Article:
Lulla, V, Dinan, AM, Hosmillo, M et al. (8 more authors) (2018) An upstream protein-coding region in enteroviruses modulates virus infection in gut epithelial cells. Nature Microbiology, 4. pp. 280-292. ISSN 2058-5276

https://doi.org/10.1038/s41564-018-0297-1

Supplementary information file

An Upstream Protein-Coding Region in Enteroviruses Modulates Virus Infection in Gut Epithelial Cells

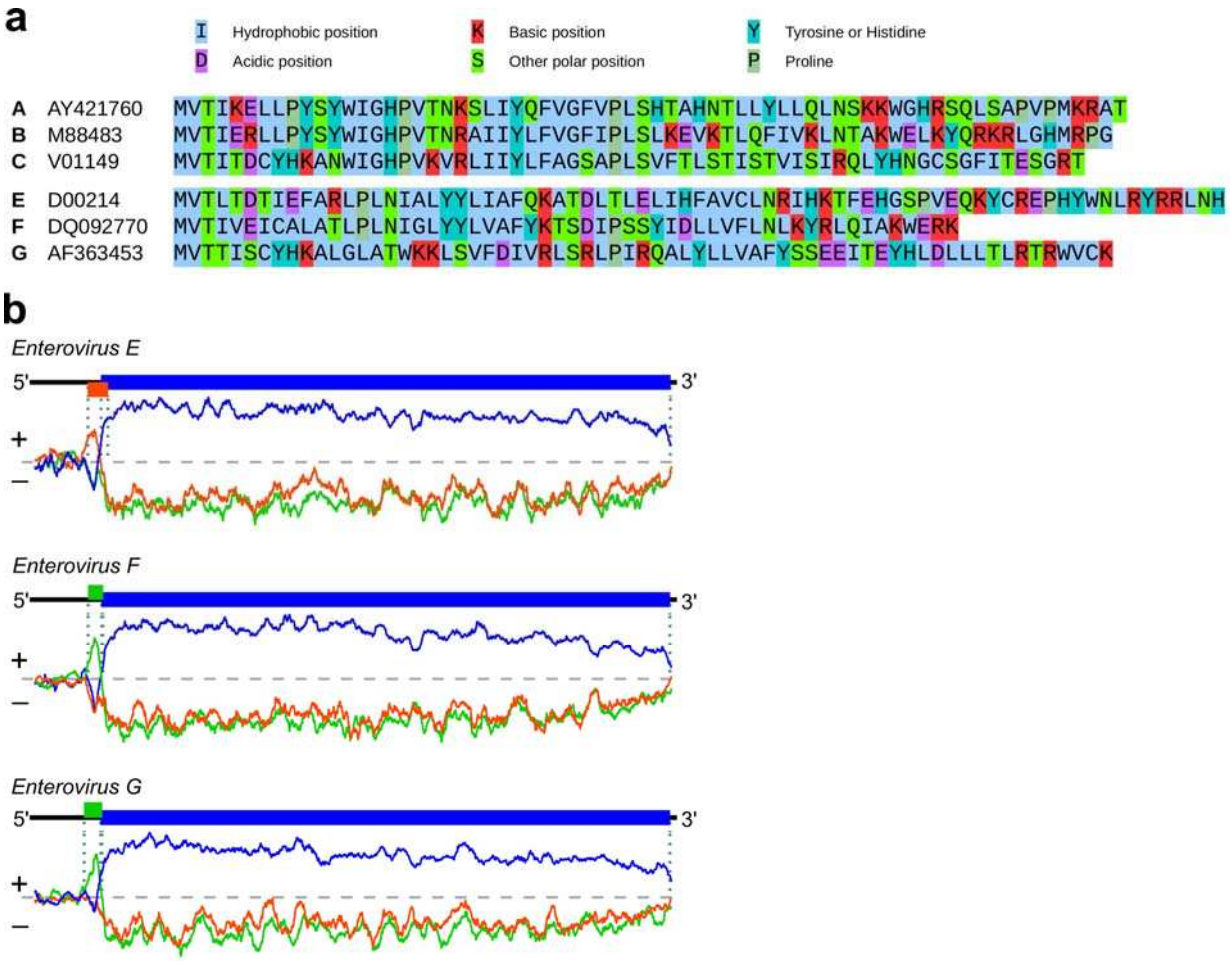
Valeria Lulla[1*], Adam M. Dinan[1], Myra Hosmillo[1], Yasmin Chaudhry[1], Lee Sherry[2], Nerea Irigoyen[1], Komal M. Nayak[3], Nicola J. Stonehouse[2], Matthias Zilbauer[3], Ian Goodfellow[1], Andrew E. Firth[1*].

[1] Division of Virology, Department of Pathology, Addenbrooke's Hospital, University of Cambridge, Cambridge, UK

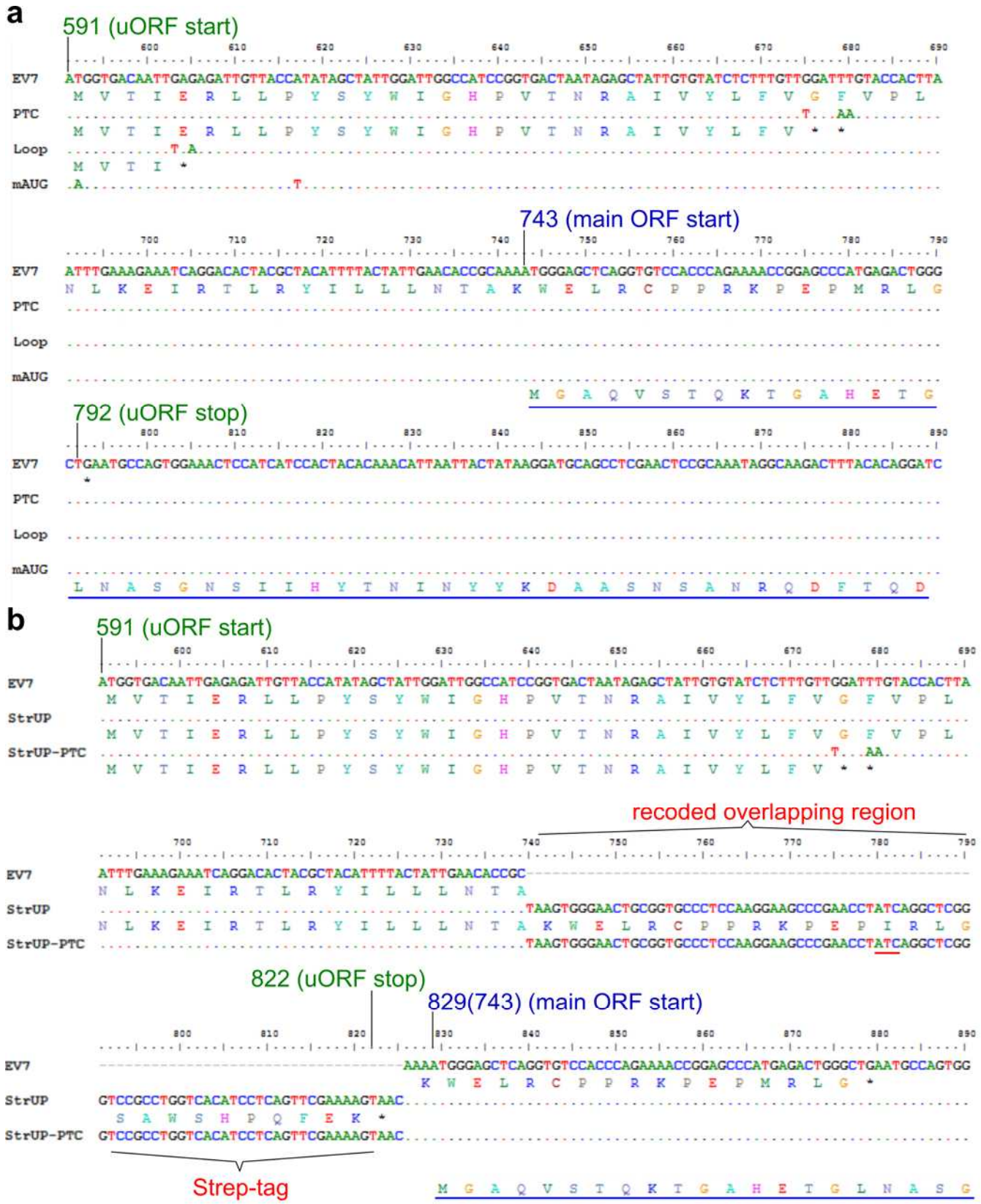
[2] School of Molecular and Cellular Biology, Faculty of Biological Sciences and Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds, UK.

[3] Department of Paediatrics, Addenbrooke's Hospital, University of Cambridge, Cambridge, UK

* Corresponding authors

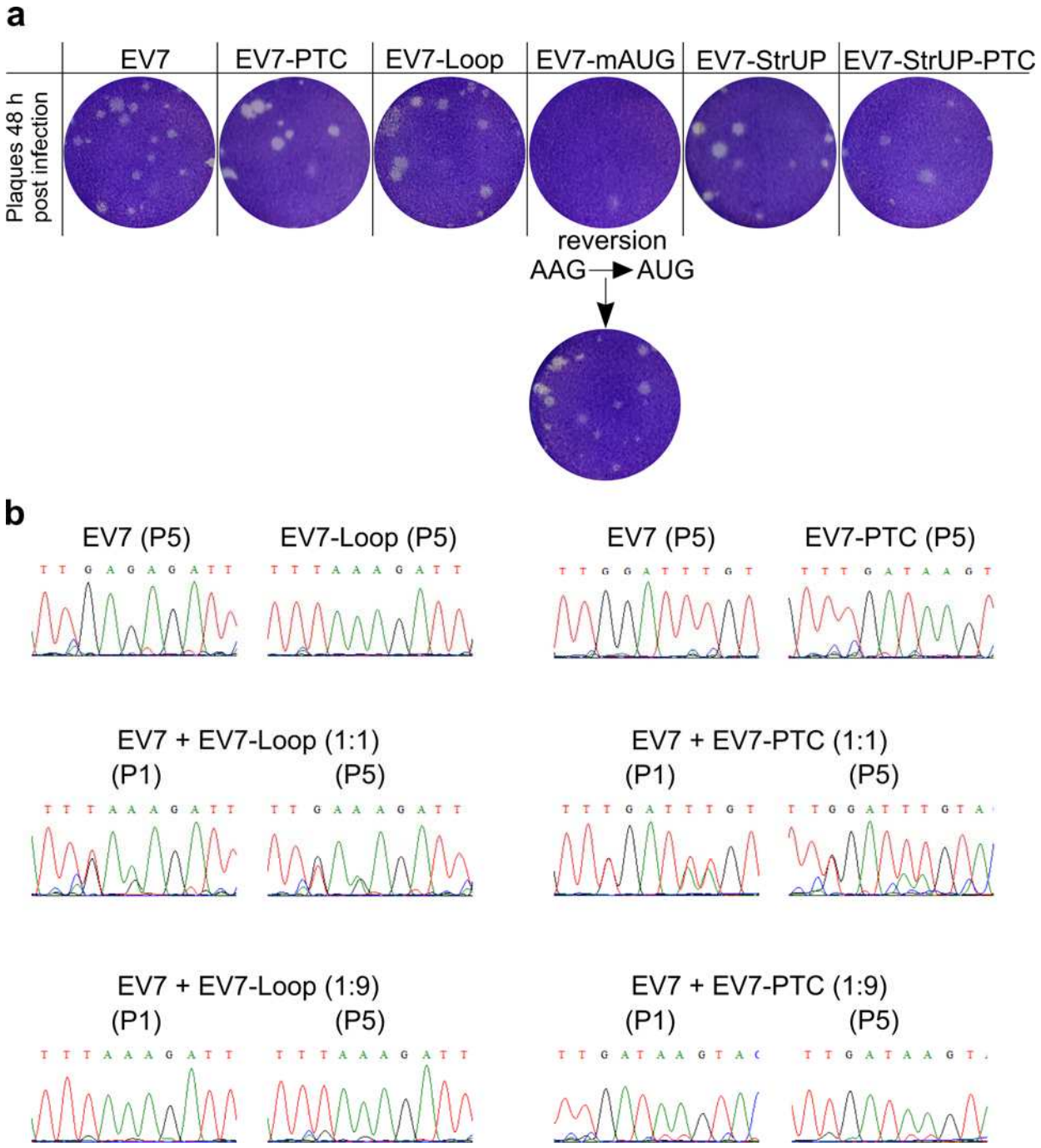


Supplementary Figure 1 │ Computational analysis of enterovirus sequences. (a) Sequences of the UP peptide for *Enterovirus A*, *B*, *C*, *E*, *F* and *G* reference sequences. Transmembrane domains are predicted in *A*, *B* and *C* but not or only weakly in *E*, *F* and *G*. Thus UP may have a different function in *Enterovirus E*, *F* and *G*, or may simply have an unusual poorly predicted transmembrane region. **(b)** Coding potential in the three reading frames (indicated by the three colors) for *Enterovirus E*, *F* and *G* as measured with MLOGD, for sequences that contain the uORF. Positive scores indicate that the sequence is likely to be coding in the given reading frame. Reading frame colors correspond to the genome maps shown above each plot, indicating the polyprotein ORF and uORF in the corresponding reference sequences.

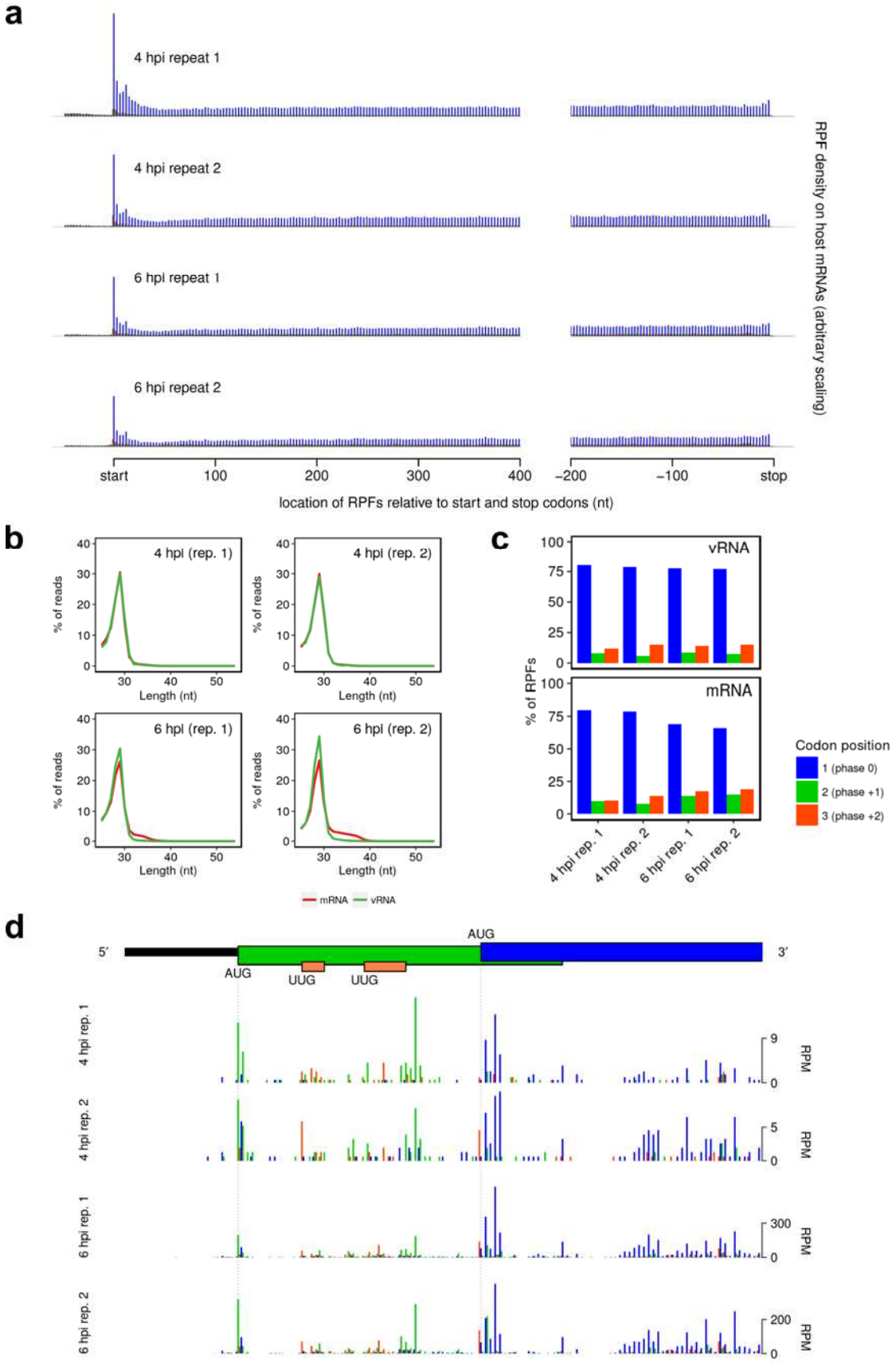


Supplementary Figure 2 | Nucleotide and amino acid alignments of EV7 mutants. (a) Alignment of EV7, EV7-PTC, EV7-Loop and EV7-mAUG mutants. Only nucleotide differences from wt EV7 are shown for the mutants. Amino acid sequences are shown for wt UP (EV7) and the truncated UP in the PTC and Loop mutants. The polyprotein amino acid sequence is underlined in blue. **(b)** Alignment and

tagging strategy for EV7-StrUP and EV7-StrUP-PTC mutants. Since the uORF overlaps the polyprotein ORF, in order to add a C-terminal Strep tag to UP the overlap region was duplicated upstream of the polyprotein ORF initiation site and the duplicated region was synonymously mutated to prevent recombination. Sequence encoding a Strep tag was appended to the end of the duplicated uORF sequence, the polyprotein ORF initiation context was maintained, and an AUG codon in the duplicated region was mutated (AUG to AUC, underlined; Met to Ile change in UP) to avoid inhibiting scanning 43S preinitiation complexes from reaching the polyprotein AUG.

