SARS-CoV-2 uses a multipronged strategy to impede host protein synthesis

(article begins on next page)

SARS-CoV-2 utilizes a multipronged strategy to suppress host protein synthesis

Yaara Finkel[1, 5], Avi Gluck[1, 5], Aharon Nachshon[1,5], Roni Winkler[1], Tal Fisher[1], Batsheva Rozman[1], Orel Mizrahi[1], Yoav Lubelsky[2], Binyamin Zuckerman[2], Boris Slobodin[3], Yfat Yahalom-Ronen[4], Hadas Tamir[4], Igor Ulitsky[2], Tomer Israely[4], Nir Paran[4], Michal Schwartz[1*] and Noam Stern-Ginossar[1*]

[1] Department of Molecular Genetics, Weizmann Institute of Science, Rehovot 76100, Israel.

[2] Department of Biological Regulation, Weizmann Institute of Science, Rehovot 76100, Israel

[3] Department of Biomolecular Sciences, The Weizmann Institute of Science, Rehovot, Israel

[4] Department of Infectious Diseases, Israel Institute for Biological Research, Ness Ziona 74100, Israel.

[5] These authors contributed equally to this work

[*] To whom correspondence should be addressed: michalsc@weizmann.ac.il or noam.stern-ginossar@weizmann.ac.il

Abstract:

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the cause of the ongoing coronavirus disease 19 pandemic[1,2]. Translation of viral proteins relies solely on the cellular translation machinery and coronaviruses developed varied mechanisms to repress host mRNA translation to allow the translation of viral mRNAs and concomitantly block the cellular innate immune response [3–6]. Although, different SARS-CoV-2 proteins are implicated in host expression shutoff [7–13], a comprehensive picture of the effects of SARS-CoV-2 infection on cellular gene expression is lacking. Here, we combine RNA-sequencing, ribosome profiling and metabolic labeling of newly synthesized RNA, to comprehensively define the mechanisms that are utilized by SARS-CoV-2 to shutoff cellular protein synthesis. We show that infection leads to a global reduction in translation, but viral transcripts are not preferentially translated. Instead, we find that infection leads to accelerated degradation of cytosolic cellular mRNAs which facilitates viral takeover of the mRNA pool in infected cells. Moreover, we reveal that the translation of transcripts whose expression is induced in response to infection, including innate immune genes, is impaired. We demonstrate this impairment is likely mediated by inhibition of nuclear mRNA export, preventing newly transcribed cellular mRNAs from accessing ribosomes. Overall, our results uncover the multipronged strategy employed by SARS-CoV-2 to commandeer the translation machinery and to suppress host defenses.

Main:

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the cause of the ongoing coronavirus disease 19 (COVID-19) pandemic [1,2]. Upon cell entry, two overlapping ORFs are translated, generating continuous polypeptides which are cleaved into 16 nonstructural proteins (NSPs) [14]. This facilitates the assembly of the viral RNA-dependent RNA polymerase that transcribes genomic and subgenomic RNAs [14]. The subgenomic RNAs are translated into structural and accessory proteins [15–17]. Translation of viral proteins relies on the cellular translation machinery, and coronaviruses (CoVs) have evolved diverse mechanisms including degradation of host mRNA and inhibition of host translation[3,4] to hijack the translation machinery and to inhibit antiviral defense mechanisms, including the interferon (IFN) response. The extent to which SARS-CoV-2 suppresses the IFN response is a key characteristic that distinguishes it from other respiratory viruses [18,19] and the IFN response seems to play a critical role in SARS-CoV-2 pathogenesis [20–24]. Although different SARS-CoV-2 proteins are implicated in host expression and IFN shutoff [7–13], a comprehensive depiction of the effect of SARS-CoV-2 infection on cellular gene expression and the underlying molecular mechanism is still lacking.

To gain a detailed view of the changes that occur in viral and host translation over the course of SARS-CoV-2 infection, we infected Calu3 cells with SARS-CoV-2 at multiplicity of infection (MOI) of 3, resulting in infection of the majority of the cells (Extended Data Fig. 1a) and thus a synchronous cell population. At 3, 5, and 8 hours post infection (hpi), we harvested infected cells as well as uninfected cells for RNA-seq and ribosome profiling (Figure 1a) [25,26]. In order to assess the reproducibility of our experiments we prepared two independent biological replicates for the uninfected and 8hpi time point, and both the mRNA and footprint measurements were reproducible (Extended Data Fig. 1b and 1c). Footprint read length distribution peaked at around 29 nt consistent with previous analyses (Extended Data Fig. 1d)[17,27,28]. Metagene analysis, in which gene profiles are aligned and then averaged, revealed the expected profiles of footprints and mRNAs (Extended Data Fig. 1-h). Using this data, we quantitatively assessed the expression pattern of 8627 cellular transcripts and 12 canonical viral ORFs that are expressed from the genomic and sub-genomic RNAs along SARS-CoV-2 infection. Analysis of the mRNAs and footprints originating from cellular and viral transcripts illustrates SARS-CoV-2's dominance

over the mRNA pool. At 8 hpi viral mRNAs comprise almost 80% of the mRNAs in infected cells (Figure 1b). Surprisingly, however, at the same time point, viral mRNAs account for only ~34% of the RNA fragments engaged with ribosomes in the cells (Figure 1c). In order to quantitatively evaluate the ability of SARS-CoV-2 to co-opt the host ribosomes we calculated the translation efficiency (TE, ratio of footprints to mRNAs for a given gene) of viral and cellular RNA along infection. We then compared the TE of human genes to that of viral genes at each of the time points along infection. At 3 hpi, viral gene translation efficiencies fall within the general range of cellular gene translation (Figure 1d). This indicates that when infection initiates, viral transcripts are translated with efficiencies similar to those of host transcripts. As infection progresses, viral gene translation efficiency relative to cellular genes is significantly reduced. This relative reduction in translation of viral genes at 5 and especially at 8 hpi may indicate that not all viral RNAs are accessible for translation. Since double-membrane replication compartments are formed to accommodate viral genome replication and transcription[29], an appealing possibility is that these compartments encompass a sizable fraction of the viral RNA molecules, and thus prevent them from being a part of the translated mRNA pool.

