos taurus* transcripts

In this work, we applied the following techniques for the analysis of bovine transcriptome: (1) direct cDNA sequencing (dcDNA-Seq) based on oligo(dT)-primed reverse transcription (RT), (2) amplified cDNA sequencing based on random-oligonucleotide-primed RT using nanopore sequencing on ONT MinION

platform, as well as (3) synthetic long-read sequencing (LoopSeq) on Illumina platform. All of the three techniques were used for bovine transcript annotation, whereas dcDNA-Seq was used for the time-varying analysis of the effect of BoHV-1 on host cell gene expression. For transcript detection and annotation, mapped reads were analyzed using the LoRTIA software suite developed in our laboratory (<https://github.com/zsolt-balazs/LoRTIA>).

For the annotation of introns, transcription start sites (TSSs), and transcription end sites (TESs), we set the criterion that these sequences have to be identified by the LoRTIA suit in at least two independent bovine cell samples. With this restriction, we identified altogether 11,025 TSSs, 21,317 TESs and 139,771 introns (**Supplementary Table S2**). Additionally, LoRTIA produced a total of 227,672 bovine transcripts (**Supplementary Data Item S1**). The median length of these transcripts was 1,678 nt ( $\sigma = 2,386.5$ ).

Three biological replicates were prepared for each time-point in dcDNA sequencing used for the time-lapse experiment. Seven time points post infection (p.i.) and a mock-infected sample was used in each replicate for this part of the analysis (**Supplementary Fig. S1**).

We identified consensus TATA boxes at a mean distance of 31.15 nt ( $\sigma = 2.96$ ) upstream of bovine TSSs. The polyadenylation signals (PASs) were located at a mean distance of 25.35 nt ( $\sigma = 8.26$ ) upstream of the host TES. Our data show that viral infection does not induce significant changes in the distance between promoters and TSSs as well as between PASs and TESs (**Figure 1A and 1B**). No significant modification was found in the sequence of the  $\pm 5$  nt surrounding region of the TSS and the  $\pm 50$  nt surrounding region of bovine gene TESs during the infection (**Figure 1C and 1D**).

To assess changes in splicing, and the usage of TSSs and TESs of the host cell during BoHV-1 infection, we evaluated transcripts represented by more than ten reads in the infected samples ( $n = 69,726$ ) reported by LoRTIA. We detected altogether 130 alternatively spliced transcripts (**Figure 2A**).

FOS, an immediate responder of the stress signaling pathway, is quickly degraded if its third intron is retained (Jurado et al., 2007). We detected a non-spliced variant of FOS in very low abundance and additional splice variants of the transcript lacking the above-mentioned exon, which were present starting from the first hour of the infection (**Figure 2B**). This confirms previous reports on the presence of FOS in the early stages of viral infections (Rubio and Martin-Clemente, 1999; Hu et al., 2016).

The 3'-UTRs of genes often contain miRNA targets, contributing to mRNA degradation. Thus, shorter 3'-UTR length can lead to increased transcript stability (Mayr and Bartel, 2009), whereas longer 3'-UTRs can be targeted by several miRNAs and other trans-acting elements thereby generating distinct regulation patterns (Pereira et al., 2017). We detected 72 transcripts with TESs located further downstream and 122 transcripts with TESs located more upstream compared to transcripts in mock samples. Superoxide dismutase 1 (SOD1) confers protection against oxidative damage (Miao and St. Clair, 2009), including that induced by the IFN-I signaling (Bhattacharya et al., 2015). A 3'-UTR isoform of SOD1 detected in infected cells was shorter than that of found in the mock sample (**Figure 2C**).

A previous work reported the disruption of transcript termination in the host caused by HSV-1 infection, resulting in extensive transcriptional overlaps between adjacent gene products (Rutkowski et al., 2015). According to our results, the length of polyadenylated transcripts remained constant during the infection (**Figure 3**). In order to investigate whether disruption of transcript termination also occurs in BoHV-1-infected bovine cells and results non-polyadenylated transcripts, we carried out ONT sequencing based on random oligonucleotide-primed RT, and the obtained dataset was used for the analysis of transcription activity at the intergenic regions. Despite this library yielding a comparable measure of reads mapping to *Bos taurus* (n=2,222,987), we were unable to detect any substantial amount of fragments mapping to the intergenic regions.

Using LRS, we were able to differentiate between TSS isoforms. We detected 80 transcripts with upstream and 142 with downstream TSSs.