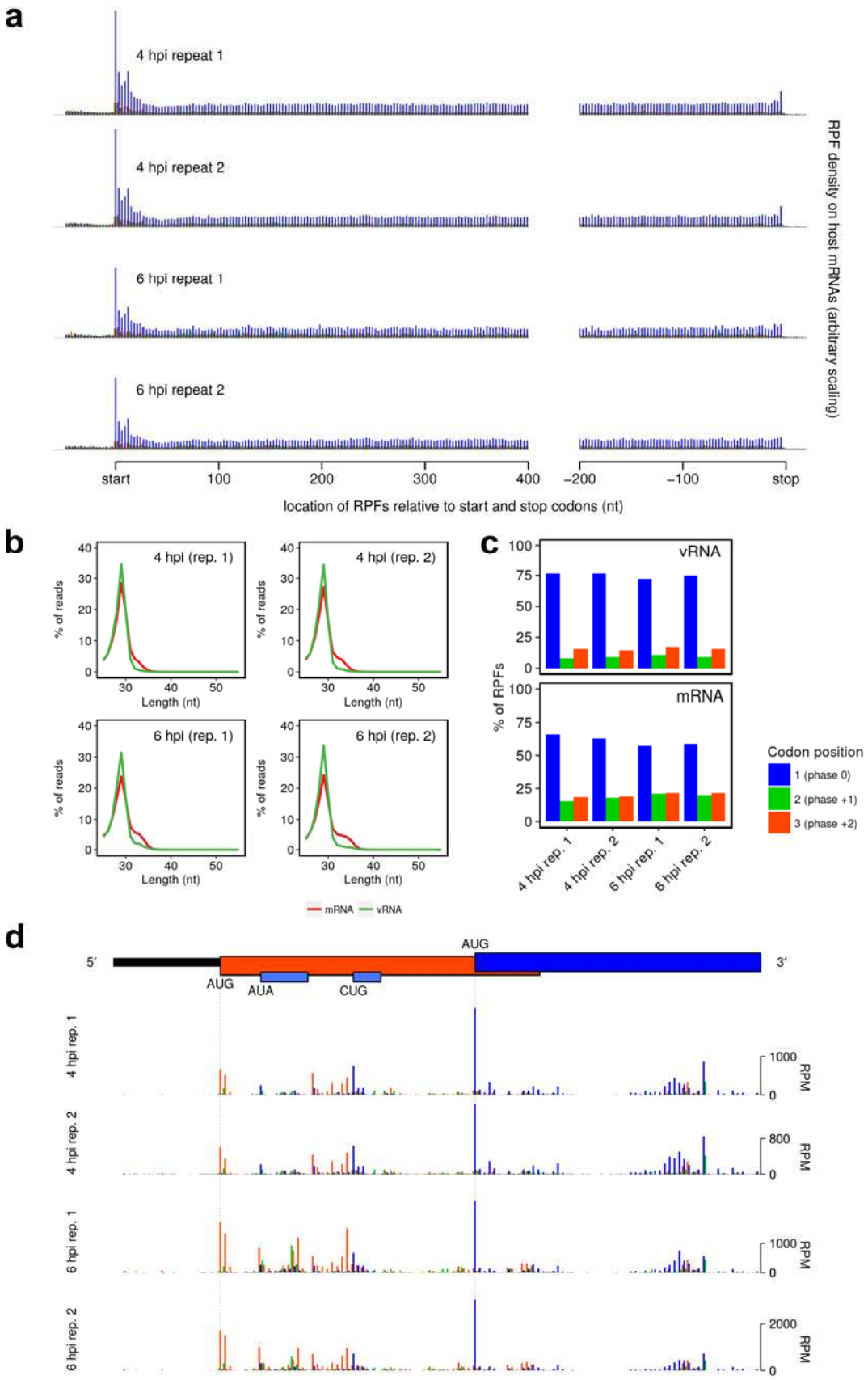


Supplementary Figure 3 │ Analysis of wt and mutant EV7 viruses. (a) Plaque morphology for EV7 and mutant viruses EV7-PTC, EV7-Loop, EV7-mAUG (before and after reversion), EV7-StrUP and EV7-StrUP-PTC. **(b)** Sequencing chromatograms of RT-PCR products for the competition experiment of EV7, EV7-PTC and EV7-Loop mutants at initial infection (P1) and after five passages (P5). Experiments were repeated three times independently with similar results (**a-b**).



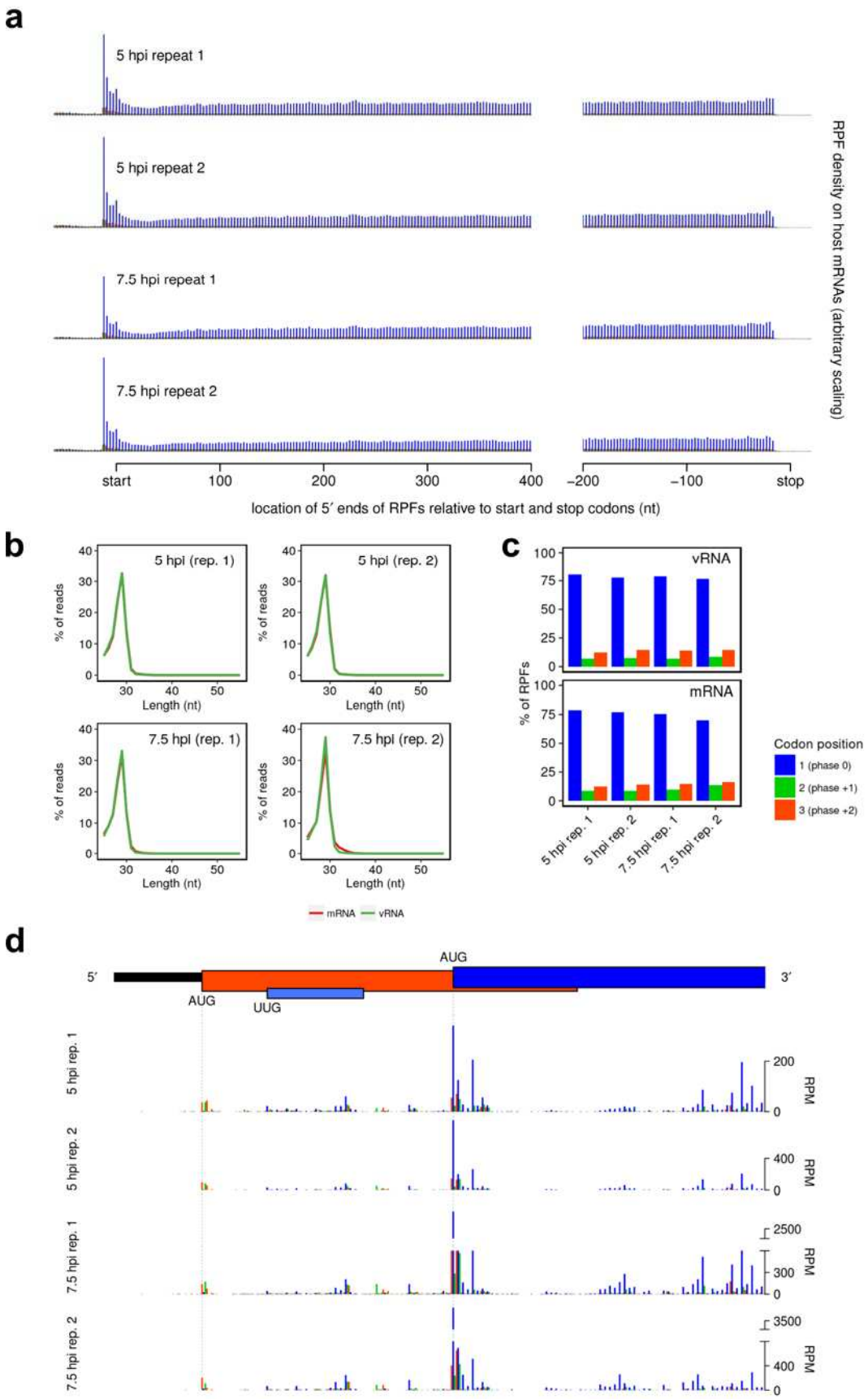
Supplementary Figure 4 | Assessment of ribosome profiling quality for EV7 libraries. (a) Histograms of RPF approximate P-site positions relative to annotated initiation and termination sites

summed over all host mRNAs. **(b)** Relative length distributions for Ribo-Seq reads mapping to virus (green) and host (red) mRNA coding regions. Note the degradation in host, but not virus, Ribo-Seq quality at the later timepoint (cf. ref. [1]). **(c)** Phasing of 5′ ends of Ribo-Seq reads that map to the viral polyprotein ORF (vRNA) or host coding regions (mRNA). **(d)** Single-nucleotide resolution view of RPF densities (reads per million mapped reads; RPM) in the uORF and flanking regions of the EV7 genome. RPF phasing colors are shown relative to the polyprotein reading frame. Note that nucleotide-to-nucleotide variation in RPF counts may be influenced by technical biases besides ribosome codon dwell-times. Reads in the orange phase may derive from translation of short non-AUG initiated overlapping ORFs (annotated), though we have no reason to suspect that these encode functional protein products and ribosomes which translate such short ORFs may be competent for reinitiation at the polyprotein AUG.



Supplementary Figure 5 | Assessment of ribosome profiling quality for PV1 libraries. (a) Histograms of RPF approximate P-site positions relative to annotated initiation and termination sites summed over all host mRNAs. **(b)** Relative length distributions for Ribo-Seq reads mapping to virus

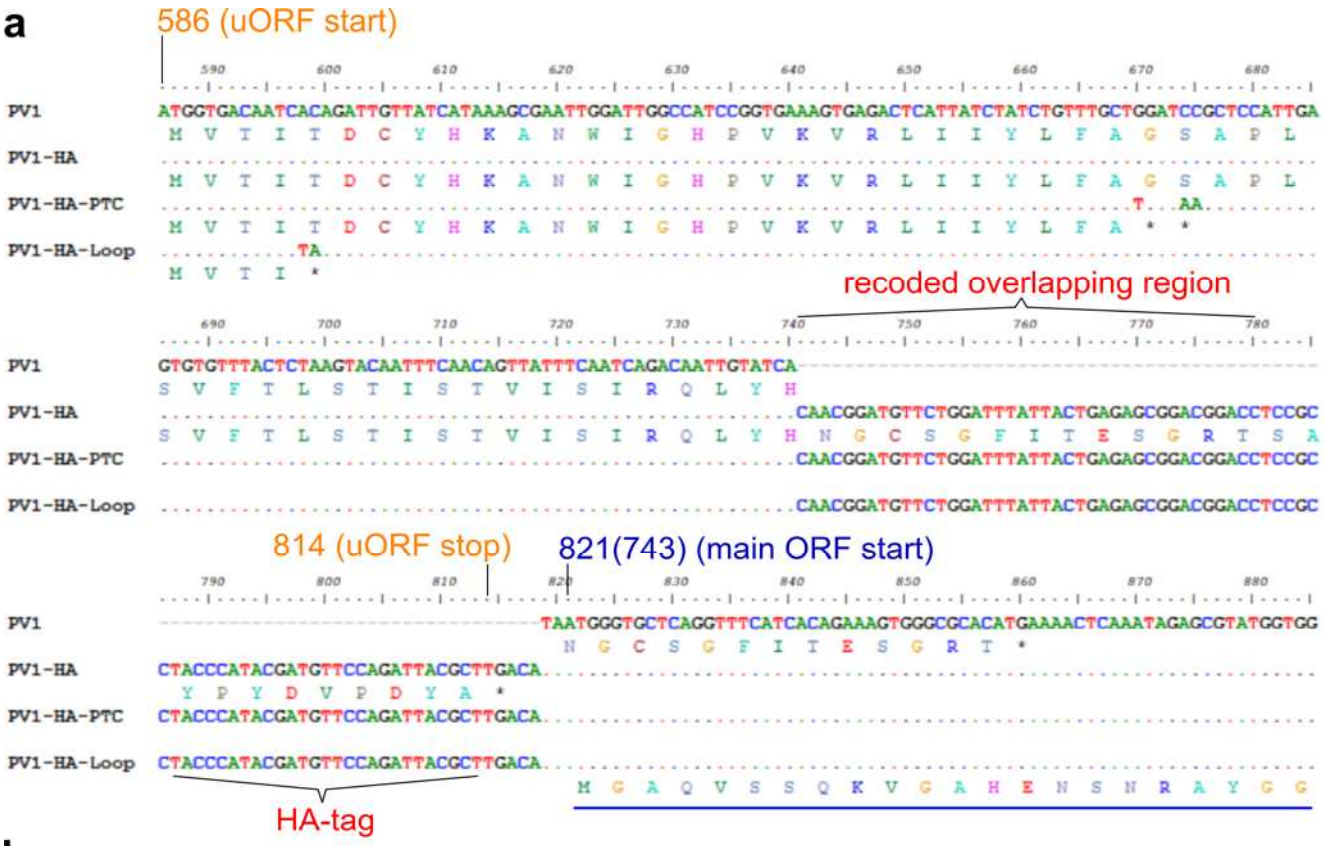
(green) and host (red) mRNA coding regions. Note the higher quality (sharper peak) of the virus reads compared to the host reads (cf. ref. [1]). **(c)** Phasing of 5′ ends of Ribo-Seq reads that map to the viral polyprotein ORF (vRNA) or host coding regions (mRNA). **(d)** Single-nucleotide resolution view of RPF densities (reads per million mapped reads; RPM) in the uORF and flanking regions of the PV1 genome. RPF phasing colors are shown relative to the polyprotein reading frame. Reads in the blue phase may derive from translation of short non-AUG initiated overlapping ORFs (annotated), though we have no reason to suspect that these encode functional protein products and ribosomes which translate such short ORFs may be competent for reinitiation at the polyprotein AUG.



Supplementary Figure 6 | Assessment of ribosome profiling quality for EV-A71 libraries. (a) Histograms of RPF approximate P-site positions relative to annotated initiation and termination sites

summed over all host mRNAs. **(b)** Relative length distributions for Ribo-Seq reads mapping to virus (green) and host (red) mRNA coding regions. **(c)** Phasing of 5′ ends of Ribo-Seq reads that map to the viral polyprotein ORF (vRNA) or host coding regions (mRNA). **(d)** Single-nucleotide resolution view of RPF densities (reads per million mapped reads; RPM) in the uORF and flanking regions of the EV-A71 genome. RPF phasing colors are shown relative to the polyprotein reading frame. Reads in the blue phase may derive from translation of a short non-AUG initiated overlapping ORF (annotated), though we have no reason to suspect that this encodes a functional protein product and ribosomes which translate such short ORFs may be competent for reinitiation at the polyprotein AUG.