Deep sequencing measurements inherently provide relative values but not absolute quantification of RNA and translation levels. Since the SARS-CoV-2 encoded protein, NSP1, was recently shown to interfere with translation by blocking the mRNA entry channel of ribosomes [7–10], and since the extent to which SARS-CoV-2 interferes with the overall levels of cellular mRNA was not assessed, we next examined if SARS-CoV-2 infection affects global translation and RNA levels. To quantify absolute translation levels, we measured nascent protein synthesis levels using an analogue of puromycin, O-Propargyl Puromycin (OPP), which incorporates into elongating polypeptide chains [30], facilitating fluorescent labeling of nascent polypeptides via a Click reaction (Extended Data Fig.2a and 2b). We infected Calu3 cells with SARS-CoV-2 and measured nascent protein synthesis levels in uninfected cells and at 3, 5 and 8 hpi. We observed a significant reduction in global translation levels already at 3 hpi which was augmented with time, and at 8 hpi translation activity was reduced by 70% (Figure 1e). In parallel, we measured the levels of rRNA and of total RNA extracted from uninfected cells and along infection. This analysis illustrated there are no major changes in total RNA or in rRNA levels along SARS-CoV-2 infection (Extended Data Fig.3a and 3b). Since the vast majority of RNA in cells originates from rRNA and this dominance of rRNA may mask changes in mRNA levels, we

sequenced total RNA, without rRNA depletion, to assess the relative abundance of cellular and viral mRNAs in uninfected cells and at 3, 5, and 8 hpi. This analysis demonstrates that the pool of mRNA molecules relative to rRNA is growing during infection, due to the massive production of viral transcripts. However, at the same time, the relative fraction of cellular mRNA is reduced by approximately 2-fold (Fig.1F). This suggests that during infection there are both massive production of viral transcripts, and a concomitant substantial reduction in the levels of cellular transcripts. We next assessed the expression pattern of cellular genes along SARS-CoV-2 infection. We clustered the mRNA levels of genes that showed the most significant changes along infection using partitioning clustering, allowing grouping of cellular transcripts into four main classes based on temporal expression profiles in RNA-seq. Overall, the majority of transcripts were reduced during SARS-CoV-2 infection, at different kinetics, accompanied by concurrent reduction in the footprints, but there were also numerous transcripts that were elevated (Figure 1g). Clustering of genes that were elevated along infection (278 genes), revealed mRNAs that were upregulated early, late or showed transient induction (Extended Data Fig. 4a, cluster A, B and C, correspondingly). The genes that were induced in late kinetics were significantly enriched with genes related to immune response, including Toll receptor signaling, chemokine signaling, and cytokine signaling (Extended Data Fig. 4b and Table S1). These genes include IL6 and IL8 which play a significant role in the pathogenesis of SARS-CoV-2 [31], as well as several IFN stimulated genes such as IFIT1, 2 and 3, IRF1, ISG15 and TNF alpha induced proteins.

Our results indicate that the levels of the majority of cellular RNAs are reduced during SARS-CoV-2 infection and this reduction likely contributes to the shutoff of cellular protein synthesis. Reduction in cellular RNA levels could be due to interference with RNA production and/or accelerated RNA degradation. To explore the molecular mechanism, we examined if the reduction of cellular transcripts is associated with their subcellular localization. We found that transcripts that mostly localize to the cytoplasm are reduced more in infected cells compared to transcripts that are mostly nuclear (Figure 2a) and there was a clear correlation between subcellular localization and the extent of reduction in transcript levels following SARS-CoV-2 infection (Extended Figure 4c). Furthermore, mitochondrial encoded transcripts are much less affected by SARS-CoV-2 infection when compared to nuclear encoded transcripts (Figure 2b). The specific sensitivity of cytosolic transcripts indicated these transcripts may be targeted for

degradation during SARS-CoV-2 infection. To directly evaluate mRNA decay in infected cells we employed SLAM-seq [32]. This approach allows to measure endogenous mRNA half-lives based on 4-thiouridine (4sU) incorporation into newly synthesized RNA. After RNA extraction, 4sU is converted to a cytosine analogue using iodoacetamide, and these U to C conversions are identified and quantified by RNA sequencing [32,33]. We applied SLAM-seq to uninfected and SARS-CoV-2 infected Calu3 cells (Extended Data Fig. 5a). We obtained all characteristics of high-quality SLAM-seq libraries; >3000 quantified genes, U- to C-mutations rates starting at 0.2% and rising to 2.6% and an increase with time in the portion of labeled RNA, which was stronger in infected cells, indicating a faster turn-over of RNA in infected cells (Extended Data Fig. 5b-e). There was strong correlation between half-lives estimated from our SLAM-seq measurements from uninfected Calu3 cells and estimates conducted previously in a different cell type[34] (Extended Data Fig. 5f). Importantly, we observed a substantial reduction in cellular mRNA half-lives upon SARS-CoV-2 infection (Figure 2c), indicating increased cellular mRNA degradation in infected cells. Furthermore, the reduction in half-life correlated with the reduction we measured in total RNA expression, indicating RNA decay dominates changes in total RNA during infection (Extended Data Fig. 5g). In agreement with the changes in RNA expression, half-lives of cytoplasmic transcripts were more reduced when compared to transcripts that are mostly nuclear (Figure 2d and Extended Data Fig. 5h). In CoVs, the most prominent and well characterized cellular shutoff protein is NSP1[35]. So far, studies on SARS-CoV-2 NSP1 have demonstrated that it restricts translation by directly binding to the ribosome 40S subunit[7–10], thereby globally inhibiting translation initiation. For SARS-CoV, on top of this translation effect, NSP1 interactions with the 40S was also shown to induce cleavage of translated cellular mRNAs, thereby accelerating their turnover [3,36–38]. We therefore examined if the degradation of cellular transcripts in SARS-CoV-2 infected cells is related to their translation. We observed weak but significant correlation between the translation efficiency of cellular genes and their half-life reduction following infection (Extended Data Fig. 5i), indicating that accelerated turnover of cellular transcripts in infected cells may be related to their translation. To directly assess the role of NSP1 in RNA degradation we analyzed RNA-seq data from cells transfected with NSP1[39], revealing that ectopic NSP1 expression leads to weaker but similar signatures to the ones we identified in infected cells; stronger reduction of cytosolic transcripts compared to