### **Overall host cell gene expression during the 12 hours of virus infection**

This study investigated the effect of viral infection on the cultured bovine cells by a time-course transcriptome analysis using ONT LRS analysis. We carried out direct cDNA sequencing using three biological replicates in each of the six time-points (1h, 2h, 4h, 6h, 8h, 12h) and in the mock-infected sample. We identified a total of 8,342 host genes that produced more than ten transcripts in each of the three biological replicates. Applying differential expression (DE) analysis with a 0.01 false discovery rate (FDR) threshold, we identified 686 genes among the 8,342 host genes that exhibited significantly altered expression levels during the course of virus infection. Genes were clustered by their expression profile and not by their absolute expression levels. In this part of the analysis, we transformed the time series of expression levels to a relative scale representing the expression changes between sampling points. This allowed to cluster the genes by their expression profiles during the course of viral transfection instead of their absolute abundance. We identified six clusters of genes with distinctive expression profiles (Figure 4A and 4B and Supplementary Table S3). By analyzing mean expression profiles of gene clusters, we identified four groups of genes (clusters 2-5) that were constantly upregulated, a single group of genes where expression levels were steadily downregulated throughout the entire period of virus infection (cluster 6), and finally, one group that showed initial upregulation followed by downregulation (cluster 1).

We performed an over-representation analysis using the 8,342 genes as reference with the PANTHER software tool. We summarized the results of this analysis using GO (Gene Ontology) biological processes and GO molecular functions annotation datasets in **Supplementary Table S3** (an FDR<0.05 was used). Over-represented genes were categorized into six functional groups according to the GO database (**Figure 4C**) as follows: 296 genes play a role in cellular metabolism, 257 are involved in transcription and RNA decay, 242 in developmental and morphogenetic processes, 187 in immune response and host defense, 161 in translation and protein folding, whereas 61 genes are specifically associated with in viral transcription related processes.

Genes of the first cluster (n=53) had medium expression preceding the infection (which was transiently slightly upregulated at the 1h, and 2h p.i. time points) followed by downregulation at later measurements.

Genes in this cluster were over-represented in pathways controlling a wide variety of developmental and morphogenetic processes. Several genes coding transcription regulatory proteins present in this cluster show diminishing expression throughout the infection. Genes involved in the cytokine regulation of the immune response and inflammatory processes are also affected. The second and third cluster of genes (n=64, n=82 respectively) had medium expression preceding the infection that rose at each consecutive time points. The genes of these clusters were over-represented in functions and molecular processes that can be associated with viral gene expression and the virion assembly. An upregulation of genes involved in transcriptional and translational processes, as well as RNA decay was also observed. RNA decay can be an immediate response of the host cell to counteract the accumulation of viral transcripts, or it may be an effort of the virus to eliminate competing host mRNAs in order to facilitate the translation of viral transcripts (Smiley, 2004; Moon and Wilusz, 2013). Some of the over represented genes in these cluster are the members of GO molecular functions that have overlapping sets of genes. For example, the 12 genes (RPS26, RPL5, RPL30, RPS29, RPL31, RPS6, RPL36, RPL37, RPL8, RPS10, RPS21, RPL19) that were significantly upregulated during infection are the members of both the “viral transcription” and the “SRP-dependent co-translational protein targeting to membrane” pathways. Many of these genes are also members of the “nuclear-transcribed mRNA catabolic process, nonsense-mediated decay” pathway.

Genes in the fourth cluster (n=64) had low relative expression preceding the infection. These genes were upregulated following a sigmoid curve during the infection. Genes in the fourth cluster were not significantly over-represented in any of the GO molecular functions or GO biological processes. The fifth cluster of genes (n=88) had zero or negligible expression preceding infection but showed an exponential increase in expression during the course of infection. The over-represented genes of this cluster were associated with anti-viral cellular and defense responses such as the type I interferon signaling pathways. The sixth cluster of genes (n=335) consisted of a huge variety of host genes with high expression preceding viral infection that showed sharp downregulation during the infection. These genes were over-represented in pathways associated with protein folding, cell cycle regulation and mitochondrial processes including aerobic respiration.

### **Key response host genes**

We performed DE analysis (with FDR=0.01) on the Mock and 1-hour expression values to describe the immediate response of host cells. We identified 6 bovine genes that were significantly down regulated and 19 genes that were significantly up regulated in the three biological replicates (Supplementary table S1). Over expression analysis revealed no significant association to either of the GO biological processes or GO molecular functions using the subset of up and down regulated genes or the whole set of genes. However, STRING association analysis revealed 4 networks between these genes. The first gene network (GADD45B, GADD45A, DDIT3, ATF3, IFRD1, CARM1, SQSTM1) contains genes that are associated to host DNA damage response, transcription regulation. Furthermore, one gene plays a role in selective autophagy. The second network consists of two of the interferon gamma stimulated genes; IRF9, a transcription factor that plays an essential role in anti-viral activity and MT2A, a metallothionein protein. The third network consists of two genes (SRSF5 and HNRNPDL) associated with pre-mRNA processing,

transport and splicing regulation. The cytokine IL11 which regulates hematopoietic cells was part of the fourth network. We found IL11 to be down-regulated. In contrast, CXCL5, a gene associated with neutrophil activation and also present in network 4, was up-regulated following virus infection. The remaining four (LASP1, HDAC7, SLC44A2 HSPG2) out of 6 down-regulated genes and eight (ID2, HMG3, TMEM190, TSC22D1, PRKAR2B, LOC100847759, LOC100847143, LOC100174924) out of 19 up-regulated genes include signaling, transcriptional regulator, developmental genes.