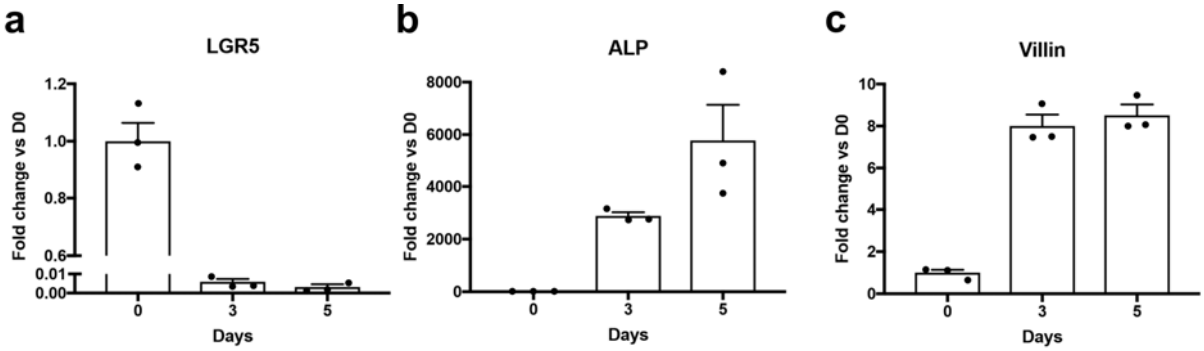
a



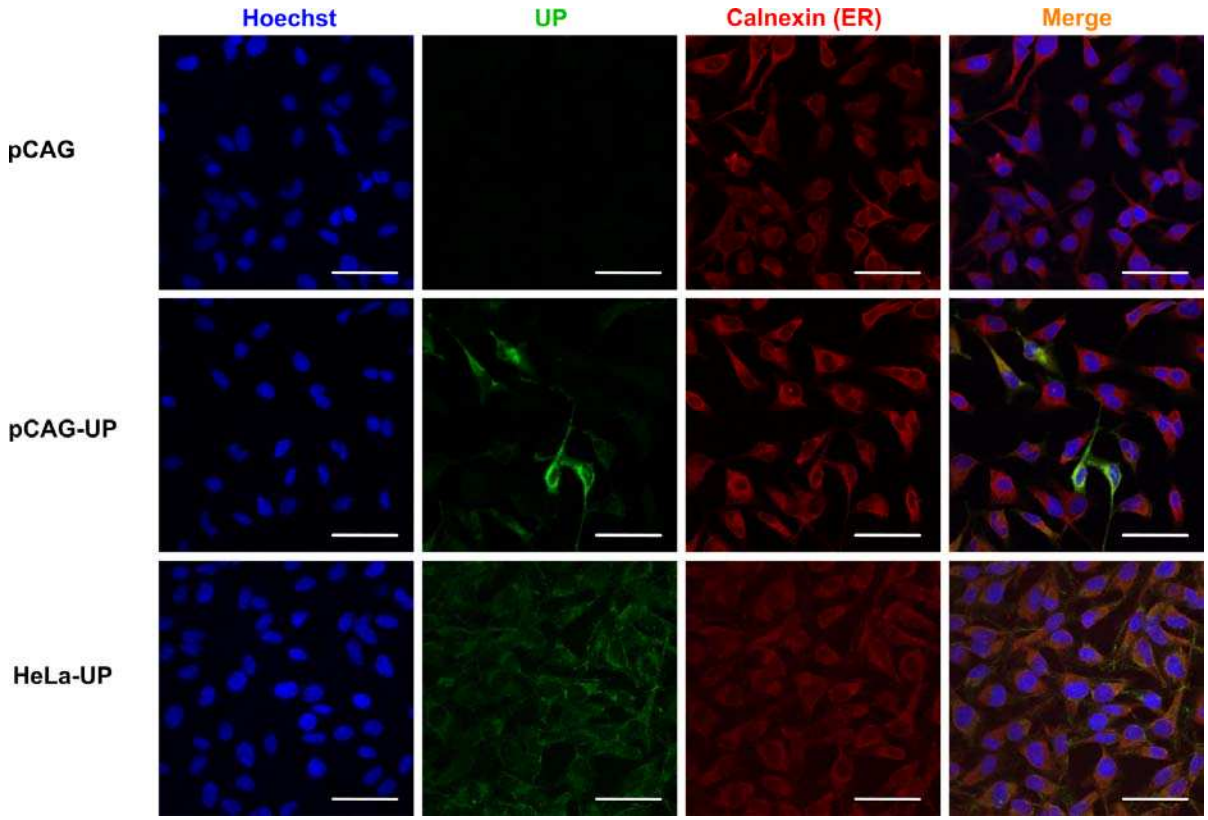
b

Construct	Virus titer (P1) PFU/ml	Plaque size	RT-PCR analysis
PV1 (wt)	3×10^9	wt	wt
PV1-HA	3×10^8	small	HA
PV1-HA-PTC	8×10^8	small	HA-PTC
PV1-HA-Loop	7×10^8	small	HA-Loop

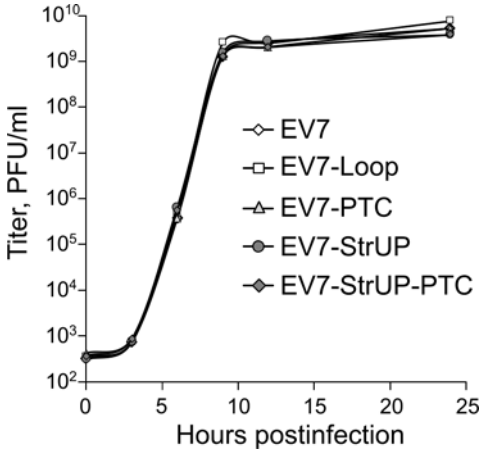
Supplementary Figure 7 | Nucleotide and amino acid alignments and properties of HA-tagged PV1 mutants. (a) Alignment and tagging strategy for PV1-HA, PV1-HA-PTC and PV1-HA-Loop mutants. Since the uORF overlaps the polyprotein ORF, in order to add a C-terminal HA tag to UP the overlap region was duplicated upstream of the polyprotein ORF initiation site, the duplicated region was synonymously mutated to prevent recombination, and sequence encoding an HA tag was appended to the end of the duplicated uORF sequence. **(b)** Titers of the viruses after one passage in RD cells (P1), comparative plaque sizes, and RT-PCR analysis of viral RNA isolated after three passages in RD cells.



Supplementary Figure 8 │ Differentiation of human intestinal organoid cultures. Results of qRT-PCR showing fold change in transcript levels in differentiated terminal ileum organoids (day 3 and 5) relative to transcript levels in undifferentiated organoids (day 0). HPRT1 transcript was used for normalization. Shown are **(a)** a stem cell marker, leucine-rich repeat-containing G-protein coupled receptor 5 (LGR5), **(b)** a mature enterocyte marker, alkaline phosphatase (ALP), and **(c)** an epithelial cell marker, villin (means ± s.d.; $n = 3$ biologically independent experiments each).

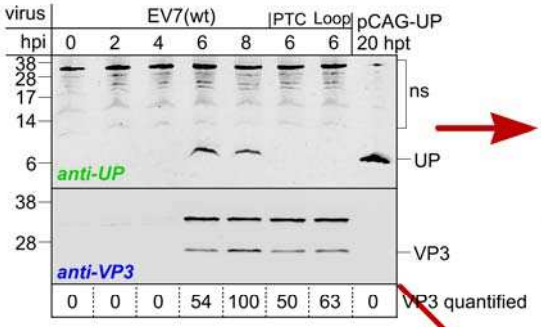


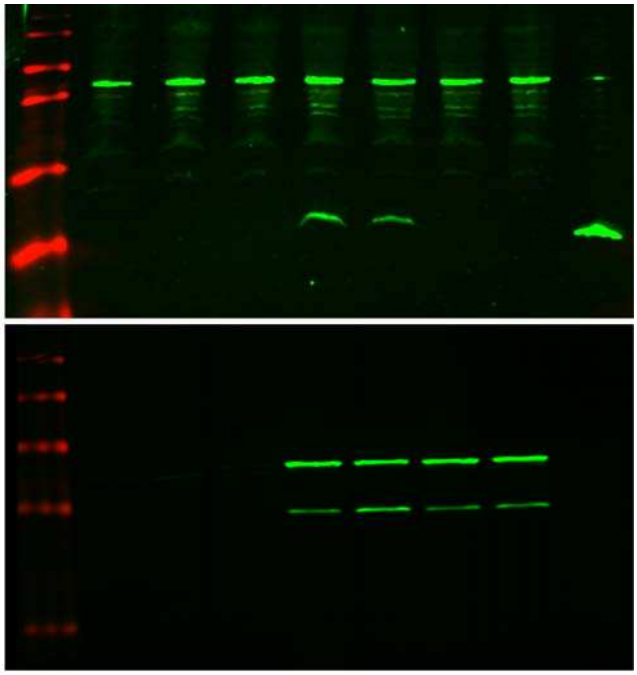
Supplementary Figure 9 │ Immunofluorescence analysis of EV7 UP expressed in HeLa cells. Confocal images of HeLa cells transfected with pCAG or pCAG-UP, and the HeLa-UP cell line, stained for UP (green), ER (calnexin, red) and nuclei (Hoechst, blue). The images are a projection of a *z*-stack. Scale bar: 25 μm. The experiment was repeated three times independently with similar results.



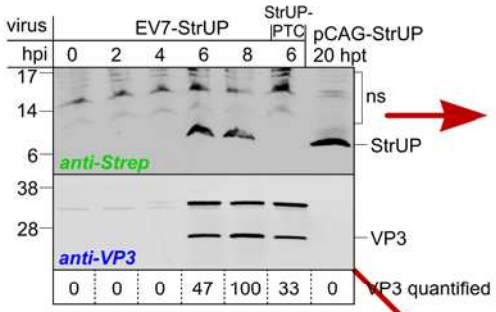
Supplementary Figure 10 │ A repeat of the Fig. 3c growth curve assay for EV7 and mutants.

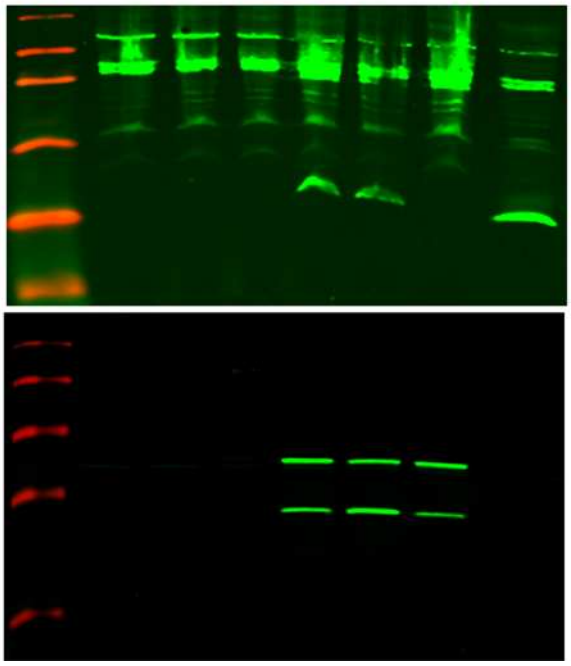
a


virus
EV7(wt)
PTC
Loop
pCAG-UP
hpi
0 2 4 6 8 6 6
20 hpt
38
28
17
14
6
ns
UP
anti-UP
38
28
VP3
anti-VP3
0 0 0 54 100 50 63 0
VP3 quantified

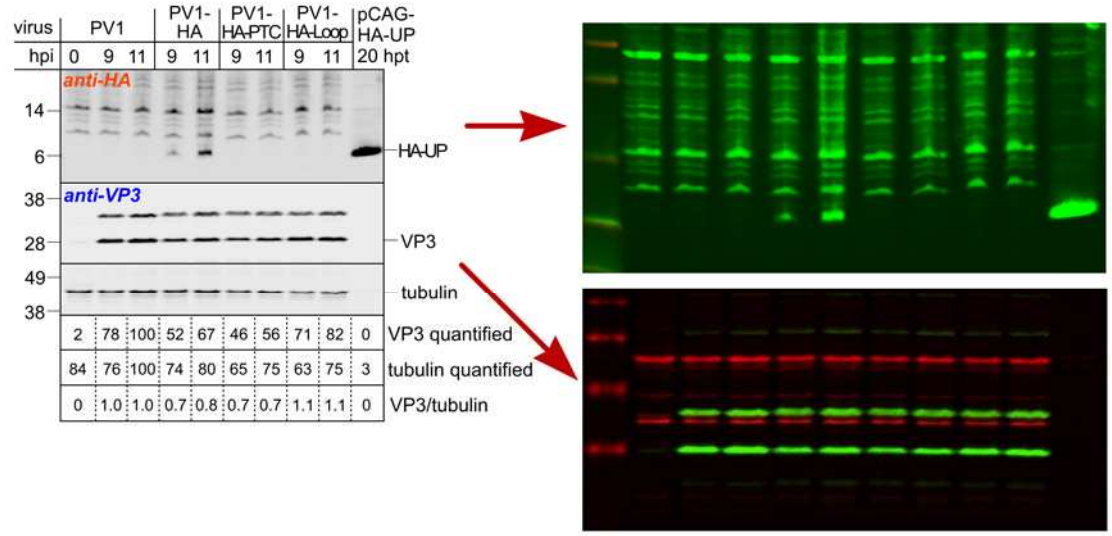


b

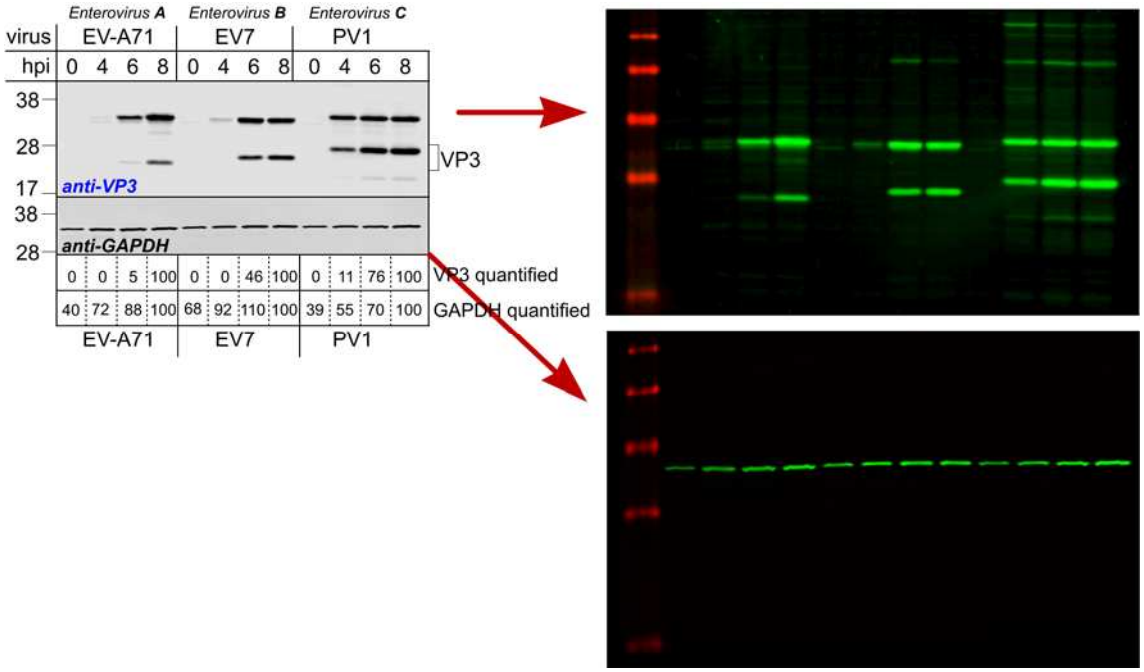

virus
EV7-StrUP
StrUP-
PTC
pCAG-StrUP
hpi
0 2 4 6 8 6
20 hpt
17
14
6
ns
StrUP
anti-Strep
38
28
VP3
anti-VP3
0 0 0 47 100 33 0
VP3 quantified