nuclear transcripts and stronger sensitivity of nuclear encoded transcripts (Extended Data Fig. 6a and 6b).

We noticed SARS-CoV-2 infection leads to increased levels of intronic reads in many cellular transcripts (Figures 2e and 2f), indicating SARS-CoV-2 may interfere with cellular mRNA splicing, as was recently suggested [11]. However, massive degradation of mature cytosolic mRNAs may also generate a relative increase in intronic reads. We therefore analyzed the ratio of intronic and exonic reads to rRNA, which we show are unperturbed in infection (Extended Data Fig. 3b). Whereas relative to rRNA levels, exonic reads showed drastic reduction along SARS-CoV-2 infection, the intronic read levels showed only a subtle change (Figure 2g) and in our SLAM-seq measurements we did not detect major changes in the turn-over of intronic RNA in infected cells (Extended Data Fig. 7a). Furthermore, we observed a correlation between the reduction in transcript half-lives and the relative increase in intronic reads (Figure 2h). Likewise, the increase in the ratio of intronic to exonic reads was greater in genes whose expression was reduced along infection compared to genes whose expression was induced (Extended Data Fig. 7b). Finally, we also detected more intronic reads in cells that exogenously expressed NSP1[39] (Extended Data Fig. 7c). These results imply that the increase in intronic reads compared to exonic reads during SARS-CoV-2 infection is mostly driven by accelerated degradation of mature cellular transcripts, which leads to relative reduction in exonic reads. Overall, these findings demonstrate that SARS-CoV-2 infection leads to accelerated degradation of cytosolic cellular mRNAs.

An important aspect of host shutoff during infection is the ability of the virus to hamper the translation of cellular transcripts while recruiting the ribosome to its own transcripts. It had been suggested that SARS-CoV-2 mRNAs are refractory to the translation inhibition induced by NSP1[10,11,40]. However, our measurements indicate that RNA degradation, which is likely mediated by NSP1, plays a prominent role in remodeling the mRNA pool in infected cells and that SARS-CoV-2 dominates the mRNA pool. All of the SARS-CoV-2-encoded subgenomic RNAs contain a common 5′ leader sequence that is added during negative-strand synthesis [41]. We therefore explored whether the genomic 5’UTR or the 5′ leader protect viral mRNAs from NSP1 induced degradation. We fused the viral 5’ leader, the genomic 5’UTR sequence, or a control host 5’UTR, to the 5′ end of a GFP reporter (Extended Data Fig. 8a) and transfected these constructs together with expression vectors encoding NSP1, NSP2 or mCherry (the latter

two were used as controls) into 293T cells (Extended Data Fig. 8b). We found that NSP1 expression suppresses the production of the host-5'UTR-GFP but not of the viral genomic 5'UTR or 5' leader-containing GFP (Figure 3b, 3c, Extended Data Fig. 8c and 8d). We extracted RNA from these cells and observed that the NSP1 induced reduction in host-5'UTR-GFP level was associated with ~12-fold reduction in the GFP mRNA levels whereas the levels of GFP mRNA fused to the SARS-CoV-2 5'UTR or to a 5'leader were only mildly reduced by NSP1 expression (Figure 3d and Extended Data Fig. 8e). The GFP reporter plasmid we used also contains an mCherry reporter expressed from an independent promoter. Reassuringly NSP1 also induces a reduction in mCherry protein and RNA levels when compared to NSP2 (Extended Data Fig. 8f-h). These results indicate that the 5′ leader of viral RNAs provides them protection from NSP1 induced degradation and that this protection contributes to the ability of the virus to dominate the mRNA pool in infected cells.

Our results so far exemplify how SARS-CoV-2 remodels the transcript pool in infected cells. To quantitatively evaluate the role of translational control along SARS-CoV-2 infection, we calculated translation efficiency (TE) of cellular genes along infection. We then centered on genes that showed the strongest reduction or elevation in their relative TE along infection. These genes were clustered into four clusters, based on similarity in their temporal TE profiles, which largely reflects either increased or decreased relative TE along infection. The mRNA and footprint temporal profiles of these genes revealed a clear signature; the genes whose relative TE along infection was reduced were genes whose mRNA increased during infection without a corresponding increase in footprints (Figure 4a and Extended data Fig. 9a). These genes were enriched in immune response genes (FDR $< 10^{-4}$) such as IRF1, IL-6 and CXCL3. Comparing changes in mRNA and TE levels of cellular genes along infection, demonstrates that generally, transcripts which are transcriptionally induced following infection, show a reduction in their relative TE and vice versa (Figure 4a and 4b). These data indicate that newly generated transcripts are less likely to engage with ribosomes. One molecular mechanism that can explain these measurements is inhibition of nuclear mRNA export. Indeed, ORF6 was shown to co-purify with host mRNA export factors [42], and by over expression it was suggested to disrupt nucleocytoplasmic mRNA export [12]. To test if SARS-CoV-2 interferes with nuclear mRNA export, Calu3 cells were left uninfected or infected with SARS-CoV-2 and subcellular localization of polyadenylated transcripts was assessed by cytoplasmic/nuclear