## Discussion

High-throughput long-read sequencing approaches are able to read full-length transcripts, and therefore allow a more comprehensive annotation of RNA molecules. LRS-based studies led to the discovery that the transcriptomes are much more complex than previously thought.

In this study, we annotated a large number of bovine transcripts and analyzed the effect of viral infection upon host gene expression. We found no significant change in the usage of promoters or PASs of the host genes. However, we observed an altered usage of transcript length and splice isoforms of the host RNA molecules. This indicates a modulation of cellular mRNA turnover. The analysis of TSS isoforms suggests that viral infection may have an effect on host mRNA translation, potentially through uORFs (Kronstad et al., 2013), or through other cis-acting elements, such as miRNA binding sites of 5'-UTRs. However, downstream TSSs can also result in truncated in-frame ORFs, which might code for N-terminally truncated polypeptides (Crofts et al., 1998; Tombácz et al., 2019). Unlike in HSV-1-infected cells (Rutkowski et al., 2015), we found no increase in the extent of transcriptional readthroughs in BoHV-1-infected-cells.

Based on the alteration of expression kinetics, we detected six distinct gene clusters that had significantly changed expression during the course of virus infection. Based on the overrepresentation analysis of these clusters we distinguished three functional groups. Genes involved in basic cell functions, including morphogenesis, cell cycle regulation, signaling, catabolic pathways and aerobic respiration, are generally downregulated during viral infection. On the other hand, we observed a considerable upregulation of genes involved in antiviral response. Additionally, genes playing a role in transcription, RNA decay, translation and protein folding were also upregulated. Our analysis shows that most of these genes are associated to distinct molecular functions and biological processes indicating general response to virus infection. However, the rest of the unassociated genes could also be associated with either susceptibility to or defense against viral infection. We also identified a small set of immediate response genes that exhibited significantly altered expression 1 hour after viral infection.

Altogether, our data provides valuable resources for future functional studies and for understanding how the virus can overcome host defense mechanisms. Furthermore, these results may be helpful for the development of novel antiviral therapies.

## Declarations

## Funding

This study was supported by OTKA K 128247 granted to ZB, by the OTKA FK 128252 and by the Lendület (Momentum) I Program of the Hungarian Academy of Sciences (LP-2020/8) granted to DT. The funding body had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

## Authors contribution

Z.M., N.M., G.T., M.B., GG, and TK. carried out bioinformatic analysis of the viral transcripts. Z.C., D.T. and ÁD prepared ONT MinION libraries, carried out ONT sequencing and participated in the analysis. D.T. generated Loop-seq libraries. Z.C. and V.J. isolated RNA, maintained cell cultures, propagated viruses and participated in nucleic acid isolation and library preparation. F.M. and DT drafted the manuscript. Z.B. integrated the data, performed the analysis, and wrote the final version of the manuscript. All authors reviewed and approved the final version of the manuscript.

## Data availability statement

The sequencing datasets generated during this study are available at the European Nucleotide Archive's SRA database under the accession PRJEB33511 (<https://www.ebi.ac.uk/ena/browser/view/PRJEB33511>).

## Code availability

The LoRTIA software suite is available on GitHub: <https://github.com/zsolt-balazs/LoRTIA>.

## Conflict of interest

The authors declare no conflict of interest.

## Competing interests

The authors declare no competing interests.

## References

- Balázs, Z., Tombácz, D., Szucs, A., Csabai, Z., Megyeri, K., Petrov, A. N., et al. (2017). Long-Read Sequencing of Human Cytomegalovirus Transcriptome Reveals RNA Isoforms Carrying Distinct Coding Potentials. *Sci. Rep.* 7. doi:10.1038/s41598-017-16262-z.
- Bhattacharya, A., Hegazy, A. N., Deigendesch, N., Kosack, L., Cupovic, J., Kandasamy, R. K., et al. (2015). Superoxide Dismutase 1 Protects Hepatocytes from Type I Interferon-Driven Oxidative Damage. *Immunity* 43, 974–986. doi:10.1016/j.immuni.2015.10.013.