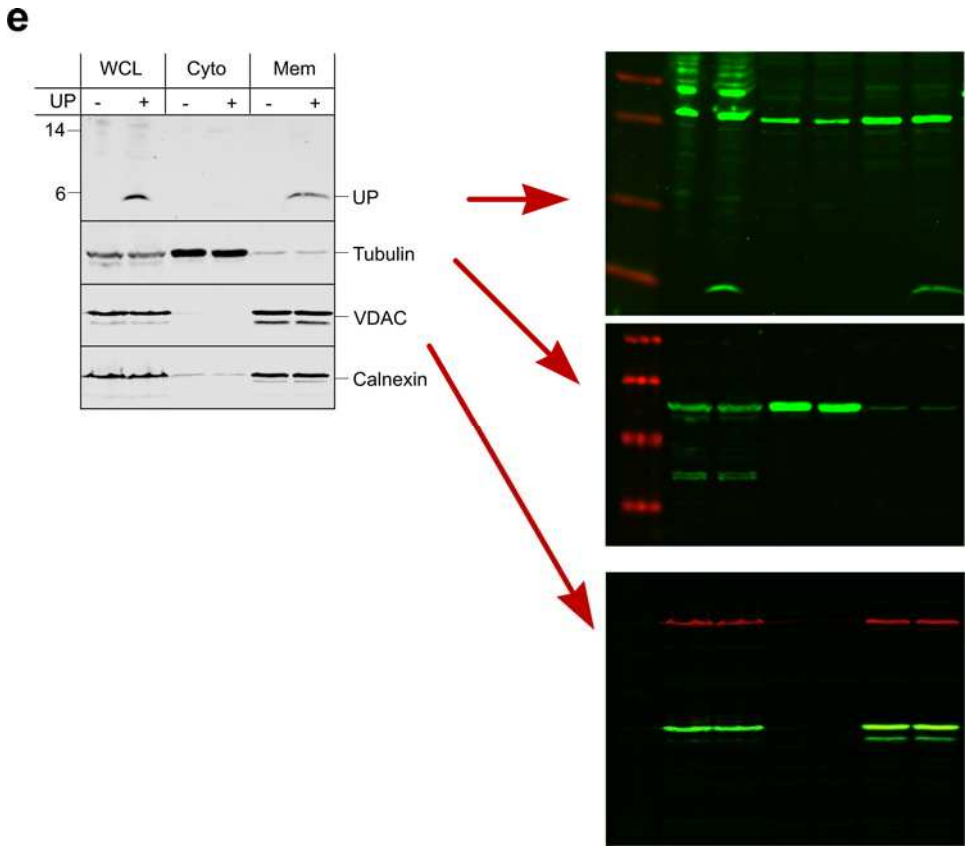


c

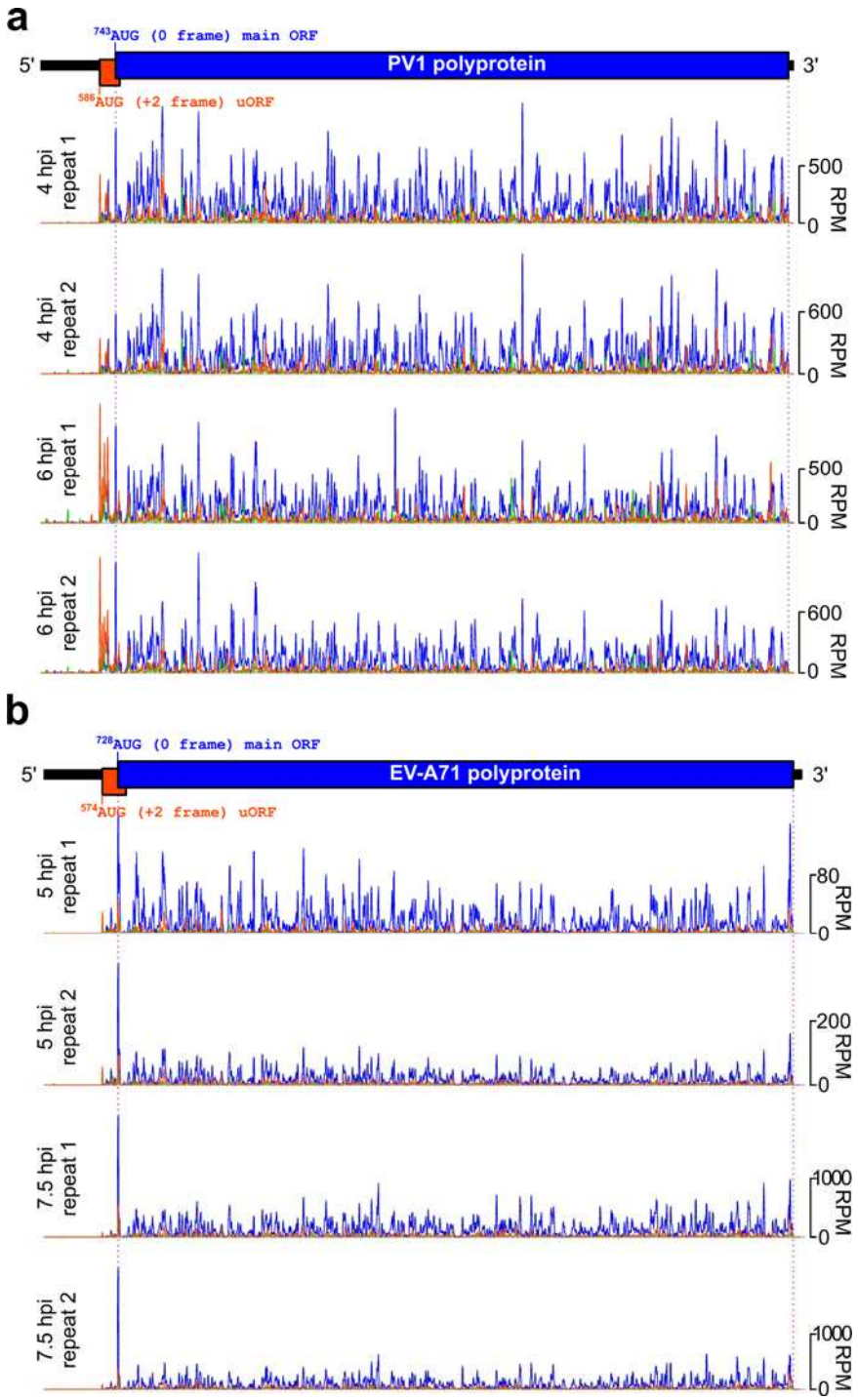

virus
PV1
PV1-HA
PV1-HA-PTC
PV1-HA-Loop
pCAG-HA-UP
hpi
0 9 11
9 11
9 11
9 11
20 hpt
anti-HA
14
6
HA-UP
anti-VP3
38
28
VP3
49
38
tubulin
2 78 100 52 67 46 56 71 82 0 VP3 quantified
84 76 100 74 80 65 75 63 75 3 tubulin quantified
0 1.0 1.0 0.7 0.8 0.7 0.7 1.1 1.1 0 VP3/tubulin

d

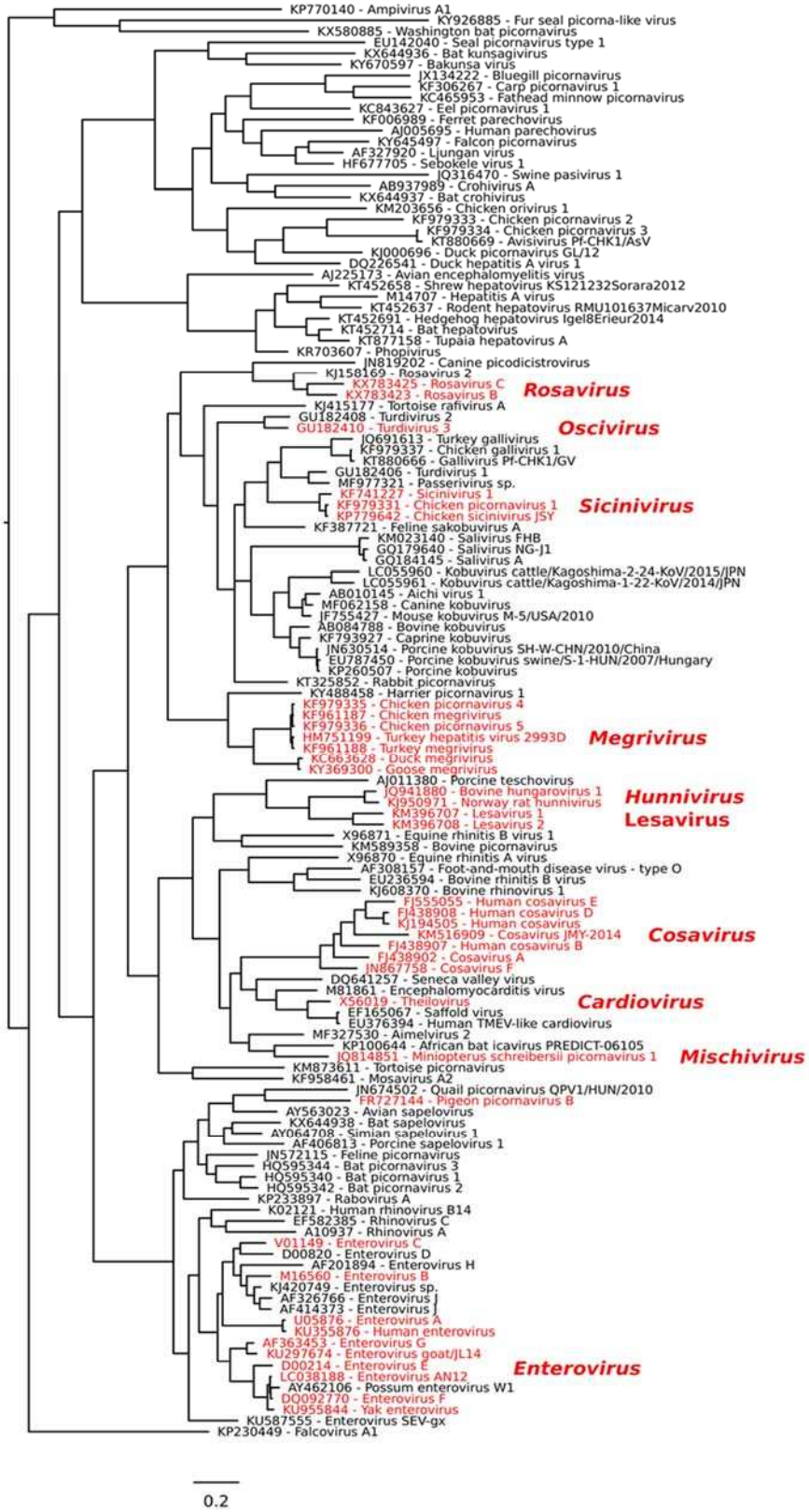

Enterovirus A
Enterovirus B
Enterovirus C
virus
EV-A71
EV7
PV1
hpi
0 4 6 8
0 4 6 8
0 4 6 8
38
28
17
anti-VP3
VP3
38
28
anti-GAPDH
0 0 5 100 0 0 46 100 0 11 76 100 VP3 quantified
40 72 88 100 68 92 110 100 39 55 70 100 GAPDH quantified
EV-A71
EV7
PV1



Supplementary Figure 11 │ Original western blot scans for Fig. 3a (a), 3b (b), 4d (c), 4f (d), 6c (e). Gels for structural protein immunoblotting were loaded with the same samples used for UP immunoblotting but at 1:20 dilution (a, b, c, d). The bands were quantified where possible (a, b, c, d).



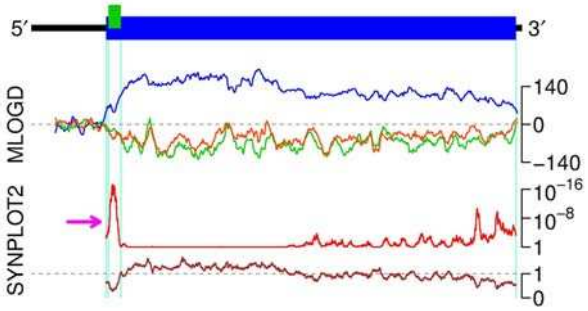
Supplementary Figure 12 │ Ribosome profiling repeat datasets. (a) Ribosome profiling of PV1-infected cells at 4 and 6 hpi. **(b)** Ribosome profiling of EV-A71-infected cells at 5 and 7.5 hpi. Note that the repeat 1 datasets are the same as those shown in Fig. 4. Ribo-Seq RPF densities in reads per million mapped reads (RPM) are shown with colors indicating the three phases relative to the main ORF (blue – phase 0, green – phase +1, orange – phase +2), each smoothed with a 3-codon sliding window.



Supplementary Figure 13 | Phylogenetic tree of the *Picornaviridae* family. All RefSeqs annotated as family *Picornaviridae* were obtained from NCBI on 28 June 2018, the polyprotein amino acid sequences were extracted and aligned with MUSCLE, and the RNA-dependent RNA polymerase (RdRp) region was excised and used for phylogenetic analysis. A maximum likelihood phylogenetic tree was estimated using the Bayesian Markov chain Monte Carlo method implemented in MrBayes version 3.2.3, sampling across

the default set of fixed amino acid rate matrices, with 10 million generations, discarding the first 25% as burn-in. The tree was visualized with FigTree. Taxa that have a putative additional 5′ or 3′ ORF are indicated in red; it is possible that other taxa also have additional ORFs as limited sequence data for some taxa precludes the ability to analyse purifying selection.

a



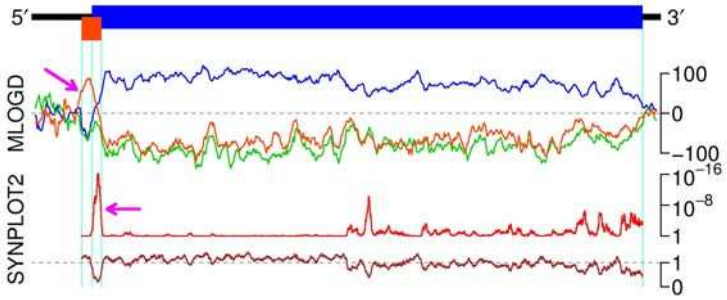
b

```
JN867757   MATRGLLLITFMLTSIMALLMLHPKVSVLTLLLRMVMYLAFLDLLLLHSTLLDFWQTQRLRTQPIWRTEF---------
FJ555055   MAIKGLLLITFMLTSIMALLMLLPRGSVAILLLKMVLSLAFLDLQLLLLMRPDFLQPRAQKTLRTWVIEF---------
FJ438908   MAIRVLLSIIFMLTHIMLLLTLLPPQSGAILLLKMVLYLAFSEVLLLLSLPLHY-------------------------
KJ194505   MAIRELLSTTFMLTHIMLLLMLLPPQSGAILLLKMVLYLVFSEVLLLHSLPLHYWLNQR--------------------
FJ438907   ---MAQLLITFMLIIIMALLMLLPRELGPLLLPKTVLYLVFLEWQVLLLMHFPYWLPQR--------------------
KX545380   ---MELQLIIFMQTITMALLMLLPKALGPLLLLKMVMSLVFLDLHLVLLMPYLSLQHPV--------------------
FJ438903   MATRELLLITFIQTSTMLLLMLLPKVLGPLLLLKMVMFLAFLDLQVVLLMLSLFSPLPE--------------------
JN867759   MATRELLVITFIQTSIMLLLMLLPKVLGPLLLLKTAMYLAFLDLQAVLLRLSLFSPHHKGRNKTSKKIESLLEMLATHL
JN867756   MATREQLLITFIQTSIMLLLMLLPKVLGPLLLLKTAMYLAFLDLQVVLLMLSLFSHPPE--------------------
AB920345   MATREQLLITFIQTSIMLLLMLLPKVLGPLLLLKTAMYLAFLDLQVVLLMLSLFSPPQ--------------------
GU968209   MATREQLLITFIQTSIMLLLMLLPKVLGPLLLLKTAMYLAFLDLQVVLLMLSLFSPPRG-------------------
FJ438904   LATREQLLITFIQTSIMLLLMLLPKVLGPLLLLKTAMYLAFLDLQVVLLMLSLFSPPP--------------------
KJ396940   MATREQLLITFIQTSIMLLLMLLPKVLGPLLLLKTATYLAFLDLQAVLLMLSLFSPPRE-------------------
FJ438902   MATREQLLITFIQTSIMLLLMLLPKVLGPLLLLKTATYLAFLDLQVVLLMLSLFSPHHEPRQE----------------
JN867758   MAMRELQLTIFMQTITMALLMLLPKELAPVLLLKMVMFLAFSDLQVVLLMLYLYWPTPGQRIVHILRTES---------
KM516909   --MRALWLTIFMQTIIMALLMLLPRELGPLLLPKMVMFLAFSDLQVVLLMLYLF-----------------------
                    *:    * ** * *      ** . .  *.* :   :
```

I	Hydrophobic position	K	Basic position	Y	Tyrosine or Histidine
D	Acidic position	S	Other polar position	P	Proline

Supplementary Figure 14 | Putative additional 5′ ORF in cosaviruses. (a) Comparative genomic analysis of genus *Cosavirus* sequences (see part B for accession numbers; reference sequence in red). Top – Genome map of reference sequence indicating the polyprotein ORF (blue) and the putative additional ORF (green). Middle – Coding potential in the three reading frames (indicated by the three colors corresponding to the genome map colors). Positive scores indicate that the sequence is likely to be coding in the given reading frame. Bottom – Analysis of conservation at polyprotein-frame synonymous sites. The red line shows the probability that the observed conservation could occur under a null model of neutral evolution at synonymous sites, whereas the brown line depicts the ratio of the observed number of substitutions to the number expected under the null model. MLOGD is generally more appropriate for detecting non-overlapping ORFs whereas SYNPLOT2 is appropriate for detecting overlapping ORFs (pink arrow). Other peaks in synonymous site conservation likely correspond to functional RNA elements embedded within the polyprotein coding sequence (cf. ref. [2]). **(b)** Amino acid alignment of the putative additional protein. Amino acids are color-coded according to their physicochemical properties. Transmembrane regions predicted by Phobius to be present in *every* sequence in the alignment are indicated with pink bars above the alignment.