(cyto/nuc) fractionation followed by RNA-seq. We obtained strong correlation between our cyto/nuc measurements and measurements conducted previously in a different cell type [34] (Extended Data Fig. 9b). Infection led to relative nuclear enrichment of most cellular transcripts (Figure 4c). Furthermore, genes whose relative TE was reduced in infection showed a stronger nuclear enrichment, suggesting the inability of induced transcripts to reach the ribosome may be explained by nuclear retention. Since there is massive cytosolic mRNA degradation in infected cells, relative nuclear enrichment may be expected even in lack of interference with nuclear mRNA export. To more accurately assess if this nuclear enrichment is also related to inhibition in nuclear export we used whole-cell extract samples to normalize the cyto/nuc ratios, allowing us to obtain absolute RNA localization values for each compartment [43]. We observed that transcripts that are transcriptionally induced show significant increase specifically in the nuclear fraction in infected cells (Figure 4d), indicating SARS-CoV-2 infection disrupts their nucleocytoplasmic export. Since cytokines and IFN induced genes are induced upon infection, this inability of nascent transcripts to exit the nucleus and reach the ribosomes, may explain why infected cells fail to launch a robust anti-viral response [20,21].

Finally, we also used our measurements to examine viral translation dynamics along SARS-CoV-2 infection. Viral ORFs are translated from the genomic RNA or from a nested series of subgenomic RNAs that contain a common 5′ leader fused to different segments from the 3′ end of the viral genome (Figure 4e and [14]). Since, as indicated above, the translation efficiency of viral genes compared to their cellular counterparts is relatively reduced along infection (Figure 1d), we examined how the translation of viral genes is distributed between different viral transcripts at different times post infection. This focused analysis revealed that ORFs that are located at the 5’ end of the genome tend to show a relative increase in their translation efficiency along infection. In contrast, ORFs that are encoded towards the 3’ end of the genome show relative reduction and ORFs located in the middle of genome showed no major changes in their relative translation efficiency (Figure 4f; p-value = 0.002 for differences in the slope of the TE changes between the groups, using linear mixed model). Analysis of the expression and translation of Mouse Hepatitis Virus strain A59 (MHV) ORFs along MHV infection using published MHV infection RNA-seq and ribosome profiling data [27], revealed a weaker but parallel trend (Extended Data Fig. 10a and 10b). The dynamics of viral translation efficiency along infection needs to be further explored, but since all viral subgenomic RNAs share the same

5'UTR, one possible explanation for these potential differences in translation is an unappreciated role for the 3'UTR or for viral transcripts length, which varies greatly between viral transcripts (Figure 4e).

Most SARS-CoV-2 ORFs are 3′-proximal and translated from dedicated subgenomic mRNAs (Figure 4e). However, several subgenomic RNAs encode for additional out-of-frame ORFs, likely via a leaky scanning mechanism. These include ORF7b and ORF9b, which are translated from the ORF7a and N subgenomic RNAs, and two additional ORFs, ORF3c and ORFS.iORF (ORF2b), that we and others recently identified [17,44–46], and are translated from ORF3a and ORF-S subgenomic RNAs. Since scanning efficiency can be regulated by stress conditions [47], we examined whether the ratio between the translation of the main 3′-proximal ORF and its corresponding out-of-frame ORF (encoded by the same subgenomic RNA) changes during infection. Since ORF9b expression was low in our measurements, it was excluded from this analysis. The translation of ORF7b, ORF3c and ORF2b correlated with the expression of the 3′-proximal main ORF, indicating there are no major changes in the efficiency of ribosome scanning of viral transcripts along infection (Extended Data Fig. 10c).

Using unbiased measurements of translation and RNA expression along SARS-CoV-2 infection, we identified three major courses by which SARS-CoV-2 interferes with cellular gene expression in infected cells; 1. global inhibition of protein translation, 2. degradation of cytosolic cellular transcripts, and 3. specific translation inhibition of newly transcribed mRNAs, which is likely explained by inhibition of nuclear mRNA export. Disruption of cellular protein production using these three components, represents a multi-pronged mechanism that synergistically acts to suppress the host antiviral response (Figure 4g). These mechanisms may explain the molecular basis of the potent suppression of IFN response observed in animal models and in severe COVID-19 patients [18,22].

We reveal here that similarly to what had been described for SARS-CoV NSP1, SARS-CoV-2 NSP1 shuts down host protein translation by two mechanisms: first, it stalls mRNA translation as was reported by others [7–11], leading to general reduction in the translation capacity of infected cells. Second, NSP1 leads to accelerated cellular mRNA degradation. SARS-CoV NSP1 induces endonucleolytic cleavage and subsequent degradation of host mRNAs and this activity depends on its binding to the 40S ribosome subunit [3,8]. Our results are consistent with a similar

mechanism operating in SARS-CoV-2 infected cells as we show cytosolic RNAs are more susceptible to SARS-CoV-2-mediated degradation. Several studies have shown that mRNAs with viral 5′ leader are translated more efficiently compared to control 5'UTR in the presence of NSP1 [10,11,40], but it was further demonstrated that NSP1 inhibits translation of both cellular and a viral 5′ UTR-containing reporter mRNA [8,9], implying that viral mRNAs may not simply evade translation inhibition in the context of the 5′ UTR sequence. Our results support a model in which NSP1 acts as a strong inhibitor of translation and at the same time NSP1 leads to accelerated degradation of cellular but not of viral mRNAs, thus providing the means for viral mRNA to quickly take over and dominate the mRNA pool. This accumulation of SARS-CoV-2 mRNAs may explain how infected cells divert their translation towards viral mRNAs. Overall, our study provides an in-depth depiction of how SARS-CoV-2 efficiently interferes with cellular gene expression, leading to shutdown of host protein production using a multipronged strategy.