- Boldogkői, Z., Moldován, N., Balázs, Z., Snyder, M., and Tombácz, D. (2019). Long-Read Sequencing – A Powerful Tool in Viral Transcriptome Research. *Trends Microbiol.* 27, 578–592. doi:10.1016/j.tim.2019.01.010.
- Byrne, A., Beaudin, A. E., Olsen, H. E., Jain, M., Cole, C., Palmer, T., et al. (2017). Nanopore long-read RNAseq reveals widespread transcriptional variation among the surface receptors of individual B cells. *Nat. Commun.* 8, 16027. doi:10.1038/ncomms16027.
- Chen, S.-Y., Deng, F., Jia, X., Li, C., and Lai, S.-J. (2017). A transcriptome atlas of rabbit revealed by PacBio single-molecule long-read sequencing. *Sci. Rep.* 7, 7648. doi:10.1038/s41598-017-08138-z.
- Crofts, L. A., Hancock, M. S., Morrison, N. A., and Eisman, J. A. (1998). Multiple promoters direct the tissue-specific expression of novel N-terminal variant human vitamin D receptor gene transcripts. *Proc. Natl. Acad. Sci. U. S. A.* 95, 10529–10534. doi:10.1073/pnas.95.18.10529.
- Glazov, E. A., Horwood, P. F., Assavalapsakul, W., Kongsuwan, K., Mitchell, R. W., Mitter, N., et al. (2010). Characterization of microRNAs encoded by the bovine herpesvirus 1 genome. *J. Gen. Virol.* 91, 32–41. doi:10.1099/vir.0.014290-0.
- Hu, B., Li, X., Huo, Y., Yu, Y., Zhang, Q., Chen, G., et al. (2016). Cellular responses to HSV-1 infection are linked to specific types of alterations in the host transcriptome. *Sci. Rep.* 6, 1–14. doi:10.1038/srep28075.
- Jones, C. (1998). Alphaherpesvirus latency: its role in disease and survival of the virus in nature. *Adv. Virus Res.* 51, 81–133. doi:10.1016/s0065-3527(08)60784-8.
- Jurado, J., Fuentes-Almagro, C. A., Prieto-Álamo, M. J., and Pueyo, C. (2007). Alternative splicing of c-fos pre-mRNA: Contribution of the rates of synthesis and degradation to the copy number of each transcript isoform and detection of a truncated c-Fos immunoreactive species. *BMC Mol. Biol.* 8, 1–13. doi:10.1186/1471-2199-8-83.
- Kronstad, L. M., Brulois, K. F., Jung, J. U., and Glaunsinger, B. a (2013). Dual short upstream open reading frames control translation of a herpesviral polycistronic mRNA. *PLoS Pathog.* 9, e1003156. doi:10.1371/journal.ppat.1003156.
- Li, H. (2018). Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics* 34, 3094–3100. doi:10.1093/bioinformatics/bty191.
- Mayr, C., and Bartel, D. P. (2009). Widespread Shortening of 3'UTRs by Alternative Cleavage and Polyadenylation Activates Oncogenes in Cancer Cells. *Cell* 138, 673–684. doi:10.1016/j.cell.2009.06.016.
- McCarthy, D. J., Chen, Y., and Smyth, G. K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* 40, 4288–4297. doi:10.1093/nar/gks042.

- Mi, H., Muruganujan, A., Casagrande, J. T., and Thomas, P. D. (2013). Large-scale gene function analysis with the panther classification system. *Nat. Protoc.* 8, 1551–1566. doi:10.1038/nprot.2013.092.
- Miao, L., and St. Clair, D. K. (2009). Regulation of superoxide dismutase genes: Implications in disease. *Free Radic. Biol. Med.* 47, 344–356. doi:10.1016/j.freeradbiomed.2009.05.018.
- Moldován, N., Tombácz, D., Szucs, A., Csabai, Z., Snyder, M., and Boldogkoi, Z. (2018). Multi-platform sequencing approach reveals a novel transcriptome profile in pseudorabies virus. *Front. Microbiol.* 8, 1–13. doi:10.3389/fmicb.2017.02708.
- Moldován, N., Torma, G., Gulyás, G., Hornyák, Á., Zádori, Z., Jefferson, V. A., et al. (2020). Time-course profiling of bovine alphaherpesvirus 1.1 transcriptome using multiplatform sequencing. *Sci. Rep.* 10, 1–14. doi:10.1038/s41598-020-77520-1.
- Moon, S. L., and Wilusz, J. (2013). Cytoplasmic Viruses: Rage against the (Cellular RNA Decay) Machine. *PLoS Pathog.* 9, 1–3. doi:10.1371/journal.ppat.1003762.
- Nataraj, C., Eidmann, S., Hariharan, M. J., Sur, J. H., Perry, G. A., and Srikumaran, S. (1997). Bovine herpesvirus 1 downregulates the expression of bovine MHC class I molecules. *Viral Immunol.* 10, 21–34. doi:10.1089/vim.1997.10.21.
- Pereira, L. A., Munita, R., González, M. P., and Andrés, M. E. (2017). Long 3'UTR of Nurr1 mRNAs is targeted by miRNAs in mesencephalic dopamine neurons. *PLoS One* 12, 1–15. doi:10.1371/journal.pone.0188177.
- Prazsák, I., Moldován, N., Balázs, Z., Tombácz, D., Megyeri, K., Szűcs, A., et al. (2018). Long-read sequencing uncovers a complex transcriptome topology in varicella zoster virus. *BMC Genomics* 19, 873. doi:10.1186/s12864-018-5267-8.
- Rubio, N., and Martin-Clemente, B. (1999). Theiler's murine encephalomyelitis virus infection induces early expression of c-fos in astrocytes. *Virology* 258, 21–29. doi:10.1006/viro.1999.9684.
- Rutkowski, A. J., Erhard, F., L'Hernault, A., Bonfert, T., Schilhabel, M., Crump, C., et al. (2015). Widespread disruption of host transcription termination in HSV-1 infection. *Nat. Commun.* 6. doi:10.1038/ncomms8126.
- Smiley, J. R. (2004). Herpes Simplex Virus Virion Host Shutoff Protein: Immune Evasion Mediated by a Viral RNase? *J. Virol.* 78, 1063–1068. doi:10.1128/jvi.78.3.1063-1068.2004.
- Tombácz, D., Moldován, N., Balázs, Z., Gulyás, G., Csabai, Z., Boldogkői, M., et al. (2019). Multiple Long-Read Sequencing Survey of Herpes Simplex Virus Dynamic Transcriptome. *Front. Genet.* 10, 1–20. doi:10.3389/fgene.2019.00834.