a



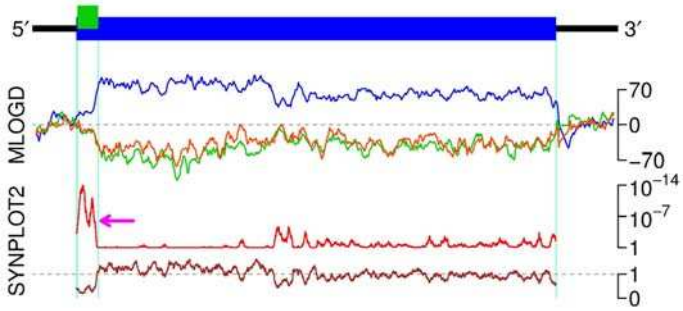
b

```
MG846480   MFYE-LLIIAALVVILVSPHDTQCLIVVCLVNFLVYFLWTIPGLDLPLYPAIQHGTFYWTSADLCEALHLLFWLCLRWLA
MG846487   MLLELLLVVVTLVIVVASPGEFYQLLAICTYNFLLIFLWLVGLIELSNYPTIQHGAYYWTNNDLCETLHLLFWLCLRWLA
MG846482   MLIEISLVVVALVVVVVSHSESHQLLTVCIYNFALTFLWLFGFLELPNYPAVQHGVYYWTSSDLCETLHLLFWLCLRWLA
KF741227   MFLEISLVTWTLIIVVVSPSELSQLLIICAYNFVLIFFWLIGLIELPGYPAIQHGAYYWTNTDLCETLYLLFWLCLRWLA
KT880665   MLVEIFLVIIALIAVVVSPNELIQLLIVCVYNFALIFLWLFGILELPDYPTIQHGALYWTNVDLCETLHLLFWLCLRWLA
MG846484   MFLESLLVAITLVIVVASPSELHQLFTICVYNFALIFLWLVEILELPNYPAIQHGALYWTNTDLCETLHLLFWLCLRWLA
MG846481   MLVETLLIIIALALVVVSSSELHQLLAICVYNFALIFLWLLGILELPSYPAIQHGALYWTSADLCETLHLLFWLCLRWLA
MG846485   MLLEFLLVTTALTIVVLSPSELHQLLAICVYNFALIFLWLLGILELPSYPAIQHGALYWTSADLCETLHLLFWLCLRWLA
MG846486   MLLEFLLVTTALAIVVLSPSELHQLLAICVYNFALIFLWLLGILELPSYPAIQHGALYWTSADLCETLHLLFWLCLRWLA
MG846483   MFLEVFLVIVALVVVVLSPSELSQLLTICIYNFVLIFLWLFGTLELPNYPTIHHGAFYWTHTDLCETLHLLFWLCLRWLA
KY069113   MFLEVFLVIVALVVVVLSPSELSQLLTICIYNFVLIFLWLFGTLELPNYPTIHHGAFYWTHTDLCETLHLLFWLCLRWLA
KY069112   MLLESLLVIVALVVVVVSPSELPQLFAICIYNFVLIFLWLFGTLELPNYPTIHHGAFYWTHTDLCETLHLLFWLCLRWLA
KP779642   MIVEVLLVVIVLAIVISSPNELTQLLAICIYNFVLIFLWLFGILELPNYPTVRHGAQYWTRTDLCETLHLLFWLCLRWLA
           *: *  *:  .*  :: *  :   *: :*  ** : *:* .  ::*. **::.**. ***  ****:*:***********

MG846480   FMLWFISPISPVVFTLDPPTEY
MG846487   LLLWFILPTSPIVFALDPPTQH
MG846482   FMFWFTLPTSPIVFTLDPPTQY
KF741227   FMLWFILPTSPVVFTLDPPTEH
KT880665   FMFWFIAPTSPVVFTLDPPTEH
MG846484   FLLWFILPASPIVFTLDPPAQY
MG846481   FMLWFILPTSPIVFTLDPPAQY
MG846485   FMLWFILPTSPIVFTLDPPAQY
MG846486   FMLWFILPTSPIVFTLDPPAQY
MG846483   FMFWFILPTSPIVFTLDPPAQH
KY069113   FMFWFILPTSPIVFTLDPPAQH
KY069112   FMFWFILPTSPIVFTLDPPAQH
KP779642   FMFWFILPTSPIVFTLDPPAQY
           :::**  * **:**:****:::
```

Supplementary Figure 15 | Putative additional 5′ ORF in siciniviruses. (a) Comparative genomic analysis of genus *Sicinivirus* sequences (see part B for accession numbers; reference sequence in red). Top – Genome map of reference sequence indicating the polyprotein ORF (blue) and the putative additional ORF (orange). Middle – Coding potential in the three reading frames (see Fig. S14 for details). Bottom – Analysis of conservation at synonymous sites (with the reading frame of the polyprotein being used where the two ORFs overlap) (see Fig. S14 for details). MLOGD and SYNPLOT2 detect, respectively, the non-overlapping and the overlapping portions of the additional ORF (pink arrows). **(b)** Amino acid alignment of the putative additional protein. Amino acids are color-coded according to their physicochemical properties (see Fig. S14). Transmembrane regions predicted by Phobius to be present in *every* sequence in the alignment are indicated with pink bars above the alignment.

a



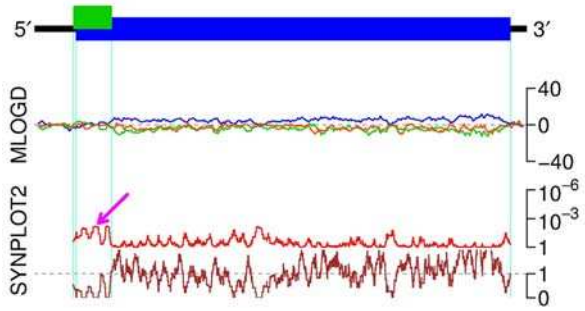
b

```
KX783422  -MATRLVRSTLSLVSQLPVPLLLFSSQSLLLVRLQEQYRDLHLLGFLLSLMVCLILLLLKSVRVQLLLIAFLRFLLVLLL
KX783421  -MAARLVRSTLSLVFQLPMPLLLFSSQSLLLVRLQEQYRDLHLLGFLLSLMACLILLLLRSVRVQLLLIAFLKFLLVLLL
KX783423  -MAARLVRSTLSLVSQLPMPLLLFSSQSLLLARLQEQYRDLHLLGFLLSLMVCLILLLLRSVRVQLLLIVFLRFLLVLLL
KX783431  MENALRTPLTTSPVCQLQVQLLLFSAQFLWLGKSLELSRDWLLLGSLLPLMVFLMIQALKLAMLQALLIGFLKWLLVQLL
KX783432  MENALRTPLTTSPVCQLQVQLLLFSAQFLWLGKSLELSRDWLLLGFLLPLMVSLMIQALKLAMLQALLIGFLKWLLVQLL
KX156156  MENVQRTLLMMSPVSRLQVQLLLFSAQFLWLGKSLELSRDWLLLGSLLPLMVCLMIQALKLEMCQGLLIGCLKWLLVLPL
KX783428  MENAQRTLLMMSPVCRLQVQLLLFSAQFLWLGKSLELSRDWLLLGFLLPLMVCLMIQVLKLAMFQALLIGFLKWLLVLPL
KX783425  MENVQRTLLMMSPVCRLQVQLLLFSAQFLWLGKSLELSRDWLLLGSLLPLMVCLMIQVLKLAMLQALLIGFLKWLLVLPL
KX783429  MENAQRILLMMSPVFQLQVQLLLFSAQFLWLGRSLELSRDWLLLGSLLPLMVSLTIQALKLAMFQALLIGCLKWLLVLPL
KX783430  MENAQRILLMMSPVFQLQVQLLLFSAQFLWLGRSLELSRDWLLLGSLLPLMVSLTIQALKLAMFQALLIGCLKWLLVLPL
KX783433  MENVQRTLLMMSPVSRLQVQLLLFSAQFLWLGKSLELSRDWLLLGSLLPLMGCLMIQALKLAMFQALLIGCLKWLLVLPL
KX783426  MENVQRTLLMMSPVSRLQVQLLLFSAQFLWLGKSLELSRDWLLLGSLLPLMGCLTIQALKLAMFQALLIGCLKWLLVLPL
KX783427  MENVQRTLLMMSPVSRLQVQLLLFSAQFLWLGKSLELSRDWLLLGSLLPLMGCLTIQALKLAMFQALLIGCLKWLLVLPL
KX783424  MENVQRTLLMMSPVSRLQVQLLLFSAQFLWLGKSLELSRDWLLLGSLLPLMGCLMIQALKLAMFQALLIGCLKWLLVLPL
            .      * * .* : *****:* * *  .  *  **  *** **.**  * :  *.    * ***  *.:***  *

KX783422  CLLRMLLLLEMGG-QAMARYLAPYLLP
KX783421  CLLRMLLLLEMGG-QVMAKYLAPYLLP
KX783423  CLLRMLLLLGMGG-QAMAKYLAPYLLP
KX783431  FLHRRLLQLDRDTLQTLTLFLMSPNPL
KX783432  FLHRRLLQLDRDTLQTLILFLMSPNLP
KX156156  YLLRRLLQLDRDTLQTLIQSLMIPSPP
KX783428  YLLRRLLQLDRDTLQTLTQSLMIPNPP
KX783425  YLLRRLLQLDRDTLQTLIQSLMIPNPP
KX783429  YLLRRLLQLDRDTQQILIQSLMIPNPL
KX783430  YLLRRLLQLDRDTQQILIQSLMIPNPL
KX783433  YLHRRLLQLDRDIQKILILFLMSPNPQ
KX783426  YLHRRLLQLDRDTQKILILFLMSPNPQ
KX783427  YLHRRLLQLDRDTQKILILFLMSPNPQ
KX783424  YLHRRLLQLDRDTQKILILFLMSPNPQ
           * * ** *  .  : :   *
```

Supplementary Figure 16 | Putative additional 5′ ORF in rosaviruses B and C. (a) Comparative genomic analysis of rosavirus B and C sequences from genus *Rosavirus* (see part B for accession numbers; reference sequence in red). Top – Genome map of reference sequence indicating the polyprotein ORF (blue) and the putative additional ORF (green). Middle – Coding potential in the three reading frames (see Fig. S14 for details). Bottom – Analysis of conservation at polyprotein-frame synonymous sites (see Fig. S14 for details). MLOGD is generally more appropriate for detecting non-overlapping ORFs whereas SYNPLOT2 is appropriate for detecting overlapping ORFs (pink arrow). **(b)** Amino acid alignment of the putative additional protein. Amino acids are color-coded according to their physicochemical properties (see Fig. S14). Transmembrane regions predicted by Phobius to be present in *every* sequence in the alignment are indicated with pink bars above the alignment.

a



b

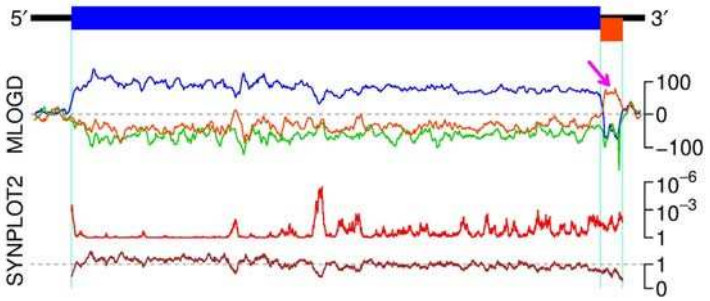
```
GU182410   TALWSVKKIIYRIIKWTLLDILLSTTSLLAEILLSGLRLIAACHLIPNHFPAGIQNLLYLLHLTAKPSTPWIPISLLARG
GU182411   TALWSVKKTIYRVIKWTLLDILLSTTSLLAETLLSGLRLIAACHLIPNHFPAGIQNLLYLLHLTAKPSTPWIPISLLARG
           ******** ***:***************** ***********************************************

GU182410   RLTRWSLEMVTHLRLTRVQMGLLRQFRSTQPFRVRLLLGLLAILQPCPNRRRPIRRRNRRLLMGKSLGTLLRNCSPTRRM
GU182411   RLIRWSLEMVTRLRLTRVQMGLLRQFRSTQPFRVRLLLGLLAILQPCPNRRRPIRRRNRRLLMGRSLGTLLRNYSLTHRM
           ** ********.*************************************************.******** * *.**

GU182410   RSLCSRLIEFLLSPLLVLLLLCNLIYHLCLLMRENLSTHH
GU182411   RSLCSRLIEFLLSPLLVLLLLCNLIYHLCLLMRENLSTHH
           ****************************************
```

Supplementary Figure 17 | Putative additional 5′ ORF in turdivirus 3. (a) Comparative genomic analysis of turdivirus 3 sequences from genus *Oscivirus* species *Oscivirus A2* (see part B for accession numbers; reference sequence in red). Top – Genome map of reference sequence indicating the polyprotein ORF (blue) and the putative additional ORF (green). Middle – Coding potential in the three reading frames (see Fig. S14 for details). Bottom – Analysis of conservation at synonymous sites (with the reading frame of the polyprotein being used where the two ORFs overlap) (see Fig. S14 for details). MLOGD is generally more appropriate for detecting non-overlapping ORFs whereas SYNPLOT2 is appropriate for detecting overlapping ORFs (pink arrow). **(b)** Amino acid alignment of the putative additional protein. Amino acids are color-coded according to their physicochemical properties (see Fig. S14). Transmembrane regions predicted by Phobius to be present in *every* sequence in the alignment are indicated with pink bars above the alignment. The ORF is predicted to start at an ACG codon (a known non-AUG initiator) [3] in a strong initiation context (A at −3, G at +4); while the ACG is expected to be decoded as Met by initiator Met-tRNA, for convenience the standard decoding (Thr) is shown in the above alignment.

a



b

Penguin, duck and goose megriviruses

```
MF405436   MGNTITTTIQQTIVNIYNYYSPSSDISIDSKQTSVSMVVVFSTISLLFLLTFIFLIYCFCKNRKSLAHDLKLSSRGESAI
KC663628   MGNTVTQTFNIVNNYIILNHSPNSQVSTGSSQSATSTVVVFSALGVFLFICIIFFSYCIFRNCRKEQKQLKFLG-----V
KY369299   MGNTITQTINQTIHNIFITNSPNSHVSSGSSQSASSSVVVFSILGVFIFFLVVFFSYCIFKNCRKEQKQLKLLE-----V
KY369300   MGNTITQTINQTIHNIFITNSPNSHVSSGSSQSASSSVVIFSVLGVFVFFLVVFFLYCIFKNCRKEQKQLKLLE-----V
           ****:* *:: .   *    **.* :* .*.*:: * **:** :::::: .:*: **: .* ..  ::**:      :
```

Chicken megriviruses

```
KF961187   MGNIASTNNYHIVNNLAMGTGATAAAVTFNTTMIVLITIGALVLVILCGVIIYYFWRKRRRSLRRKKRERKIFERVLMQR
KF961186   MGNIASTNNYHIVNNLAMGTGATAAAVTFNTTMIVLITIGALVLVILCGVIIYYFWRKRRRSLRRKKRERKIFERVLMQR
MG846464   MGNIASTNNYHIVNNLAMGTGATAAAVTFNTTMIILITIGALVLVILCGVIIYYFWRKRRRSLRRKKRERKIFERVLMQR
MG846465   MGNIASTNNYHIVNNLAMGTGATAAAVTFNTTMIVLITIGALVLVILCGVIIYYFWRRRRRSLRRKKREKKIFERVLMQR
MG846463   MGNIASTNNYHIVNNLAMGTGATAAAVTFNTTMIVLITIGALVLVILCGVIIYYFWRRRRRSLRRKKRERKIFERVLMQR
           **********************************:**********************.***********.**********

KF961187   ARLPPDHARAFAESIYAGNIRRSTTDEDGYEVVY
KF961186   AHLPPDHARAFAESIYAGNIRRSTTDEDGYEVVY
MG846464   ARLPPDHARAFAESVYAGNIRRSTTDEDGYEVMY
MG846465   AHLPPDHARAFAESVYAGNIRRSTTDEDGYEVMY
MG846463   AHLPPDHARAFAESVYAGNIRRSTTDEDGYEVMY
           *.************:****************:*
```

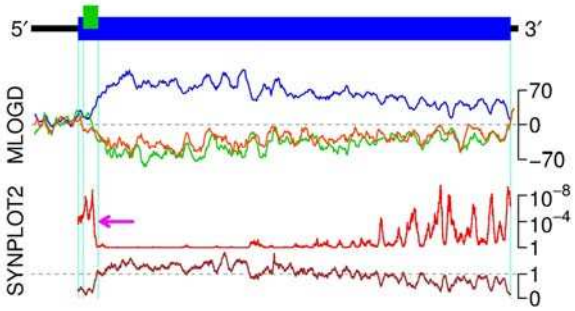
Turkey and chicken megriviruses

```
KF961188   MGITVSTINNINQIFATGSTVTSFTLQSASIVVMTIGVLVLIIVAGVVIYYFWKKHRRGRKSRRREEKILARLIEKGIVS
KF979336   MGITVSTINNINQIFAVGSTVTSFTLQSASIVVMSIGILVLVILAGVVIYYFWKKHRRGRGSRRREEKILARLIEKGLVS
KF979335   MGITVSTINNINQIFAVGSTVTSFTLQSASIIVMSIGVLVLVILAGVVIYYFWKKHRRGQSSRRREEKILARLIEKGLVS
           ***************.*************:**:**:***:*:***************. ****************:**

KF961188   PDTRLPDVEDEVIESIKL
KF979336   PDTRLPDYGNEVIESIKL
KF979335   PDTRLPDGGNEVIESIKL
           *******  :********
```

Supplementary Figure 18 | Putative additional 3′ ORF in megriviruses. (a) Comparative genomic analysis of genus *Megrivirus* sequences (see part B for accession numbers; reference sequence in red). Top – Genome map of reference sequence indicating the polyprotein ORF (blue) and the putative additional ORF (orange). Middle – Coding potential in the three reading frames (see Fig. S14 for details). Bottom – Analysis of conservation at synonymous sites (see Fig. S14 for details). MLOGD is generally more appropriate for detecting non-overlapping ORFs (pink arrow) whereas SYNPLOT2 is appropriate for detecting overlapping ORFs. **(b)** Amino acid alignment of the putative additional protein in three subgroups of megriviruses. Amino acids are color-coded according to their physicochemical properties (see Fig. S14). Transmembrane regions predicted by Phobius to be present in *every* sequence in the alignment are indicated with pink bars above the alignment. The 3′ ORF was previously noted in [4], and is likely expressed via a ribosomal reinitiation mechanism.