Figure legends:

Figure 1: Global reduction of translation and of cellular mRNA levels along SARS-CoV-2 Infection

(A) Calu3 cells were left uninfected or infected with SARS-CoV-2 (MOI=3) for 3, 5 or 8 hours and harvested for RNA-seq and for Ribo-seq. **(B and C)** Percentage of reads that aligned to human or viral coding regions from the sum of aligned reads shown for mRNAs **(B)** and footprints **(C)** in uninfected cells and in cells harvested at 3, 5 and 8 hpi. **(D)** Cumulative frequency of well-expressed human genes (with more than zero reads in each sample, black points) and viral genes (colored points) according to their relative TE at 3, 5 and 8 hpi. **(E)** Protein synthesis measurement by flow cytometry of Calu3 cells infected with SARS-CoV-2 (MOI = 3) for 3, 5 and 8 hpi or an uninfected control following O-Propargyl Puromycin (OPP) incorporation and fluorescent labelling. **(F)** Percent of reads that aligned to the human or viral transcripts from the sum of total RNA reads in uninfected cells and in cells harvested at, 3, 5 and 8 hpi. **(G)** Heat map presenting relative mRNA and footprint levels of well-expressed human transcripts that showed the strongest changes in their mRNA levels at 8 hpi relative to uninfected, across time points during SARS-CoV-2 infection. Shown are expression levels scaled by gene after partitioning clustering. Four main clusters are marked on the right.

Figure 2: Cytosolic cellular RNAs are degraded during SARS-CoV-2 infection

(A) RNA level fold change (FC) of cellular RNAs at different time points after infection relative to uninfected cells. RNAs were grouped to ten bins based on their cytosol to nucleus localization ratio in uninfected Calu3 cells. p-values calculated using t-test comparing the first and last bins in each time point, fold-change between first and last bins mean values: 3 hpi 0.4-fold, 5 hpi 0.8-fold and 8 hpi 0.7-fold. **(B)** The fold change in RNA levels of nuclear encoded or mitochondrial encoded RNAs at different time points after infection relative to uninfected cells. p-values calculated using Wilcoxon tests, fold-change between mitochondrial and nuclear genes mean values: 3 hpi 0.7-fold, 5 hpi 0.5-fold and 8 hpi 0.4-fold. **(C)** Scatter plot of mRNA half-life in SARS-CoV-2 infected cells (MOI=3) relative to uninfected cells as was calculated from SLAM-

seq measurements (**D)** Scatter plot depicting changes in transcript half-life between uninfected and infected cells relative to cytosol/nuclear ratio of cellular transcripts. Pearson's R and p-value are shown. **(E)** RNA reads on exons and introns of the end of IL-32 gene from uninfected cells and at 8 hpi. **(F)** Box plots presenting the ratio of intronic to exonic reads for each gene in uninfected cells and at the different time points along SARS-CoV-2 infection. p-values calculated using t-test on log values. Fold-change between uninfected and infected mean values: 3 hpi 0.8-fold, 5 hpi 1.2-fold and 8 hpi 1.2-fold. **(G)** The % of reads that align to exonic or intronic regions relative to rRNA abundance along SARS-CoV-2 infection. **(H)** Scatter plot showing the change in the ratio of intronic to exonic reads of cellular genes at 7hpi relative to uninfected cells relative to changes in transcript half-life between infected and uninfected cells. Pearson's R and p-value are shown.

Figure 3: SARS-CoV-2 5′-leader protects mRNA from NSP1-mediated degradation

(A) Microscopy images of 293T cells co-transfected with mCherry (top) or NSP1 (bottom) together with a GFP reporter that includes the human beta-globin as a control (Control-5'UTR), the viral genome 5'UTR (CoV2-5'UTR) or the viral 5' leader (CoV2-5'leader). Scale bars are 100μm. **(B)** Flow cytometry analysis of GFP levels in untransfected cells or cells co-transfected with NSP1 or mCherry together with Control-5'UTR, CoV2-5'UTR, or CoV2-5'leader. **(C)** Relative GFP RNA levels from Control-5'UTR, CoV2-5'UTR, or CoV2-5'leader in cells expressing NSP1 or mCherry as measured by quantitative RT-PCR. Data points show measurement of biological replicates.

Figure 4: The translation of induced transcripts is impaired during infection

(A) Heat map presenting relative TE, mRNA and footprints (FP) of human genes that showed the strongest changes in their relative TE along SARS-CoV-2 infection. Shown are relative expression ratios after partitioning clustering based on changes in relative TE values. **(B)** Scatter plot presenting cellular transcript levels in uninfected cells compared to 8hpi. Genes are colored based on the relative change in their TE between uninfected and 8hpi. Central cytokines and IFN stimulated genes are labeled. **(C)** Scatter plot depicting Cyto/Nuc ratio in infected (7 hpi,

