Supplemental Information

Genome-wide Transcript Structure Resolution

Reveals Abundant Alternate Isoform

Usage from Murine Gammaherpesvirus 68

Tina O'Grady, April Feswick, Brett A. Hoffman, Yiping Wang, Eva M. Medina, Mehmet Kara, Linda F. van Dyk, Erik K. Flemington, and Scott A. Tibbetts
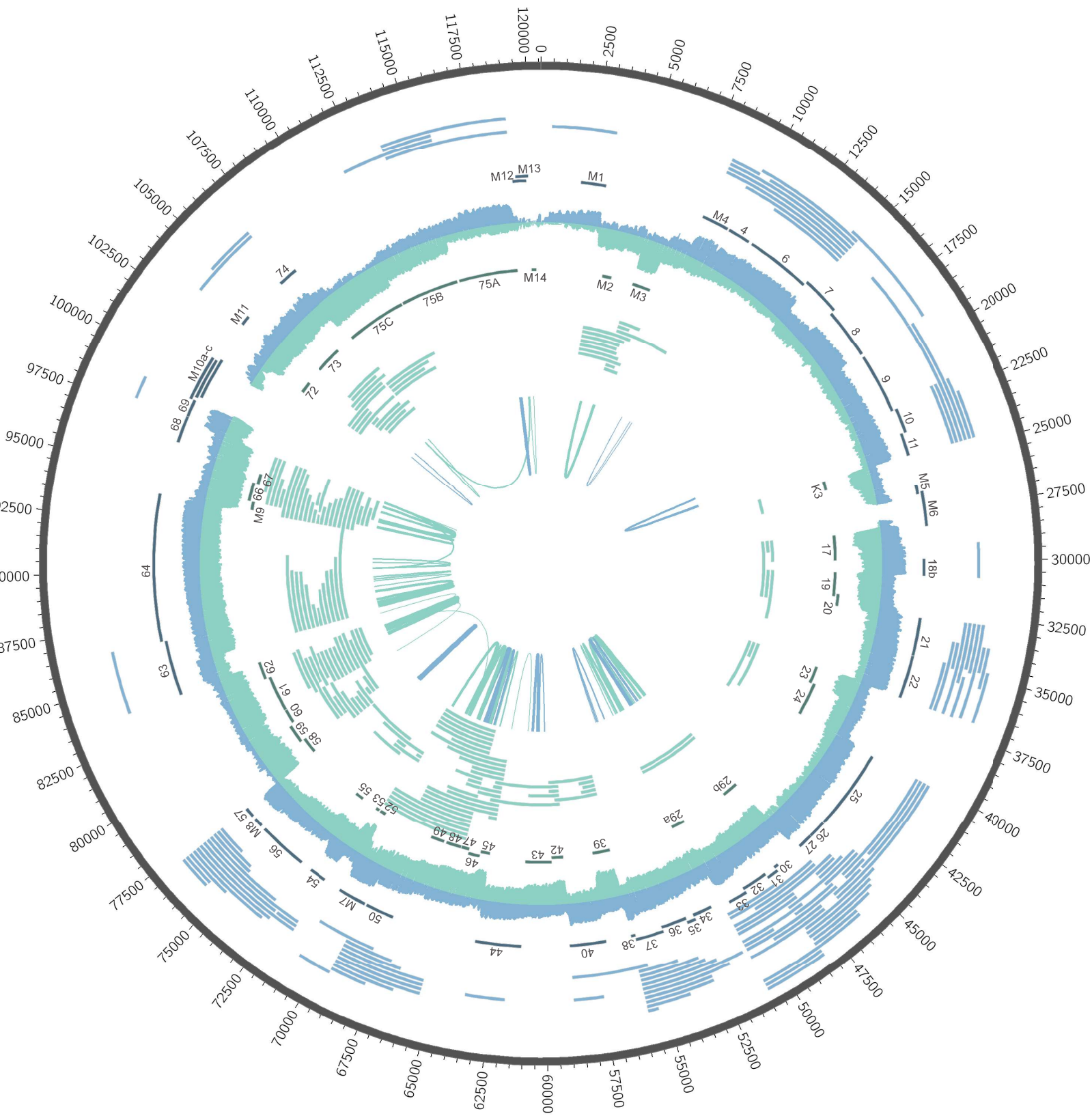
Figure S1 (related to Figure 2). Overview of MHV68 transcription. From outside to inside, tracks represent: dsDNA genome (with nucleotide position annotation), plus strand TRIMD transcripts (shown unspliced), plus strand GenBank ORFs, plus strand RNA-Seq coverage, minus strand RNA-Seq coverage, minus strand GenBank ORFs, minus strand TRIMD transcripts (shown unspliced), TRIMD splice junctions. Plus strand features are displayed in blue, minus strand features are displayed in green. Image was generated with Circos (Krzywinski, M. et al. Circos: an Information Aesthetic for Comparative Genomics. Genome Res (2009) 19:1639-1645).
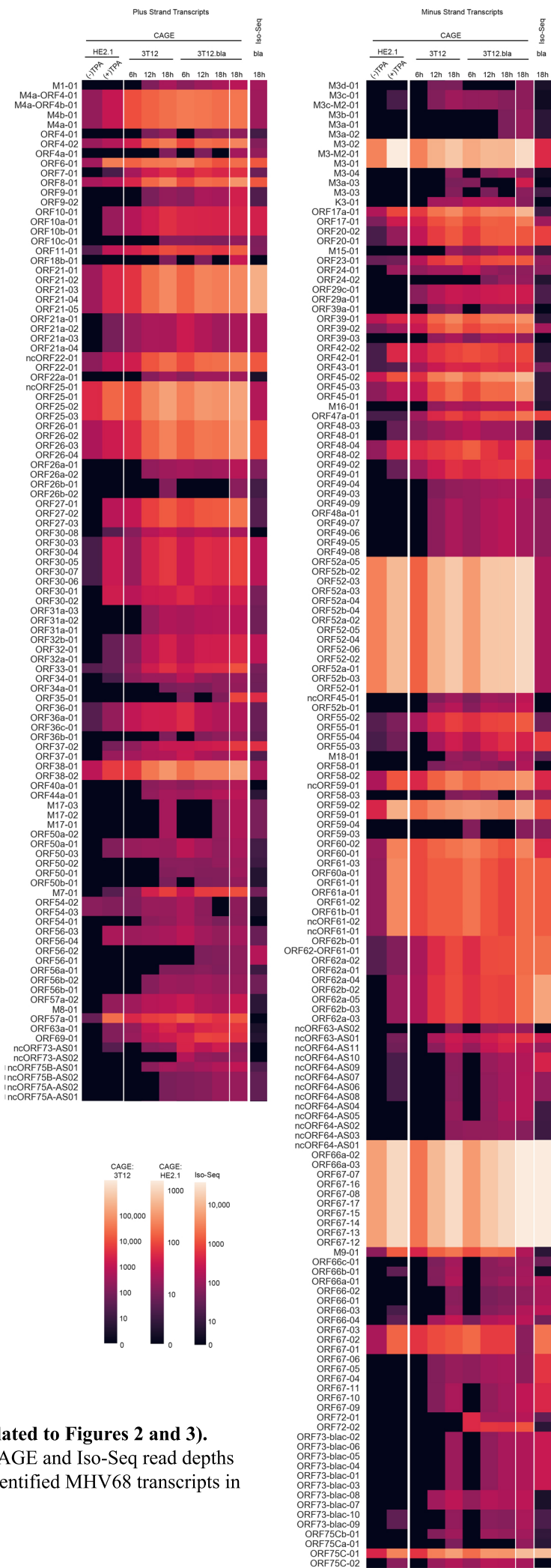
Figure S2 (related to Figures 2 and 3).
Heatmap of CAGE and Iso-Seq read depths
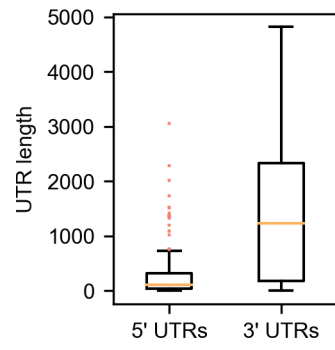for TRIMD-identified MHV68 transcripts in
all datasets.