a

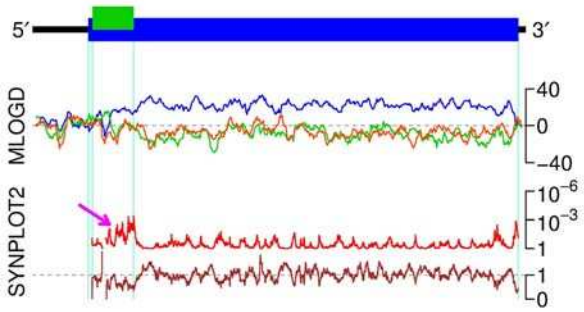


b

```
KT944214   TDLL-GKQIQQQAIWRQQVLIFL-LRTAGLRSLTSLKMSSALCFSMVRRQKNFSLSLKGLDRASRSQVIPITQETMV
KY432925   TDLF-GKQIQQQSIWRQQVSIFL-LRTAGLKXLTSLKTSSALCFSMVRRQKNFSLSLKGLDKVSRSQAIQTIQGTMV
KJ950971   TDLL-GKQIQQQSIWRQQVSIFL-LRTAGLRSLTSLKMSSALCFSMVRRQKNFSLSLKGLDKASRNQAIQTIQETMV
MF352430   TDLL-GKQIQQESIWRQQVLIFL-LRTAGLRSLTSLRMSSAQCFSMVRRQKNFSLNLKGLDKASRSQAIQTIQETMV
KT944212   TDLL-GKQIQQESIWRQQVLIFL-LRTAGLRSLTSLKMSSALCFSMVRRQKNFSLSLKGLDKANRNQATPIIQETMV
KT944213   TDLL-GKQIQQESIWRQQVLIFL-LRTAGLRSLTSLKMSSALCFSMVRRQKNFSLSLKGLDKANRNQATPIIQETMV
JQ941880   TDLLFTKQIQQQAIWRFQQWILICLVTSGMRRPTLNLMLSRLCFWMVRRQKTFSLSFRDLDKVSRNLATQTIQETME
KX156157   TELL-GKQIQQQSIWRLQSSILIWLTTSGMKTQTLSLMLSRLCFFMVRRQKTFSLSLKDLDRANRNLATPTIQATME
MF352413   TELL-GKQIQQQSIWRLQLSILIWLTTSGMKTQTLSLMLSRLCFYMVRRQKTFSLSLKDLDRANRNLAIPTIQATME
HM153767   TDLL-GKQIQQQSIWRLQSWILICLVTSGVRNPTLILMLSRLCFSLVRRQKNFSLSLKDLDRVSKNLETPITLATMA
           *:*:  *****::*** *  *:: * *:*:.  *     *  ** :*****.***.:..**.....       **
```

Supplementary Figure 19 | Putative additional 5′ ORF in hunniviruses. (a) Comparative genomic analysis of genus *Hunnivirus* sequences (see part B for accession numbers; reference sequence in red). Top – Genome map of reference sequence indicating the polyprotein ORF (blue) and the putative additional ORF (green). Middle – Coding potential in the three reading frames (see Fig. S14 for details). Bottom – Analysis of conservation at polyprotein-frame synonymous sites (see Fig. S14 for details). MLOGD is generally more appropriate for detecting non-overlapping ORFs whereas SYNPLOT2 is appropriate for detecting overlapping ORFs (pink arrow). **(b)** Amino acid alignment of the putative additional protein. Amino acids are color-coded according to their physicochemical properties (see Fig. S14). The ORF is predicted to start at an ACG codon (a known non-AUG initiator) [3] in a strong initiation context (A at −3, G at +4); while the ACG is expected to be decoded as Met by initiator Met-tRNA, for convenience the standard decoding (Thr) is shown in the above alignment.

a



b

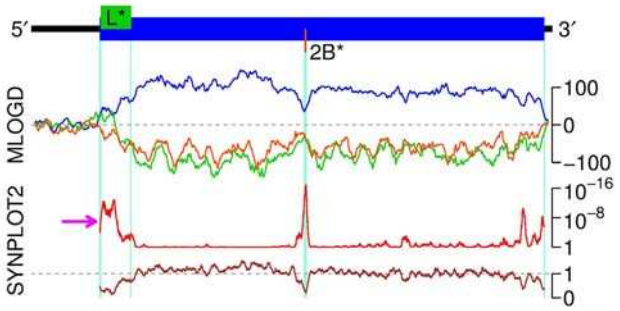
```
KM396707   MRLTMPSGLDLFF-TRVWRNPAPLTTTMTLGLRTQRAMILSGILTRPLHRPEYRLLQLSFQKRIVVLFPLLSHVVQSLLS
KM396708   MLLMQNSKKKIWFSTQTWTNRSHQTQRVTISPQTQRRLVLNGTISHPLLEQTFVILKLDFRRTAIVTLPLLSYVVRHFLK
           * *   *  .::* *..* * :  *  :*:. .*** ::*.* ::.**    : :*:*.*..  :* :****:**. :*.

KM396707   AELVGWLSRMFLLRPLRHSLNELYTWSFKARGHPGLTMETPRMMEIMESSITLFTTPTTKIPWTYRTLVALQALVMDMTV
KM396708   SELDTWLLRTLWLKPLRHSLEQLYPWNSREPEPPGWRTATPRMTEITESLITHFTTPTTRTPWTFLTLLGLRQQIMDMTV
           :**  ** * : *.******::**.*. .    **    **** ** ** ** ******. ***: **:.*.  :*****

KM396707   VLLTSTTPPGTGPTLLQPVLVQYHLSCPCWLIPKQRILRIQTGLTSTRRERPHS
KM396708   VIQTMTSPQESGLIYCQPVSVHFHSCCPCWLIPRQRMLRIQIESTSGRLERRHS
           *: * *:*  :*    *** *::* .******.**:****   ** * ** **
```

Supplementary Figure 20 │ Putative additional 5′ ORF in lesavirus 1 and 2. (a) Comparative genomic analysis of lesavirus 1 and 2 sequences (see part B for accession numbers; reference sequence in red). Top – Genome map of reference sequence indicating the polyprotein ORF (blue) and the putative additional ORF (green). Middle – Coding potential in the three reading frames (see Fig. S14 for details). Bottom – Analysis of conservation at polyprotein-frame synonymous sites (see Fig. S14 for details). MLOGD is generally more appropriate for detecting non-overlapping ORFs whereas SYNPLOT2 is appropriate for detecting overlapping ORFs (pink arrow). **(b)** Amino acid alignment of the putative additional protein. Amino acids are color-coded according to their physicochemical properties (see Fig. S14).

a



b

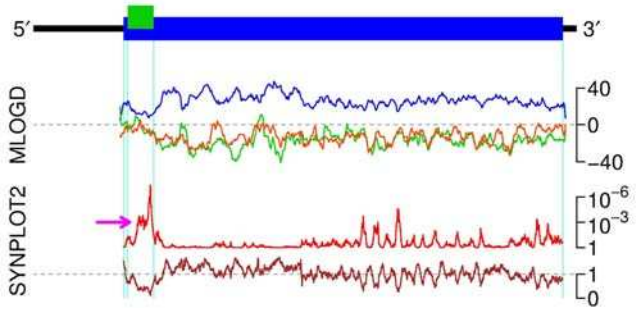
```
MF352428   MDIPMCVLGAQPLTLSLTTTTFCWQTENGSPKIFSVLTSTMMFSGLRLLVIRTTRWNGMIPWTTVMLSWNPREIPTPLIR
MF352411   MDTQACVLFAQPLTKVPTECICSWQITNGIQRTFLLWTWTMTSSGLTTRAWCLRQWTGLIFRSYTILSWNPRETPRHLTR
MF352420   MDTQTCVLYAQPLTLLLTLNTCSWQTENGSQRTFFVLTWMMTSSGLRIRAIALRQWTGLTYRAYSILSWNPRETPRHLTR
EU815052   MDTQACVLFAQPLTKVPTECICSWQITNGSQRIFLLWTWMMTSSGLMTRAMCLRQWTGLTFRSYSILSWNPRETPRHLTR
AB090161   MDTQACVLFAQPLTKVPTECICSWQITNGSQRIFLLWTWMMTSSGLMTRAMCLRQWTGLTFRSYSILSWNPRETPRHLTR
EU542581   MDTQACVLFAQPLTKVPTECICSWQITNGSQRIFLLWTWMMTSSGLMTRAMCLRQWTGLTFRSYSILSWNPRETPRHLTR
MF172923   MDIQTCVLYAQPLTLLPTLTICSWQTVNGSRRTFFVLTWMMTSSGLRIRALNLKQWNGLTYRSYAILSWNPRETPRHLTR
JX443418   MDTQMCALFAQPLTLLPDLNICSWQTVNGSQRTFFVWTWTMTSSGLRTRAINLKQWNGLTYRSYAILSWNPRETPLHLTR
M20301     MDTQMCALFAQPLTLLPDLNICSWQTVNGSQRTFFVWTWTMTSSGLRTRAINLKQWNGLTYRSYAILSWNPRETPLHLTR
EU718733   MDTQMCVPFAQPLTLLPALNICSWRTENGSLRTFFVWTWTMTSSGLRTRAINLKQWNGLTYRSYAILSWNPRETPLHLTR
KJ191558   MDTQTCALFAQPLTLLPALNICSWRTENGSQRTFFVWTWTMTSSGLRTRAINLKQWNGLTYRSYAILSWNPREIPHHLTR
MF416403   MDTQTCVLFAQPLTLLPTLNICSWQTENGSLRTFFVWTWTMTSSGLRTRAINLKQWNGLTYRSYAILSWNPRETPRHLTR
HQ652539   MDTQMCAPFAQPLTLLPDLNICSWQTVNGSQRTFFVWTWTMTSSGLRTRAINLKQWNGLTYRSYAILSWNPRETPLHLTR
M20562     TDTQTCALFAQPLTLLPTLNICSWQTENGSLRTFFVWTWTMTSSGLRTRALNLKQWNGLMYRSYAILSWNPREMPRHLIR
X56019     TDTQTCALFAQPLTLLPTLNICSWQTENGSLRTFFVWTWTMTSSGLRTRALNLKQWNGLMYRSYAILSWNPREMPRHLIR
EU718732   MDTQTCALFAQPLTLLPTLNICSWQTENGSLRTFFVWTWTMTSSGLRTRVINFKQWTGLTYRSYAILSWNPREMPRHLTR
DQ401688   MDTQMCALFAQPLTLLPALNICSWQTENGTLRTFFVWTWTMTSSGLRTRAINLKQWTGLTYRSYAILSWNPRETPRHLTR
M16020     MDTQMCALFAQPLTLLPALNICSWQTENGTLRTFFVWTWTMTSSGLRTRAINLKQWTGLTYRSYAILSWNPRETPRHLTR
EU723238   MDTQMCALFAQPLTLLPTLNICSWQTENGSLRTFFVWTWTMTSSGLRTRAINLKQWTGLTYRSYAILSWNPRETPRHLTR
            *   *.  *****        .*.  **  . * : *  *  ***   .    .*.*:  .:  :******* *  * *

MF352428   TVLKPPGMRALLSTTTTQTNTRTQLISQQTTAMPVVEPAPQGNWQTYLVILQMHFLLLAPYYLTKIQKKWKIFQTV
MF352411   TTLSPLGMRVLSLTITIPTSTKIQSTFLPMLVAPETVALALNQRVSSATSLVVLLTL-----------SRILLHS
MF352420   VTPSLQVMKELLLTTSIPISTKIQLTSLPTAATPAALQRQKDNWRAFWVMLQMHSLLWLPFFLIKIQRKWKIFQTA
EU815052   VTPSLQAMKELLLTTSIPISTKIQLTSLPTAGTPVALLKQKDNWGTYLVMLQMHFPLWLLCFLTKTQRKWKIFQTA
AB090161   VTPSLQAMKELLLTTSIPISTKIQLTSPPTAETPAALLRQKDNWGTYLVMLQMHFPLWLLYFLIKIQRKWKIFQTA
EU542581   VTPSLQAMKELLLTTSIPISTKIQLTSLPTVETPAALLKQKDNWGTYLVMLQMHFPLWLLYFLTKIQRKWKIFQIA
MF172923   VTPSLQEMKGLSLITSIPINTKTQLIFLPAAATPAELPRLKDSCPTFWVALLTPLLPWLLSSLIKTPRRWRISLTE
JX443418   VTPSPQVTKGSLLTTSIPINTKIQLICLPAVAMLATPPKTTDNCRTSWAALQMLLLLWHLSSWIKTQRRWRISLTE
M20301     VTPSPQVTKGSLSTTSIPINTKIQLICLPAVAMLATPPKTTDNCRTSWAALQMLLLLWHLSSWIKTQRRWRISLTE
EU718733   ATPSPREMKGLLLTTSIPISTKIQLICLPVVAMQAMLPKPTDNCQTSWAALQMLLLLWHLSSWTKIQRRWRISLTE
KJ191558   VTPSPQEMKGLLLTTSIPISTKIQLICLPAVAMQVTLPKPTDNCPTSWAVLQMLLLLWHLSSWTKTQRRWRISLTE
MF416403   VTPSPQEMRGLLLTTSIPINTKIQLICLPVVAALAMPPRPKDNCPTSWAELLMPLLLWHLSSWIRTQRKWRISLTE
HQ652539   VTPSPQEMRGLLLTTSIPINTKTQLICLPVVATLAMLPKTTDSCPTFWVELQMLLLLWHLSSWTKTRRRWKTFLTE
M20562     VTPSPQEMRGLSLITSIPINTRTQLICLPVVATLAMLPRTMDNCPAFWVELQMLLLLWHLSSWTRTQRRWKTSLTE
X56019     VTPSPQEMRGLSLITSIPINTRTQLICLPVVATLAMLPRTMDNCPAFWVELQMLLLLWHLSSWTRTQRRWKTSLTE
EU718732   VTPSPQEMKGLLLTTSIPINTKTQLICLPVVATLAMLPRTADNCPTSWVELQMLLLLWHLSSWTKTQRRWKTSLTE
DQ401688   VTPSPQEMKGLLLTTSIPINTKTQLICLPVVATLAMLPRPMDNCPTSWVELQMLLLLWHLSFWTKTQRRWKTSLTE
M16020     VTPSPQEMKGLLLTTSIPINTKTQLICLPVVATLAMLPRPMDNCPTSWVELQMLLLLWHLSFWTKTQRRWKTSLTE
EU723238   VTPSSQEMKGLLLTTSIPINTRTQLICLPVVATLAMLPRTMDNCPTSWVELQMLLLLWHLSSWTKTQRRWKTSLTE
           .. .   .      :   .*. *                  ..  :  . *                    .
```

Supplementary Figure 21 | The established L* ORF in Theiler's murine encephalomyelitis virus and relatives. (a) Comparative genomic analysis of Theiler's murine encephalomyelitis virus, rat theilovirus and related sequences from genus *Cardiovirus* species *Cardiovirus B* (see part B for accession numbers; reference sequence in red). Top – Genome map of reference sequence indicating the polyprotein ORF (blue) and the L* ORF (green) [5,6]. Middle – Coding potential in the three reading frames (see Fig. S14 for details). Bottom – Analysis of conservation at polyprotein-frame synonymous sites (see Fig. S14 for details). MLOGD is generally more appropriate for detecting non-overlapping ORFs whereas SYNPLOT2 is appropriate for detecting overlapping ORFs (pink arrow). **(b)** Amino acid alignment of the L* protein. Amino acids are color-coded according to their physicochemical properties (see Fig. S14). In two sequences, the ORF starts at an ACG codon in a strong initiation context (A at −3, G at +4) [6]; while

the ACG is expected to be decoded as Met by initiator Met-tRNA, for convenience the standard decoding (Thr) is shown in the above alignment. Note that cardioviruses have another additional ORF – 2B* (orange) – that is accessed via a −1 ribosomal frameshift [1,7].

a



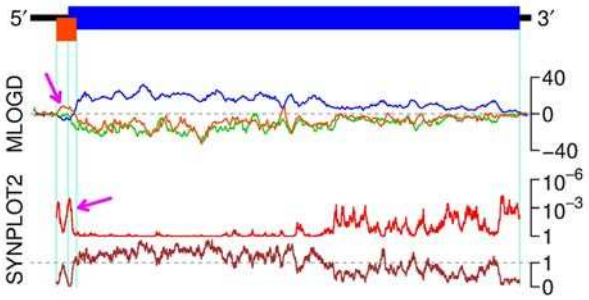
b

```
JQ814851   LVLTH-QSLLRMFLLHPSPLRLAKGYIYQALAVVLPMNWTRNLRKDSQDLLVPERWFLWCHLTFPSTTSRTRVVILLNRQ
KP054278   LVLIHLQCLKRLTHYHHSHQRKHASSIFPVLALVLPSHWMRNLKNQESSLQIQERWFLWFHLRFPSTTSRTRVAILLSHP
KP054277   LVLIHLQCLKNLTHLHLFRPRTSINSIFHVLARALPLPWMRSLKNQESSLQIQEKWFLWFHLKFPSMTSRTRVAILLSHP
KP054274   LVLIHLQCLKNLTHLHLFRPRTSINSIFHVLARALPLPWMRSLKNQESSLQIQEKWFLWFHLKFPSMTSRTRVVILLSHP
KP054275   LVLIHLQCLKNLTHLHLFRPRTSINSIFHVLARALPLPWMRSLKNQESSLQIQEKWFLWFHLKFPSTTSRTRVVILLSHP
KP054273   LVLIHLQCLKNLTHLHLFRPRTSINSIFHVLARALPLPWMRSLKNQESSLQIQEKWFLWFHLKFPSMTSRTRVVILLSHP
KP054276   LVLIHLQCLKNLTHLHLFRPRTSINSIFHVLARALPLPWMRSLKNQESSLQIQEKWFLWFHLKFPSMTSRTRVVILLSHP
                        *   *   . *: .** .**  * ..*..:...* : *.**** **  ** ******.***..

JQ814851   HLAITMKVPLSTTSMLITIMDPSTQLAQPLAQAGALSLLWVPSLVLLVLWLEPY---------------
KP054278   HLLITMKVSSSTTIMPTTIMVPSMQLAQPLVQAGALSLLWVPSLVLLDLWLDHFSWTKKLRRRLIYQTEY
KP054277   HLLITMKVSSSTTIMPTIIMVPSMQPAQPLVQAGALSLLWVPSLVLLDLWLDHCSWTKTLRRRLIYQTES
KP054274   HLLITMKVSSSTTIMPTIIMVPSMQPAQPLVQAGALSLLWVPSLVLLDLWLDHCSWTKTLRRRLIYQTES
KP054275   HLLITMKVSSSTTIMPTIIMVPSMQLAQPLVQAGALSLLWVPSLVLLDLWLDHCSWTKTLRRRLIYQTES
KP054273   HLLITMKVSSSTTIMPTIIMVPSMQLAQPLVQAGALSLLWVPSLVLLDLWLDHCLWTKTLRRRLIYQTES
KP054276   HLLITMKVSSSTTIMPTIIMVPSMQLAQPLVQAGALSLLWVPSLVLLDLWLDHCSWTKTLRRRLIYQTES
           ** *****. **  *   ** ** * ****.****** ********* ***:
```