MOI=3) and uninfected cells. Genes with reduced TE (decreased TE clusters) are shown in purple **(D)** Effects of infection on cytosolic and nuclear normalized RNA abundance. Transcripts are divided according to the clusters shown in Extended data Fig. 4a (representing pattern of induced transcripts along infection) and all the rest of the transcripts (ATR). Colored rectangles represent the media fold-change between infected and uninfected samples (red decreasing, green increasing, "--" fold-change < -2, "-" -2< fold-change< 0, "+" 0< fold-change <1, "++" fold-change > 2), p-values were calculated from the interaction term in a linear model. Log10 fold-change between infected and uninfected mean values in each cluster are: Induced late cyto 0.2, nuc 0.5; Induced early cyto 0.02, nuc 0.2; Induced mix cyto 0.04, nuc 0.1, All the rest cyto 0.5, nuc 0.08. **(E)** Schematic presentation of the SARS-CoV-2 genome organization, the subgenomic mRNAs and the main ORFs. **(F)** Relative translation efficiency of each canonical viral ORF along infection. Genes are divided to three groups based on their physical location along the genome. **(G)** A model of how SARS-CoV-2 suppresses host gene expression through multi-pronged approach: 1. Global translation reduction; 2. Degradation of cytosolic cellular mRNAs; 3. Specific translation inhibition of newly synthesized cellular mRNAs, likely through inhibition of nuclear mRNA export.

Acknowledgements

We thank Stern-Ginossar lab members for providing valuable feedback. We thank Nevan Krogan for the SARS-CoV-2 ORFs expression plasmids, Ghil Jona and Weizmann Bacteriology and Genomic Repository Units for technical assistance, members of the virology research group at the IIBR for their contribution and support and Shay Weiss for biosafety guidance. This study was supported by Miel de Botton and by a research grant from the Corona Response Fund. Work in the Stern-Ginossar lab is supported by a European Research Council consolidator grant (CoG-2019-864012). N.S-G is an incumbent of the Skirball Career Development Chair in New Scientists and is a member of the European Molecular Biology Organization (EMBO) Young Investigator Program. The authors declare no competing interests.

Methods

Cells and viruses

Calu3 cells (ATCC HTB-55) were cultured in 6-well or 10cm plates with RPMI supplemented with 10% fetal bovine serum (FBS), MEM non-essential amino acids, 2mM L-Glutamine, 100Units/ml Penicillin and 1% Na-pyruvate. Monolayers were washed once with RPMI without FBS and infected with SARS-CoV-2 virus, at a multiplicity of infection (MOI) of 3, in the presence of 20 µg per ml TPCK trypsin (Thermo scientific). Plates were incubated for 1 hour at 37°C to allow viral adsorption. Then RPMI medium supplemented with 2% fetal bovine serum, was added to each well. SARS-CoV-2 BavPat1/2020 Ref-SKU: 026V-03883 was kindly provided by C. Drosten, Charité–Universitätsmedizin Berlin, Germany. It was propagated (5 passages) and titered on Vero E6 cells and then sequenced [17] before it was used. Handling and working with SARS-CoV-2 virus was conducted in a BSL3 facility in accordance with the biosafety guidelines of the Israel Institute for Biological Research. The Institutional Biosafety Committee of Weizmann Institute approved the protocol used in these studies.

Preparation of ribosome profiling and RNA sequencing samples

SARS-CoV-2 infected cells were harvested at 3, 5, and 8 hours post infection (hpi), two independent biological replicates were done for the 8hpi time point and for uninfected cells that were harvested in parallel at 3 and 5hpi. For RNA-seq, cells were left uninfected or infected as described above and at the indicated time points washed with PBS and then harvested with Tri-Reagent (Sigma-Aldrich), total RNA was extracted, and poly-A selection was performed using Dynabeads mRNA DIRECT Purification Kit (Invitrogen) mRNA sample was subjected to DNAseI treatment and 3’ dephosphorylation using FastAP Thermosensitive Alkaline Phosphatase (Thermo Scientific) and T4 PNK (NEB) followed by 3’ adaptor ligation using T4 ligase (NEB). The ligated products used for reverse transcription with SSIII (Invitrogen) for first strand cDNA synthesis. The cDNA products were 3’ ligated with a second adaptor using T4 ligase and amplified for 8 cycles in a PCR for final library products of 200-300bp. For Ribo-seq libraries, cells were treated with 100µg/mL CHX for 1 minute. Cells were then placed on ice, washed twice with PBS containing 100µg/mL CHX, scraped from 10cm plates, pelleted and lysed with lysis buffer (1% triton in 20mM Tris 7.5, 150mM NaCl, 5mM $MgCl_2$, 1mM

dithiothreitol supplemented with 10 U/ml Turbo DNase and 100µg/ml cycloheximide). After lysis, samples stood on ice for 2h and subsequent Ribo-seq library generation was performed as previously described [28]. Briefly, cell lysate was treated with RNAseI for 45min at room temperature followed by SUPERase-In quenching. Sample was loaded on sucrose solution (34% sucrose, 20mM Tris 7.5, 150mM NaCl, 5mM $MgCl_2$, 1mM dithiothreitol and 100µg/ml cycloheximide) and spun for 1h at 100K RPM using TLA-110 rotor (Beckman) at 4c. Pellet was harvested using TRI reagent and the RNA was collected using chloroform phase separation. For size selection, 15uG of total RNA was loaded into 15% TBE-UREA gel for 65min, and 28-34 nt footprints were excised using 28 nt and 34 nt flanking RNA oligos, followed by RNA extraction ribosome profiling library construction as previously described[28]

Sequence alignment, metagene analysis

Sequencing reads were aligned as previously described [17]. Briefly, linker (CTGTAGGCACCATCAAT) and poly-A sequences were removed and the remaining reads were aligned to the hg19 and to the SARS-Cov-2 genome (Genebank NC_045512.2) with 3 changes to match the used strain (BetaCoV/Germany/BavPat1/2020 EPI_ISL_406862), 241:C−>T, 3037:C−>T, 23403:A−>G]. Alignment was performed using Bowtie v1.1.2 [48] with maximum two mismatches per read. Reads that were not aligned to the genome were aligned to the transcriptome (using the known canonical isoform UCSC gene annotations) and to SARS-CoV-2 junctions that were recently annotated [41]. The aligned position on the genome was determined as the 5' position of RNA-seq reads, and for Ribo-seq reads the p-site of the ribosome was calculated according to read length using the off-set from the 5' end of the reads that was calculated from canonical cellular ORFs. The offsets used are +12 for reads that were 28-29 bp and +13 for reads that were 30-33 bp. Footprint reads that were in other lengths were discarded. In all figures presenting ribosome density data, only footprint lengths (28-33nt) are presented.