Figure S3 (related to Figure 4). Lengths of 5' and 3' UTRs of TRIMD-identified MHV68 transcripts.
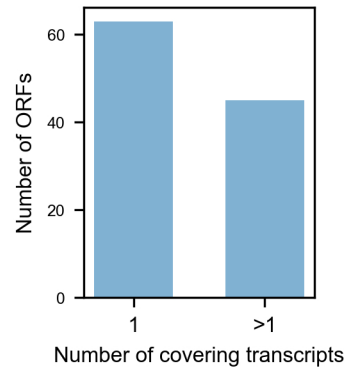
Figure S4 (related to Figure 4). Number of TRIMD-identified transcripts containing each identified ORF.
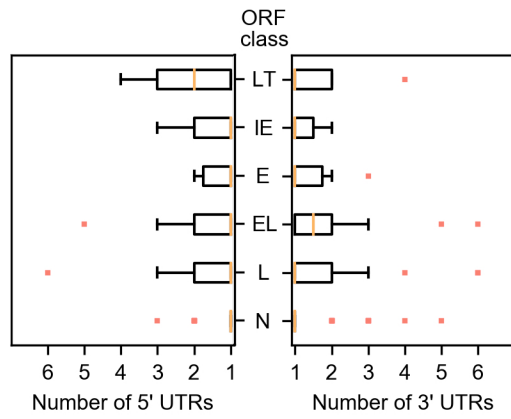
Figure S5 (related to Figure 4). Number of 5' and 3' UTR variants per ORF, by ORF class. LT = latent, IE = immediate early, E = early, EL = early-late, L = late, N = non-GenBank.
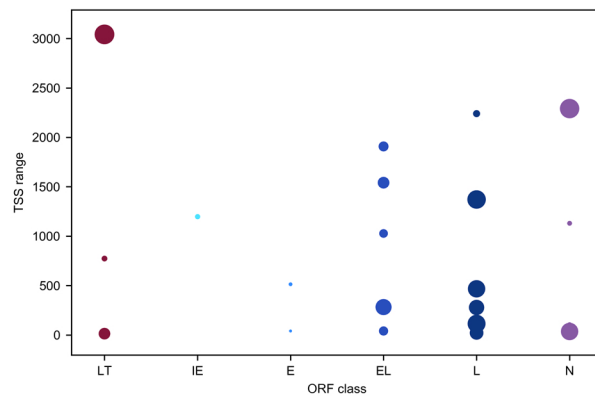
Figure S6 (related to Figure 4). Maximum genomic distance between TSSs for each ORF with multiple TSSs, scaled by the ratio of highest number of CAGE tags to lowest number of CAGE tags and arranged by ORF class. LT = latent, IE = immediate early, E = early, EL = early-late, L = late, N = non-GenBank.
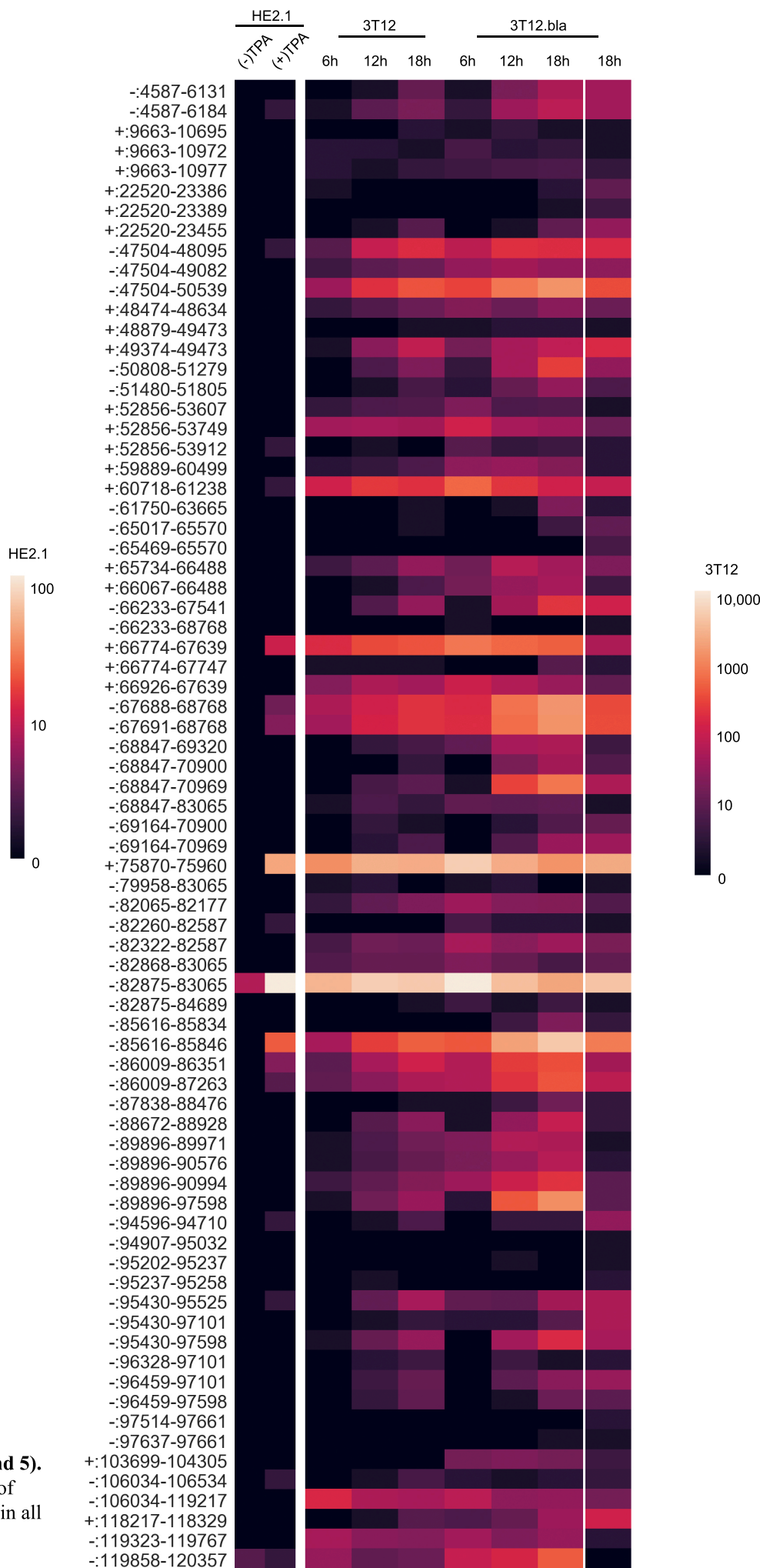
Figure S7 (related to Figures 2 and 5).
Heatmap of RNA-Seq read depths of TRIMD-identified splice junctions in all datasets.