Supplementary Figure 22 | Putative additional 5′ ORF in Miniopterus schreibersii picornavirus 1 and relatives. (a) Comparative genomic analysis of Miniopterus schreibersii picornavirus 1 and related sequences from genus *Mischivirus* (see part B for accession numbers; reference sequence in red). Top – Genome map of reference sequence indicating the polyprotein ORF (blue) and the putative additional ORF (green). Middle – Coding potential in the three reading frames (see Fig. S14 for details). Bottom – Analysis of conservation at polyprotein-frame synonymous sites (see Fig. S14 for details). MLOGD is generally more appropriate for detecting non-overlapping ORFs whereas SYNPLOT2 is appropriate for detecting overlapping ORFs (pink arrow). **(b)** Amino acid alignment of the putative additional protein. Amino acids are color-coded according to their physicochemical properties (see Fig. S14). Transmembrane regions predicted by Phobius to be present in *every* sequence in the alignment are indicated with pink bars above the alignment. The ORF is predicted to start at a CUG codon (a known non-AUG initiator) [3] in a strong initiation context (A or G at −3, G at +4); while the CUG is expected to be decoded as Met by initiator Met-tRNA, for convenience the standard decoding (Leu) is shown in the above alignment.

a



b

```
FR727144   MASTRFCCRFGMSHVGDVVTGVCFTCARIRKQSLEFWKLMDNIKFPLRSSCILGQKIDHKIENGVCKTCDARFWRDQCCK
KC560801   MASTRFCCRFGMSHVGDVVTGVCFTCARIRKQSLEFWKLMDNIKFPLRSSCILGQKIDHKIENGVCKTCDARFWRDQCCK
           ********************************************************************************

FR727144   RYSTFYDQLPKGETYSTRDILQGT
KC560801   RYSTFYDQLPKGEAHPPCDVLQGA
           *************:::: *:***:
```

Supplementary Figure 23 | Putative additional 5′ ORF in pigeon picornavirus B. (a) Comparative genomic analysis of pigeon picornavirus B sequences (see part B for accession numbers; reference sequence in red). Top – Genome map of reference sequence indicating the polyprotein ORF (blue) and the putative additional ORF (orange). Middle – Coding potential in the three reading frames (see Fig. S14 for details). Bottom – Analysis of conservation at synonymous sites (with the reading frame of the polyprotein being used where the two ORFs overlap) (see Fig. S14 for details). MLOGD and SYNPLOT2 detect, respectively, the non-overlapping and the overlapping portions of the additional ORF (pink arrows). **(b)** Amino acid alignment of the putative additional protein. Amino acids are color-coded according to their physicochemical properties (see Fig. S14).

Supplementary Table 1 │ uORF and UP statistics in different enterovirus groups

species	*n*	median length (aa)	median mass (kDa)	median pI	dN/dS (mean ± s.d.)	median % overlap with main ORF	frame of overlap with main ORF (+1/+2)
Enterovirus A	1182	64	7.3	9.5	0.14 ± 0.03	27	431/751
Enterovirus B	357	67	8.1	11.2	0.12 ± 0.03	26	357/0
Enterovirus C	345	65	7.5	8.5	0.22 ± 0.04	23	16/329
Enterovirus E	9	76	9.0	9.2	0.08 ± 0.03	32	1/8
Enterovirus F	11	56	6.5	8.9	0.04 ± 0.02	13	11/0
Enterovirus G	16	67	7.9	9.1	0.09 ± 0.03	10	16/0

Supplementary Table 2 │ Host and virus read counts for different Ribo-Seq samples

virus	repeat	time point	total reads	host rRNA	host mRNA	vRNA
EV7						
	1	4 hpi	9,316,308	4,508,754	1,751,701	2,452
	2	4 hpi	9,843,860	5,706,711	1,545,561	4,273
	1	6 hpi	9,941,585	3,644,716	1,821,542	401,949
	2	6 hpi	10,626,765	4,074,181	2,641,240	459,564
PV1						
	1	4 hpi	20,899,214	13,652,594	992,512	944,353
	2	4 hpi	19,753,522	13,452,101	697,185	683,735
	1	6 hpi	8,400,577	4,225,244	223,021	178,996
	2	6 hpi	11,745,343	6,431,870	543,444	411,096
EV-A71						
	1	5 hpi	19,901,615	10,178,049	2,618,450	320,634
	2	5 hpi	19,164,709	8,252,591	2,647,234	303,239
	1	7.5 hpi	12,172,109	3,836,010	1,473,426	310,292
	2	7.5 hpi	13,629,131	3,968,251	1,820,243	338,542

Supplementary Table 3 │ Statistical analysis of data from organoid-derived samples
Please see the accompanying Supplementary Table S3 excel file.

Supplementary Table 4 │ Statistical analysis of dual luciferase data
Please see the accompanying Supplementary Table S4 excel file.

References

1. Napthine, S. *et al.* Protein-directed ribosomal frameshifting temporally regulates gene expression. *Nat. Commun.* **8,** 15582 (2017).

2. Firth, A. E. Mapping overlapping functional elements embedded within the protein-coding regions of RNA viruses. *Nucleic Acids Res.* **42,** 12425–39 (2014).

3. Firth, A. E. & Brierley, I. Non-canonical translation in RNA viruses. *J. Gen. Virol.* **93,** 1385–1409 (2012).

4. Boros, Á. *et al.* Comparative complete genome analysis of chicken and Turkey megriviruses (family picornaviridae): long 3' untranslated regions with a potential second open reading frame and evidence for possible recombination. *J. Virol.* **88,** 6434–43 (2014).

5. Roos, R. P., Kong, W. P. & Semler, B. L. Polyprotein processing of Theiler's murine encephalomyelitis virus. *J. Virol.* **63,** 5344–53 (1989).

6. van Eyll, O. & Michiels, T. Non-AUG-initiated internal translation of the L* protein of Theiler's virus and importance of this protein for viral persistence. *J. Virol.* **76,** 10665–73 (2002).

7. Loughran, G., Firth, A. E. & Atkins, J. F. Ribosomal frameshifting into an overlapping gene in the 2B-encoding region of the cardiovirus genome. *Proc. Natl. Acad. Sci. U. S. A.* **108,** E1111-9 (2011).

Supplementary Table 3 | Statistical analysis of data from organoid-derived samples

All t-tests are two-tailed and assume separate variances for the two populations being compared.

TI200 - patient#1, clarified supernatants, growth curve timepoints titers

hpi	EV7	PTC	Loop	**T-test**	(*n* = 3 , 6)
0	2.6E+04	2.7E+04	2.5E+04		
	2.4E+04	2.6E+04	2.2E+04		
	2.5E+04	2.4E+04	1.8E+04	0.3917	
3	2.4E+04	1.8E+04	2.3E+04		
	2.5E+04	2.1E+04	2.2E+04		
	2.4E+04	1.9E+04	2.2E+04	0.0057	
6	8.4E+03	1.6E+04	1.1E+04		
	1.2E+04	1.4E+04	8.4E+03		
	9.2E+03	1.7E+04	1.1E+04	0.1272	
9	4.0E+05	5.2E+05	4.0E+05		
	3.2E+05	5.2E+05	8.0E+05		
	2.8E+05	4.4E+05	6.8E+05	0.0157	
12	1.2E+06	4.0E+06	6.4E+06		
	2.0E+06	3.6E+06	3.6E+06		
	3.2E+06	4.4E+06	5.6E+06	0.0237	
24	3.6E+07	2.0E+07	2.0E+07		
	4.0E+07	2.0E+07	2.8E+07		
	4.0E+07	2.0E+07	4.0E+07	0.0072	
36	6.4E+06	2.4E+06	1.2E+06		
	7.8E+06	1.6E+06	1.6E+06		
	1.2E+07	2.4E+06	1.6E+06	0.0523	
48	1.4E+06	3.0E+05	2.0E+05		
	1.6E+06	2.0E+05	2.0E+05		
	1.6E+06	2.0E+05	3.0E+05	0.0012	

TI365 - patient#2, cla

hpi	EV7
0	2.8E+04
	6.4E+04
	4.0E+04
3	3.6E+04
	4.0E+04
	2.4E+04
6	4.8E+04
	6.6E+04
	3.8E+04
9	6.8E+05
	5.6E+05
	1.0E+05
12	2.8E+06
	2.4E+06
	4.0E+06
24	2.0E+07
	1.6E+07
	3.2E+07
36	4.0E+07
	3.6E+07
	3.6E+07
48	3.6E+06
	6.4E+06
	4.4E+06

TI200 - patient#1, clarified supernatant derived virus

36h	#	Titer	Protein	Titer/prot	**T/P norm**	**T-test**	RNA	Titer/RNA
EV7	1	6.4E+06	483	13251	**0.800**	**0.0266**	0.0981	6.53E+07
	2	7.8E+06	494	15789	**0.953**	(*n* = 3 , 6)	0.1016	7.67E+07
	3	1.2E+07	581	20654	**1.247**		0.1292	9.29E+07
PTC	1	2.4E+06	347	6916	**0.418**		0.1060	2.26E+07
	2	1.6E+06	331	4834	**0.292**		0.0725	2.21E+07
	3	2.4E+06	363	6612	**0.399**		0.0865	2.77E+07
Loop	1	1.2E+06	450	2667	**0.161**		0.0855	1.40E+07
	2	1.6E+06	343	4665	**0.282**		0.0975	1.64E+07
	3	1.6E+06	306	5229	**0.316**		0.0863	1.85E+07
48h	#	Titer	Protein	Titer/prot	**T/P norm**	**T-test**	RNA	Titer/RNA
EV7	1	1.4E+06	96.3	14538	**0.943**	**0.0086**	0.0705	1.99E+07
	2	1.6E+06	87.6	18265	**1.185**	(n = 3 , 6)	0.0736	2.17E+07
	3	1.6E+06	119	13445	**0.872**		0.0897	1.78E+07
PTC	1	3.0E+05	91.5	3279	**0.213**		0.0664	4.52E+06
	2	2.0E+05	65.9	3035	**0.197**		0.0636	3.15E+06
	3	2.0E+05	74.4	2688	**0.174**		0.0639	3.13E+06

Loop	1	2.0E+05	90.3	2215	**0.144**		0.0493	4.06E+06
	2	2.0E+05	49	4082	**0.265**		0.0567	3.53E+06
	3	3.0E+05	53.8	5576	**0.362**		0.0344	8.72E+06

36h wt avg	16565	7.83E+07
48h wt avg	15416	1.98E+07

TI365 - patient#2, clarified supernatant derived virus

36h	#	Titer	Protein	Titer/prot	**T/P norm**	**T-test**	RNA	Titer/RNA
EV7	1	4.0E+07	249	160643	**0.917**	**0.0083**	1.2468	32083218
	2	3.6E+07	175	205714	**1.174**	(n = 3 , 6)	0.7923	45435892
	3	3.6E+07	226	159292	**0.909**		0.9599	37502508
PTC	1	3.2E+06	206	15534	**0.089**		0.8341	3836265
	2	4.8E+06	165	29091	**0.166**		0.5672	8462655
	3	4.0E+06	167	23952	**0.137**		0.8740	4576807
Loop	1	4.8E+06	161	29814	**0.170**		0.9521	5041687
	2	5.6E+06	194	28866	**0.165**		0.6345	8825748
	3	4.0E+06	260	15385	**0.088**		0.7034	5686656
48h	#	Titer	Protein	Titer/prot	**T/P norm**	**T-test**	RNA	Titer/RNA
EV7	1	3.6E+06	203	17734	**0.924**	**0.0020**	1.0571	3405629
	2	6.4E+06	293	21843	**1.137**	(n = 3 , 6)	0.9772	6549220
	3	4.4E+06	244	18033	**0.939**		0.9688	4541490
PTC	1	1.4E+06	228	6140	**0.320**		0.8040	1741238
	2	1.4E+06	176	7955	**0.414**		0.7734	1810172
	3	1.2E+06	186	6452	**0.336**		0.7324	1638412
Loop	1	1.2E+06	178	6742	**0.351**		0.7472	1605942
	2	6.8E+05	214	3178	**0.165**		0.7396	919420.8
	3	1.0E+06	258	3876	**0.202**		0.7821	1278595

36h wt avg	175216	4E+07
48h wt avg	19203	5E+06

Neutralisation assay, % of membrane-associated virus not neutralised by EV7 neutralisation serum an

	Virus	Neutralization serum (NS)				Average	St. Dev	**T-test**
Patient 1	EV7	14.4	9.4	10.0	9.0	10.7	2.5	**0.0057**
	PTC	13.0	10.7	10.0	10.0	10.9	1.4	**0.0008**
	Loop	13.3	10.8	10.0	10.0	11.0	1.6	**0.0011**
Patient 2	EV7	16.0	10.0	12.5	14.0	13.1	2.5	**0.0029**
	PTC	13.3	12.5	8.8	10.0	11.2	2.1	**0.0029**
	Loop	13.3	11.9	8.0	10.0	10.8	2.3	**0.0043**
EV7(RD)		2.2	1.6	2	2	2.0	0.3	
	Virus	anti-DAF antibody				Average	St. Dev	**T-test**
Patient 1	EV7	32.8	29.4	32.9	29.0	31.0	2.1	**0.00003**
	PTC	33.0	23.0	50.0	55.0	40.3	14.9	**0.0207**
	Loop	26.7	26.7	43.8	25.0	30.6	8.9	**0.0127**
Patient 2	EV7	50.0	50.0	47.5	45.0	48.1	2.4	**0.00001**
	PTC	40.0	47.5	50.0	41.0	44.6	4.9	**0.0004**
	Loop	56.6	52.3	42.7	36.0	46.9	9.3	**0.0032**
EV7(RD)		7.8	6.2	8	6.4	7.1	0.9	
	Virus	NS + anti-DAF antibody				Average	St. Dev	**T-test**

Patient 1	EV7	13.6	8.2	10	8	10.0	2.6	**0.0056**
	PTC	13	9.3	8	11.7	10.5	2.3	**0.0032**
	Loop	10	9.2	10	8.1	9.3	0.9	**0.0002**
Patient 2	EV7	18	10	10	7.5	11.4	4.6	**0.01859**
	PTC	13.3	12.5	8	7.5	10.2	3.1	**0.0090**
	Loop	16.6	9.5	7	8.0	10.2	4.4	**0.0232**
EV7(RD)		0.6	0.6	0.8	0.9	0.7	0.2	

T-test (RD neutralization serum vs. NS + anti-DAF antibo	**0.000444**
T-test (RD anti-DAF antibody vs. NS + anti-DAF antibody	**0.000672**

arified supernatants, growth curve timepoints titers

PTC	Loop	**T-test**	(*n* = 3 , 6)
1.6E+04	3.2E+04		
3.2E+04	2.0E+04		
3.6E+04	6.0E+04	0.4167	
1.2E+04	3.6E+04		
2.0E+04	2.8E+04		
1.6E+04	2.8E+04	0.1664	
2.1E+05	2.2E+05		
2.1E+05	2.4E+05		
2.6E+05	3.3E+05	0.0000	
2.0E+06	2.0E+06		
2.3E+06	1.7E+06		
2.4E+06	1.9E+06	0.0025	
2.4E+06	4.0E+06		
4.4E+06	2.8E+06		
2.0E+06	2.0E+06	0.8430	
1.1E+07	7.2E+06		
6.4E+06	4.4E+06		
9.2E+06	6.4E+06	0.0815	
3.2E+06	4.8E+06		
4.8E+06	5.6E+06		
4.0E+06	4.0E+06	0.00089	
1.4E+06	1.2E+06		
1.4E+06	6.8E+05		
1.2E+06	1.0E+06	0.0460	