Tombácz, D., Prazsák, I., Szűcs, A., Dénes, B., Snyder, M., and Boldogkői, Z. (2018a). Dynamic transcriptome profiling dataset of vaccinia virus obtained from long-read sequencing techniques. *Gigascience* 7. doi:10.1093/gigascience/giy139.

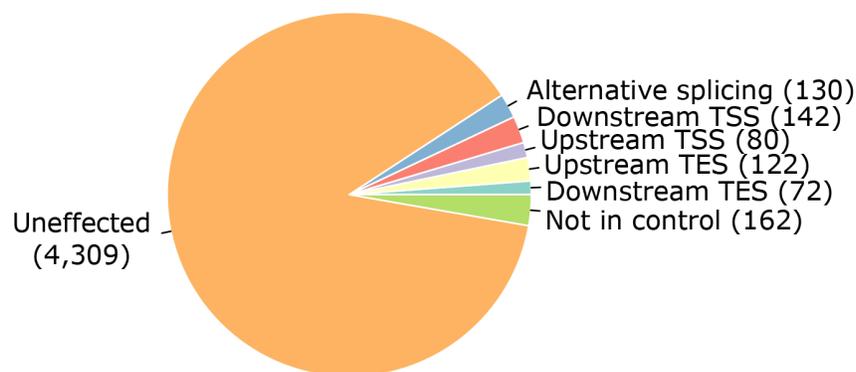
Tombácz, D., Sharon, D., Szűcs, A., Moldován, N., Snyder, M., and Boldogkői, Z. (2018b). Transcriptome-wide survey of pseudorabies virus using next- and third-generation sequencing platforms. *Sci. Data* 5, 180119. doi:10.1038/sdata.2018.119.

van Oirschot, J. T. (1995). Bovine herpesvirus 1 in semen of bulls and the risk of transmission: a brief review. *Vet. Q.* 17, 29–33. doi:10.1080/01652176.1995.9694526.

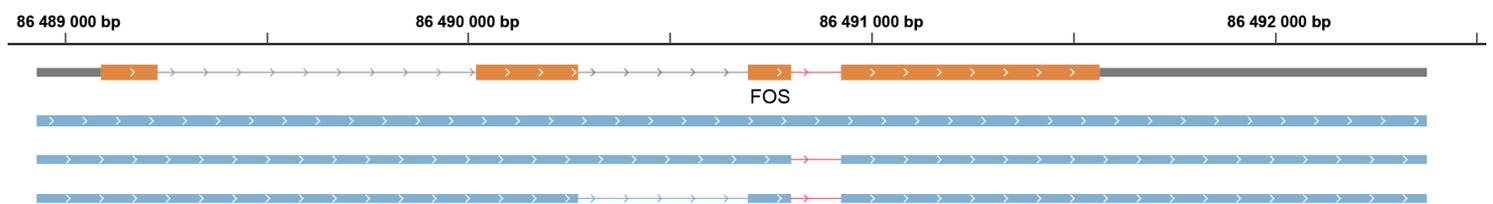
Zhao, L., Zhang, H., Kohnen, M. V., Prasad, K. V. S. K., Gu, L., and Reddy, A. S. N. (2019). Analysis of Transcriptome and Epitranscriptome in Plants Using PacBio Iso-Seq and Nanopore-Based Direct RNA Sequencing. *Front. Genet.* 10, 253. doi:10.3389/fgene.2019.00253.