For the metagene analysis only genes with CDS length of at least 300 nucleotides, UTRs of at least 50 nucleotides and more than 50 reads in the analyzed window around the start or the stop codon were used. For each gene, reads were normalized to the sum of reads in the analyzed window and then averaged.

Gene filtering, quantification and RPKM normalization

For cellular gene quantification, genes were filtered according to the number of reads as follows. The number of reads aligned to the CDS of each gene in each replicate, from at least one of the extreme conditions (uninfected or 8hr) had to be greater than 50 reads for the mRNA libraries and greater than 25 for the footprint libraries. In addition, genes with zero reads in any of the samples (mRNA or footprint of any of the time points) were excluded. Histone genes (which are not polyadenylated) were excluded from the analysis. RNA-seq read coverage on CDS was normalized to units of reads per kilobase per million (RPKM) in order to normalize for CDS length and for sequencing depth. For analysis in which host and viral gene expression were compared (Figure 1b, 1c, 1d and 1f), the RPKM was calculated based on the total number of uniquely aligned reads to the coding regions of both the host and the virus. For analysis that was focused on cellular gene expression, for RNA expression, the RPKM values were calculated based on the total number of uniquely aligned reads to the cellular coding regions. RPKM values were further scaled according to the ratio of the aligned host mRNA reads to the total aligned reads, including viral, rRNA and tRNA reads in the total RNA sequencing (without polyA selection), in order to keep the total reads equal across samples. For footprint libraries, read coverage of cellular genes was normalized to units of RPKM normalizing to the total CDS aligned ribosome profiling reads, including both viral and host reads.

Since the viral RNAs are widely overlapping, RNA-seq RPKM levels of viral genes were computed with deconvolution as was previously described for MHV[27]. First, values for each gene were calculated by subtracting the RPKM of an ORF from the RPKM of the ORF located just upstream of it in the genome. Then, for subgenomic RNAs, leader-body junctions were quantified based on the number of uniquely mapped reads that span each canonical junction using STAR 2.5.3a aligner [49]. Finally, based on the correlation between the deconvoluted RPKM and junction abundance of the subgenomic RNAs, the RPKM levels of all viral RNAs (including the genomic RNA) were estimated. Footprint coverage of canonical viral ORFs were determined as described above for cellular genes. Viral and host genes TE was calculated as the ratio between footprint RPKM and RNA RPKM. To compare viral TE along infection, viral gene TE was further normalized by dividing the TE of each viral gene by the sum of viral genes TE in each sample and multiplying by 100. For comparing the relative translation levels of canonical ORFs and overlapping viral ORFs decoded from the same subgenomicRNA ORF-RATER was

used[50]. In order to estimate the error in our expression measurements of out-of-frame ORFs, for each of these ORFs, we defined 500 random partial ORF-regions, which range in length between 50% and 100% of the original ORF, and used ORF-RATER to quantify the expression from these regions while keeping all other ORFs unmodified. Based on these values we have added standard deviation for the ORF expression measurement. Data from Irigoyen et al. [27] were analyzed for calculating changes in relative viral genes TE. Due to differences in the percent of aligned viral reads between the replicates, we analyzed only replicate 1 which showed the expected gradual increase in viral mRNA and FP reads along infection.

Clustering and heatmaps

Clustering was performed on 2000 cellular genes that showed the strongest change based on the fold-change in RNA-seq expression levels between uninfected samples and 8 hpi samples (both averaged across duplicates). For clustering of upregulated genes, cellular genes that showed at least 1.5-fold increase in expression levels between uninfected samples and 8 hpi were used. RNA-seq and footprint measurements for each of these genes were scaled so that the minimum level for across samples is zero and the maximum is 100. Hierarchical clustering of these normalized values was performed using ward.D2 method on Pearson correlations between scaled RNA-seq measurements using the means of the uninfected and 8 hpi samples.

For presenting changes in relative translation efficiency (TE) clustering was performed on the 35 most increased and 35 most decreased genes based on the fold-change in relative TE between uninfected samples and 8 hpi samples (averaged across duplicates, fold-change > 1.7 or fold-change < -2.4). TE, RNA-seq and footprint measurements for these gene were scaled so that the average level for each gene across samples is one, using mean expression from uninfected and 8 hpi duplicates. Hierarchical clustering of genes was performed using ward.D2 method on Pearson correlations between scaled relative TE measurements. With this clustering we obtained 4 homogenous groups of genes, each one of them showing a clear different pattern of behavior with time.

Quantification of intronic reads

Read density for introns was calculated as described above for exons, with intron annotations based on the known canonical isoform UCSC gene annotations. To avoid biases from intron read count, genes without introns, or genes where one of the introns overlaps with an exon of another gene were excluded. In addition, genes with low number of reads (< 20 on the exons, < 2 on the introns) were omitted. The number of reads on exons and introns was normalized by the total length of the exons and introns respectively to get quantification proportional to the number of molecules. Finally, the normalized number of reads on introns was calculated as percentage of the normalized number of reads on exons. Statistical significance (in figure 2f) was tested using a paired t-test on the log values of the percentage (with offset of 0.1 to overcome zero values).