TI200 - patient#1

T/R norm	**T-test**	36h	#	TiterNT	TiterTX	**Fold**	**T-test**
0.833	**0.0142**		1	5.6E+06	7.2E+06	**1.29**	**0.0194**
0.980	(n = 3 , 6)	EV7	2	4.2E+06	4.8E+06	**1.14**	(n = 3 , 6)
1.187			3	4.8E+06	5.6E+06	**1.17**	
0.289			1	2.0E+06	3.6E+06	**1.80**	
0.282		PTC	2	5.6E+05	9.6E+05	**1.71**	
0.354			3	2.0E+06	3.6E+06	**1.80**	
0.179			1	1.2E+06	5.2E+06	**4.33**	
0.210		Loop	2	8.0E+05	2.8E+06	**3.50**	
0.237			3	8.0E+05	2.8E+06	**3.50**	
T/R norm	**T-test**	48h	#	TiterNT	TiterTX	**Fold**	**T-test**
1.003	**0.00024**		1	9.0E+05	1.2E+06	**1.33**	**0.0059**
1.097	(n = 3 , 6)	EV7	2	1.6E+06	1.6E+06	**1.00**	(n = 3 , 6)
0.900			3	1.2E+06	1.2E+06	**1.00**	
0.228			1	2.0E+05	6.0E+05	**3.00**	
0.159		PTC	2	1.6E+05	3.2E+05	**2.00**	
0.158			3	3.2E+05	5.6E+05	**1.75**	

0.205		Loop	1	2.8E+05	6.0E+05	**2.14**	
0.178			2	1.2E+05	3.2E+05	**2.67**	
0.440			3	2.0E+05	8.0E+05	**4.00**	

NT - not treated
TX - Triton X100 treated

TI365 - patient#2

T/R norm	**T-test**	36h	#	TiterNT	TiterTX	**Fold**	**T-test**
0.837	**0.0112**	EV7	1	1.9E+07	1.9E+07	**1.00**	**0.0448**
1.185	(n = 3 , 6)		2	1.1E+07	1.2E+07	**1.09**	(n = 3 , 6)
0.978			3	1.6E+07	1.8E+07	**1.13**	
0.100		PTC	1	4.0E+06	5.6E+06	**1.40**	
0.221			2	4.4E+06	9.2E+06	**2.09**	
0.119			3	6.0E+06	1.6E+07	**2.67**	
0.131		Loop	1	5.2E+06	7.2E+06	**1.38**	
0.230			2	4.8E+06	1.4E+07	**2.92**	
0.148			3	8.0E+06	8.0E+06	**1.00**	
T/R norm	**T-test**	48h	#	TiterNT	TiterTX	**Fold**	**T-test**
0.705	**0.0653**	EV7	1	3.0E+06	4.4E+06	**1.47**	**0.0095**
1.355	(n = 3 , 6)		2	5.6E+06	5.2E+06	**0.93**	(n = 3 , 6)
0.940			3	4.8E+06	6.4E+06	**1.33**	
0.360		PTC	1	6.4E+05	1.2E+06	**1.88**	
0.375			2	1.4E+06	2.0E+06	**1.43**	
0.339			3	7.2E+05	1.5E+06	**2.08**	
0.332		Loop	1	1.1E+06	3.2E+06	**2.91**	
0.190			2	8.4E+05	2.4E+06	**2.86**	
0.265			3	8.4E+05	2.0E+06	**2.38**	

NT - not treated
TX - Triton X100 treated

d/or anti-DAF antibody

Comb. T-test	
4.79E-10	(n = 12 , 4)
1.05E-08	(n = 12 , 4)
Comb. T-test	
1.53E-06	(n = 12 , 4)
1.69E-11	(n = 12 , 4)
Comb. T-test	

2.80E-09	(n = 12 , 4)
1.75E-06	(n = 12 , 4)

(n = 4 , 4)
(n = 4 , 4)

TI200 - patient#1, lysed cell derived virus, 48hpi

48h	#	TiterNT	TiterTX	**Fold**	**T-test**
EV7	1	2.0E+06	4.0E+06	**2.00**	**0.0146**
	2	5.2E+06	9.6E+06	**1.85**	(n = 3 , 6)
	3	2.0E+06	4.4E+06	**2.20**	
PTC	1	8.0E+05	2.8E+06	**3.50**	
	2	5.6E+05	3.6E+06	**6.43**	
	3	9.2E+05	3.2E+06	**3.48**	
Loop	1	1.2E+06	4.4E+06	**3.67**	
	2	8.0E+05	4.8E+06	**6.00**	
	3	4.4E+05	4.0E+06	**9.09**	

TI200 - patient#1, clarified supernatant derived virus used in flotation

36h	TiterNT	TiterTX	**Fold**	Fract. 1-3	Fract. 8-12	**(8-12)/(1-3)**	norm wt
EV7	4.60E+06	5.40E+06	**1.17**	1.60E+04	1.73E+04	**1.08**	**1.0**
PTC	1.20E+06	5.20E+06	**4.33**	2.94E+04	1.58E+05	**5.37**	**5.0**
Loop	1.60E+06	8.80E+06	**5.50**	3.56E+04	7.30E+04	**2.05**	**1.9**

TI365 - patient#2, lysed cell derived virus, 48hpi

48h	#	TiterNT	TiterTX	**Fold**	**T-test**
EV7	1	4.8E+06	9.6E+06	**2.00**	**0.0059**
	2	7.6E+06	1.4E+07	**1.84**	(n = 3 , 6)
	3	4.8E+06	9.6E+06	**2.00**	
PTC	1	8.0E+05	6.2E+06	**7.75**	
	2	1.6E+06	6.4E+06	**4.00**	
	3	2.8E+06	9.6E+06	**3.43**	
Loop	1	2.4E+06	9.4E+06	**3.92**	
	2	2.0E+06	1.2E+07	**6.00**	
	3	2.0E+06	9.6E+06	**4.80**	

TI365 - patient#2, clarified supernatant derived virus used in flotation

36h	TiterNT	TiterTX	**Fold**	Fract. 1-3	Fract. 8-12	**(8-12)/(1-3)**	norm wt
EV7	7.20E+06	8.80E+06	**1.22**	1.43E+04	2.40E+04	**1.68**	**1.0**
PTC	8.00E+05	3.20E+06	**4.00**	3.84E+03	1.80E+04	**4.69**	**2.8**
Loop	7.60E+05	2.50E+06	**3.29**	1.32E+03	6.44E+03	**4.88**	**2.9**

T-test (membrane-bound to free virus in flotation assay)					
	patient 1	patient 2	**Fold change**	**T-test**	
EV7	1.08	1.68	3.1	0.026	(n = 2 , 4)
PTC	5.37	4.69			
Loop	2.05	4.88			

T-test
0.0428

Luciferase measurement data from Figure 2d

	EV7-wt	Loop	PTC	mAUG	StrUP	StrUP-PTC	2A-stop (NC)
FF-bkg							
#1	8067	13083	13258	2249	13494	13824	84
#2	8122	11296	13007	1980	11761	13050	84
#3	9517	11507	12873	1759	13232	13715	110
RL-bkg							
#1	24749	34268	39851	41042	51331	52094	80354
#2	24673	34904	40032	40210	47449	49232	72713
#3	27128	32714	42168	37799	46512	52577	67824
FF/RL							
#1	0.3260	0.3818	0.3327	0.0548	0.2629	0.2654	0.0010
#2	0.3292	0.3236	0.3249	0.0492	0.2479	0.2651	0.0012
#3	0.3508	0.3517	0.3053	0.0465	0.2845	0.2609	0.0016
FF/RL, mean	**0.3353**	**0.3524**	**0.3210**	**0.0502**	**0.2651**	**0.2638**	**0.0013**
FF/RL, SD	0.0135	0.0291	0.0141	0.0042	0.0184	0.0025	0.0003
IRES activity, %	**100.0**	**105.1**	**95.7**	**15.0**	**79.1**	**78.7**	**0.4**
SD	4.0	8.7	4.2	1.3	5.5	0.8	0.09

Luciferase measurement data from Figure 6e-f

NC *(Negative Control): 2A-stop (stop codon introduced before FFLuc in both ORFs)*

3 hpt		**IRES**	**cap**	**IRES/cap**	**ppORF/uORF**		
not infected	**n**	**FF-bkg**	**RL-bkg**	**FF/RL**	**ppORF/uORF**	**mean**	**SD**
IRES-ppORF	1	37041	484731	0.07642	12.64	**11.86**	**0.50**
	2	34777	476957	0.07291	12.23		
	3	34773	465106	0.07476	11.86		
	4	32136	449960	0.07142	11.65		
	5	33260	459279	0.07242	11.35		
	6	35915	485611	0.07396	11.39		
IRES-uORF	1	3380	559090	0.00605			
	2	3482	584274	0.00596			
	3	3498	555132	0.00630			
	4	3170	517291	0.00613			
	5	3752	587799	0.00638			
	6	3374	519819	0.00649	**signal/ORF**	**n = 3**	
IRES-ppORF NC	1	224	774701	0.00029	0.0039		
	2	178	709994	0.00025	0.0034		
	3	214	765378	0.00028	0.0038		
IRES-uORF NC	1	652	625990	0.00104	0.17		
	2	588	616949	0.00095	0.15		
	3	548	616855	0.00089	0.14		

6 hpt		**IRES**	**cap**	**IRES/cap**	**ppORF/uORF**		
not infected	**n**	**FF-bkg**	**RL-bkg**	**FF/RL**	**ppORF/uORF**	**mean**	**SD**
IRES-ppORF	1	51270	502471	0.10204	13.06	**13.89**	**1.04**
	2	50518	480531	0.10513	14.51		
	3	50037	484038	0.10337	12.72		
	4	51428	466946	0.11014	13.79		

	5	56596	468766	0.12073	15.59		
	6	50109	469368	0.10676	13.65		
IRES-uORF	1	3895	498584	0.00781			
	2	3807	525428	0.00725			
	3	3709	456364	0.00813			
	4	3577	447851	0.00799			
	5	3809	491876	0.00774			
	6	3771	482007	0.00782	signal/ORF	n = 3	
IRES-ppORF NC	1	217	655593	0.00033	0.0031		
	2	201	645203	0.00031	0.0029		
	3	233	670679	0.00035	0.0032		
IRES-uORF NC	1	815	618327	0.00132	0.17		
	2	793	588577	0.00135	0.17		
	3	801	635167	0.00126	0.16		

9 hpt		IRES	cap	IRES/cap	ppORF/uORF		
not infected	n	FF-bkg	RL-bkg	FF/RL	ppORF/uORF	mean	SD
IRES-ppORF	1	35224	205930	0.17105	15.90	15.96	1.07
	2	35837	194682	0.18408	17.78		
	3	35224	212861	0.16548	14.63		
	4	38680	204254	0.18937	16.43		
	5	38117	222618	0.17122	15.48		
	6	38203	205418	0.18598	15.54		
IRES-uORF	1	2196	204145	0.01076			
	2	2082	201092	0.01035			
	3	2304	203745	0.01131			
	4	2258	195930	0.01152			
	5	2530	228786	0.01106			
	6	2716	226884	0.01197	signal/ORF	n = 3	
IRES-ppORF NC	1	114	231293	0.00049	0.0028		
	2	122	260675	0.00047	0.0026		
	3	132	270523	0.00049	0.0027		
IRES-uORF NC	1	446	269947	0.00165	0.15		
	2	438	263203	0.00166	0.15		
	3	412	286905	0.00144	0.13		

12 hpt		IRES	cap	IRES/cap	ppORF/uORF		
not infected	n	FF-bkg	RL-bkg	FF/RL	ppORF/uORF	mean	SD
IRES-ppORF	1	23917	86455	0.27664	19.15	16.98	1.78
	2	18729	70331	0.26630	16.82		
	3	17210	70774	0.24317	14.65		
	4	19766	79935	0.24728	15.61		
	5	17059	69460	0.24559	18.93		
	6	16149	63409	0.25468	16.70		
IRES-uORF	1	990	68514	0.01445			
	2	1238	78216	0.01583			
	3	1100	66285	0.01660			
	4	968	61114	0.01584			
	5	1042	80303	0.01298			
	6	1038	68045	0.01525	signal/ORF	n = 3	

IRES-ppORF NC	1	44	79863	0.00055	0.0022
	2	60	82084	0.00073	0.0029
	3	36	82984	0.00043	0.0017
IRES-uORF NC	1	124	79618	0.00156	0.10
	2	130	80672	0.00161	0.11
	3	164	79923	0.00205	0.14

Titer, PFU/ml	**#1**	**#2**	**#3**	**mean**	**SD**	**log10**
3 hpi	1.6E+02	2.8E+02	4.0E+02	2.8E+02	1.2E+02	2.447
6 hpi	4.2E+04	2.8E+04	5.6E+04	4.2E+04	1.4E+04	4.623
9 hpi	4.4E+05	3.6E+05	5.2E+05	4.4E+05	8.0E+04	5.643
12 hpi	2.4E+06	2.4E+06	2.4E+06	2.4E+06	0.0E+00	6.380

n = 3

n = 6

3 hpt		IRES	cap	IRES/cap	ppORF/uORF	
infected	n	FF-bkg	RL-bkg	FF/RL	ppORF/uORF	mean
IRES-ppORF	1	110239	157996	0.69773	24.60	20.14
	2	100644	155177	0.64858	21.62	
	3	92597	151233	0.61228	17.72	
	4	93163	153331	0.60759	18.76	
	5	89623	154762	0.57910	18.62	
	6	87620	149135	0.58752	19.52	
IRES-uORF	1	5044	177856	0.02836		
	2	4746	158189	0.03000		
	3	5532	160137	0.03455		
	4	5378	166072	0.03238		
	5	5196	167048	0.03110		
	6	5026	166945	0.03011	signal/ORF	n = 3
IRES-ppORF NC	1	501	207052	0.00242	0.0039	
	2	569	216978	0.00262	0.0042	
	3	615	224732	0.00274	0.0044	
IRES-uORF NC	1	1347	190958	0.00705	0.23	
	2	1399	182686	0.00766	0.25	
	3	1377	185504	0.00742	0.24	

n = 6

6 hpt		IRES	cap	IRES/cap	ppORF/uORF	
infected	n	FF-bkg	RL-bkg	FF/RL	ppORF/uORF	mean
IRES-ppORF	1	299902	112103	2.67524	25.45	22.94
	2	294808	120601	2.44449	23.52	
	3	282333	116265	2.42836	20.62	
	4	283958	124990	2.27185	20.02	

	5	293082	119426	2.45409	23.04	
	6	319161	128332	2.48699	25.00	
IRES-uORF	1	12387	117828	0.10513		
	2	11961	115100	0.10392		
	3	13131	111513	0.11775		
	4	12279	108207	0.11348		
	5	12281	115315	0.10650		
	6	11530	115883	0.09950	signal/ORF	n = 3
IRES-ppORF NC	1	1235	151745	0.00814	0.0033	
	2	1401	148205	0.00945	0.0038	
	3	1353	151235	0.00895	0.0036	
IRES-uORF NC	1	4173	140949	0.02961	0.27	
	2	4195	141497	0.02965	0.28	
	3	4037	137781	0.02930	0.27	

n = 6

9 hpt		IRES	cap	IRES/cap	ppORF/uORF	
infected	n	FF-bkg	RL-bkg	FF/RL	ppORF/uORF	mean
IRES-ppORF	1	291030	50285	5.78761	24.45	23.05
	2	280233	50580	5.54039	24.43	
	3	293852	54065	5.43516	21.15	
	4	277057	56362	4.91567	20.56	
	5	267315	51416	5.19906	23.75	
	6	293263	54013	5.42949	23.94	
IRES-uORF	1	11134	47037	0.23671		
	2	11026	48621	0.22677		
	3	11570	45021	0.25699		
	4	11851	49565	0.23910		
	5	11222	51272	0.21887		
	6	11488	50662	0.22676	signal/ORF	n = 3
IRES-ppORF NC	1	953	59268	0.01608	0.0030	
	2	991	62701	0.01581	0.0029	
	3	997	61860	0.01612	0.0030	
IRES-uORF NC	1	3115	64860	0.04803	0.21	
	2	3297	69307	0.04757	0.20	
	3	2887	61238	0.04714	0.20	

n = 6

12 hpt		IRES	cap	IRES/cap	ppORF/uORF	
infected	n	FF-bkg	RL-bkg	FF/RL	ppORF/uORF	mean
IRES-ppORF	1	89243	15304	5.83135	20.79	19.67
	2	93167	16237	5.73794	15.47	
	3	94090	15782	5.96186	19.81	
	4	99096	16057	6.17151	20.29	
	5	81439	14694	5.54233	18.75	
	6	84627	14549	5.81669	22.91	
IRES-uORF	1	4061	14481	0.28044		
	2	5416	14605	0.37083		
	3	4444	14770	0.30088		
	4	4295	14123	0.30411		
	5	4341	14682	0.29567		
	6	3957	15586	0.25388	signal/ORF	n = 3

IRES-ppORF NC	1	275	15160	0.01814	0.0031
	2	227	14976	0.01516	0.0026
	3	335	17172	0.01951	0.0033
IRES-uORF NC	1	841	19856	0.04235	0.14
	2	717	19163	0.03742	0.12
	3	831	19556	0.04249	0.14

SD	n = 6
2.55	

SD	n = 6
2.23	

SD
1.73

n = 6

SD
2.48

n = 6