## Figures

A



B



C

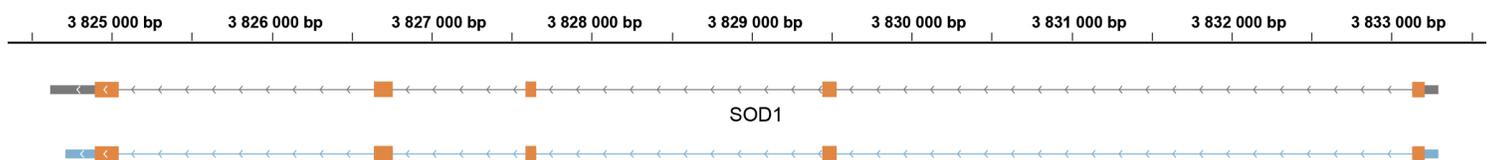
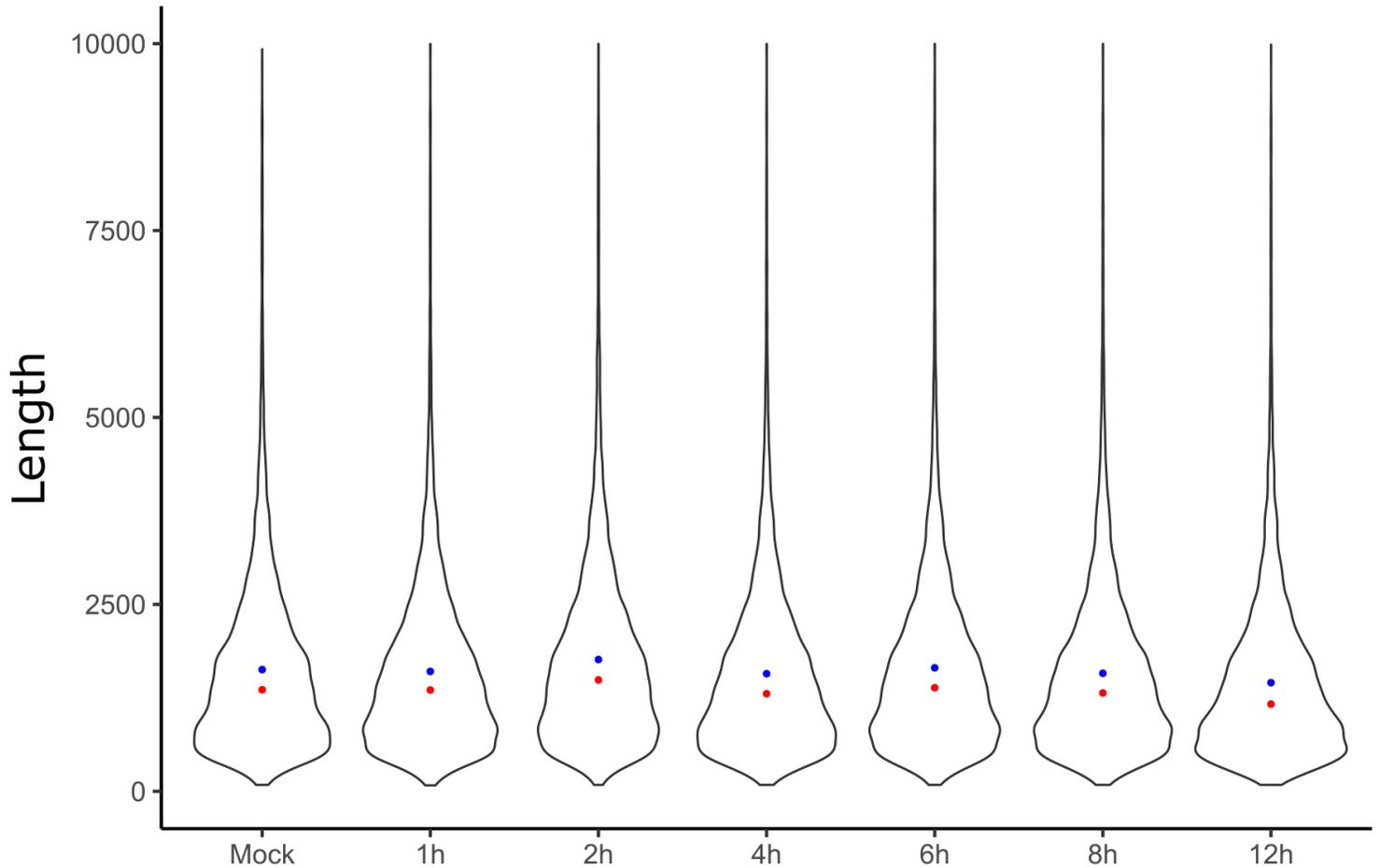