Figure S8 (related to Figures 3 and 4). Distance from TSS to ORF for each ORF, scaled by CAGE tag depth and colored by ORF class. Maroon = latent gene, teal = immediate-early gene, light blue = early gene, royal blue = early-late gene, dark blue = late gene, purple = unknown.
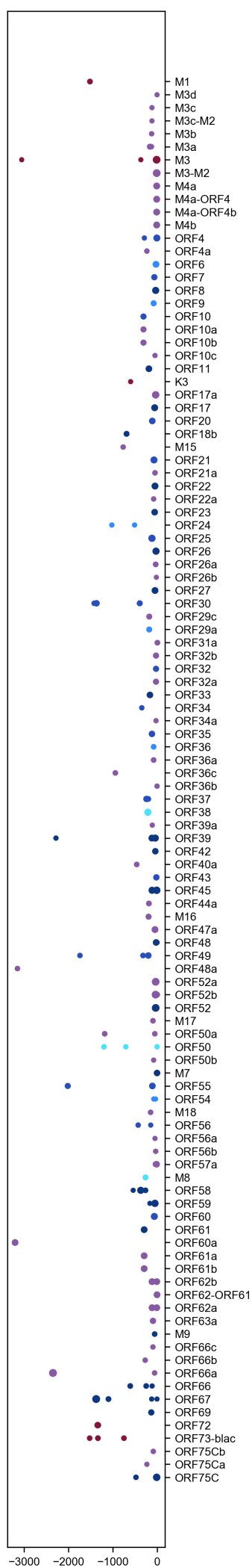
Table S1: **Primers** (related to STAR Methods).