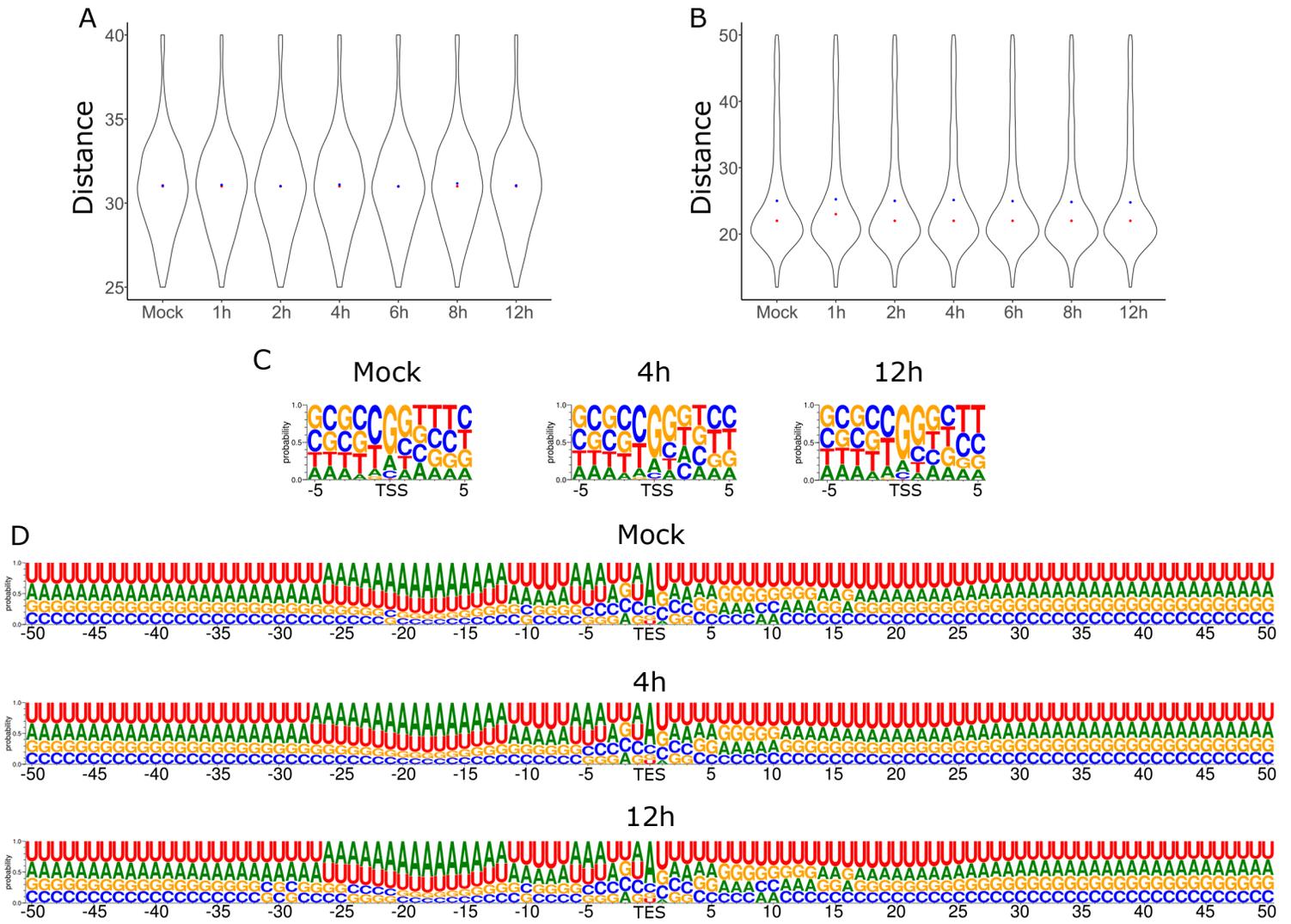
Figure 1

The effect of viral infection on transcript start and end sites. (A) The distance of TATA boxes from TSSs and (B) the distance of polyadenylation signals from TESs in base pairs. The sequence motif of the (C) TSSs and the (D) TESs in the mock infected cells, 4h and 12h following infection