Primer	Sequence (5'→ 3')
asORF75 F	TCCACCCATGTTCCCCGA
asORF75 R	<u>TAATACGACTCACTATAGGGAGA</u>GCGGGACCGTCCCTCAAC
asORF50 F	ACAGTCTGGTGGGATGTTGA
asORF50 R	<u>TAATACGACTCACTATAGGGAGA</u>ACAGCTACCTCAACCTCTGG
asORF50-SP F	GGAGGAAACTACTGTGGGTG
asORF50-SP R	<u>TAATACGACTCACTATAGGGAGA</u>GCTTACACAAACAGGCACTTAC
asORF48 F	CAGGTCATCAAGGGTCCAA
asORF48 R	<u>TAATACGACTCACTATAGGGAGA</u>ATAACTTCCCAGCCTCACCA
ORF55 F	CAAGAGGCTCGCTGATGTAG
ORF55 R	<u>TAATACGACTCACTATAGGGAGA</u>GTTTCCTCTGCCCTCTTGTC
ORF55 5'UTR F	GAAGAAATAGACGGGATGAGACG
c15777 5'UTR R	<u>TAATACGACTCACTATAGGGAGA</u>GAAGACTACTGGAAGAAGCACG
U6	ATATGGAACGCTTCACGAATT
tRNA-Tyr	CAGTCCTCCGCTCTACCAACTGAGC
gB F	GGCCCAAATTCAATTTGCCT
gB R	CCTGGACAACTCCTCAAGC
gB probe	ACAAGCTGACCACCAGCGTCAACAAC
ORF10 RT,R (TV_004)	GTAGTGTCCAAACAGGAGCAG
_c3244 F (TV_017)	TCTTCCTAAGAGCGATCACA
_c2410 F (TV_018)	TCTTCCTAAGCAGAGTGTTG
ORF61 RT (TV_007)	AATACCCATAGACCTGTCACTC
ORF61 R (TV_010)	CTGGTATGCACACTCCTTGAAA
_c8714 F (TV_025)	TGGTTCAGCGCAGGTGTGCT
_c20295 F (TV_026)	TGGTTCAGCGTTGTTTCATC

Note: For primers used to prepare template DNA for in-vitro transcription of RNA probes complementary to target region, a T7 promoter (underlined) was added to the R primer which is antisense to the transcript of interest.