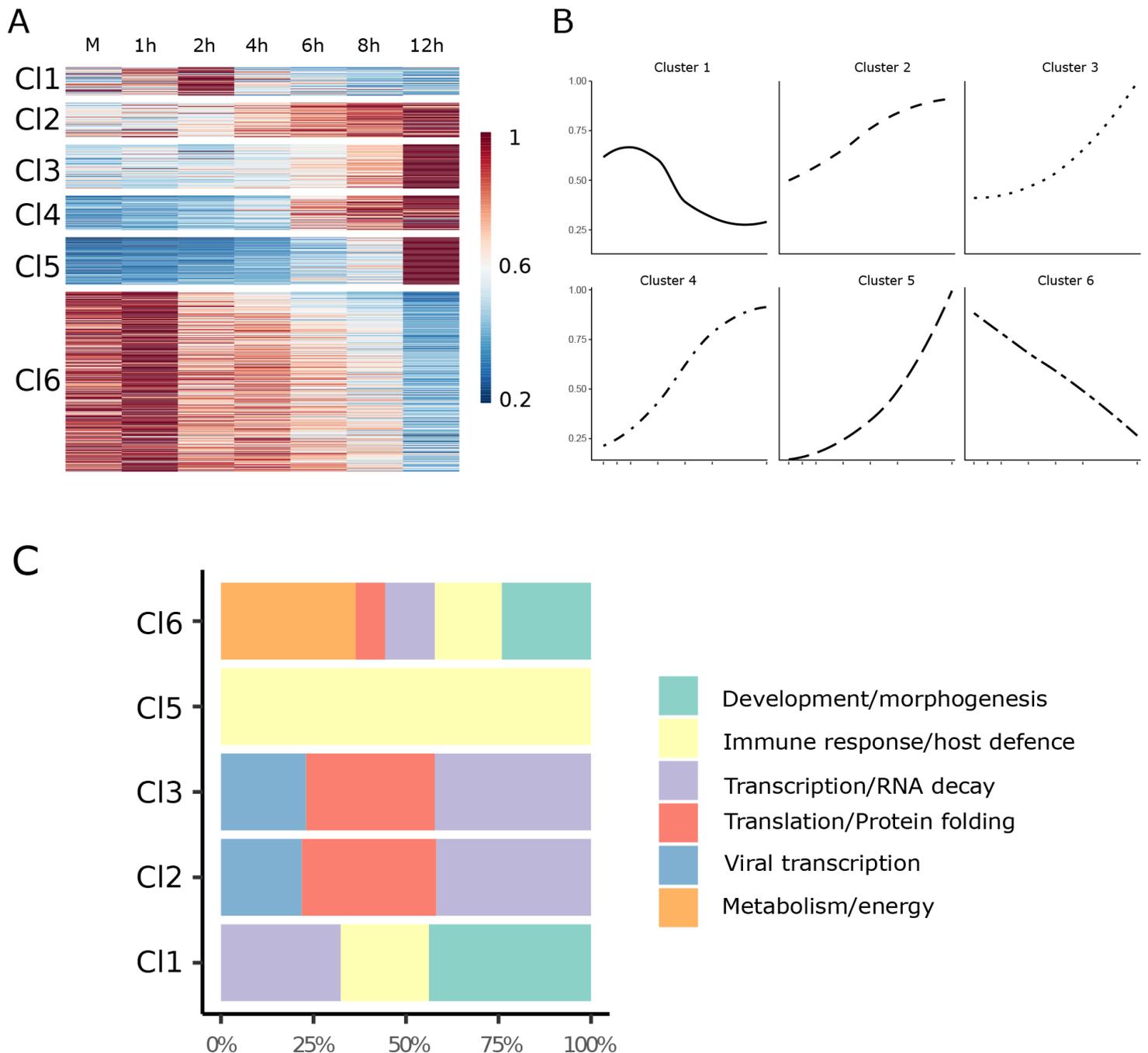
**Figure 2**

The effect of viral infection on the transcript isoforms of host cells. (A) A total of 546 transcript isoforms were detected with at least ten reads in the datasets of transcripts derived from infected bovine cells compared to the transcripts present in the mock dataset. (B) The blue rectangles represent a novel non-spliced and two new splice variants of the FOS transcript. Introns are represented by lines between rectangles. The intron which is responsible for the increased degradation rate of FOS is shown in red. Orange rectangles represent the ORF. (C) The blue rectangles represent the novel TES isoform of SOD1, the lines represent introns, and orange rectangles illustrate the ORF. A SOD1 isoform with an upstream TES detected during the infection but not in the mock samples.



**Figure 3**

The length of polyadenylated bovine transcripts. Blue dots represent the median while red dots represent the mean of datasets.



**Figure 4**

Differential expression analysis of host genes. (A) Relative gene expression represented on a heat map. Six distinct kinetic clusters (CI) were identified among host genes. The scale represents the mean relative gene expression. (B) The average change in normalized relative expression of genes present in the six kinetic clusters during infection. (C) The distribution of genes associated with the six functional categories according to gene ontology (GO) between the kinetic clusters.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- DataSheet1.zip
- DataSheet2.pdf
- Table1.xlsx
- Table2.xlsx
- Table3.xlsx