Protein synthesis measurement using O-Propargyl Puromycin (OPP)

OPP assay (OPP, Thermo Fisher Scientific) was carried out following the manufacturer's instructions. Briefly, cells were collected following treatment with 10μM O-Propargyl Puromycin for 30 minutes at 37˚C. The cells were then fixed for 15 minutes in 3.7% formaldehyde, and permeabilized in 0.1% Triton X-100 for 15 minutes. OPP was then fluorescently labeled by a 30-minute incubation in Click-iT® Plus OPP reaction cocktail with Alexa Fluor®594 picolyl azide (Thermo Fisher Scientific). Cells were analyzed using BD LSRII flow cytometer. The decrease in translation levels was calculated according to the median Alexa 594 fluorescence intensity between the uninfected and the infected, 8hpi samples.

Pathway enrichment analysis

Enrichment analysis of cellular pathways in specific gene clusters (Extended Data Fig. 4a and Figure 4a) was done with PANTHER version 15.0, with default settings and the PANTHER pathways data set [51,52].

Fractionation assay

Uninfected or SARS-CoV-2 Calu-3 infected cells (MOI=3) at 7 hpi were washed in PBS, trypsinized and resuspended in cold PBS. A fraction of 10% of the cells was then transferred to a new tube and RNA was extracted in Tri-reagent to obtain whole cellular extract. Remaining cells were pelleted for 5 minutes at 300 xg. Cells were resuspended in 150 µl fractionation buffer A (15mM Tris-Cl pH8, 15mM NaCl, 60mM KCl, 1mM EDTA pH 8, 0.5mM EGTA pH 8, 0.5mM

spermidine, and 10U/ml RNase inhibitor), and 150 µl 2X lysis buffer (15mM Tris-Cl pH8, 15mM NaCl, 60mM KCl, 1mM EDTA pH 8, 0.5mM EGTA pH 8, 0.5mM spermidine, 10U/ml RNase inhibitor and 0.5% NP-40) was added followed by 10 minutes incubation on ice. The extract was pelleted for 5 minutes at 400 g and the supernatant containing the cytoplasmic fraction was removed to a new tube. This was centrifuged again at 500 g for 1 minute, the supernatant was transferred to a new tube and RNA was extracted with Tri-reagent. The nuclear pellet was resuspended in 1ml RLN buffer (50mM Tris-Cl pH8, 140mM NaCl, 1.5mM $MgCl_2$, 0.5% NP-40, 10mM EDTA, and 10U/ml RNase inhibitor) and incubated on ice for 5 minutes. The nuclear fraction was then pelleted for 5 minutes at 500 g, the supernatant was removed and RNA was extracted from the pellet with Tri-reagent. RNA-seq libraries were then prepared from all three fractions as described above.

Fractionation assay analysis

RNA-seq reads from total, nuclear and cytosolic fractions were aligned to the human and viral reference as described above. Human gene read counts were adjusted to RPKM as described above, and then converted to transcripts per million (TPM) by normalizing to the sum of RPKM in each sample, so that the expression levels in each sample sum up to the same value.

A list of 3884 average expressed genes was defined. These genes were genes with 25 or more reads across all samples and whose sum of TPM values in the total RNA samples across replicates was within quantiles 0.4 and 0.9. Based on this list, for each replicate, a linear regression model was calculated of the total fraction as a linear combination of the cytosolic and the nuclear fractions. The regression coefficients were used to normalize the cytosolic and nuclear TPM values to obtain absolute localization values [43]. To correct for changes in total RNA levels, the absolute values were further scaled by a factor calculated from total RNA-seq as described above (see ‘Gene filtering, quantification and RPKM normalization’ section).

RNA labeling for SLAM-seq

For metabolic RNA labeling, growth medium of infected Calu3 cells (MOI=3) at 3hpi or uninfected cells was replaced with medium containing 4-Thiouridine (4sU, T4509, Sigma) at a final concentration of 200µM (a concentration that did not induce significant cell cytotoxicity at

4h labeling). Cells were harvested with Tri-reagent at 1, 2, 3, and 4 hours post medium replacement (corresponding to 4,5,6 and 7hpi for infected cells). RNA was extracted under reducing conditions and treated with Iodoacetamide (A3221, Sigma) as described previously [32]. RNA-seq libraries were prepared and sequenced as described above and paired-end reads were sequenced with 51 cycles for each end.

SLAM-seq data analysis and half-life calculation

Alignment of SLAM-seq reads was performed using STAR [49], with parameters that were previously described [17]. First, reads were aligned to a reference containing human rRNA and tRNA, and all reads that were successfully aligned were filtered out. The remaining reads were aligned to a reference of the human and the virus as described above. Reads mapped to the virus were discarded and reads mapped to the human were used in the next steps. Output bam files from STAR were used as input for the GRAND-SLAM analysis[33] with default parameters and with trimming of 5 nucleotides in the 5' and 3' ends of each read. Infected and uninfected samples were analyzed in separate runs. Each one of the runs also included an unlabeled sample (no4sU) that was used for estimating the linear model of the background mutations. The output of GRAND-SLAM is the estimated ratio of newly synthesized out of total molecules for each gene (New to Total Ratio, NTR). The old transcript fraction for each gene in each sample is 1 – NTR, this number reflect the pre-existing mRNA molecules (not labeled) and these values were used for half-life estimation of cellular genes. In the case of uninfected samples, we compared two approaches for calculating mRNA half-life: in the first, we assumed steady state and in the second, we analyzed gene composition in each sample and the old fraction was normalized to the gene composition. These two approaches yielded highly similar values and the half-life values in uninfected cells that are presented in the figures are based on the calculation that assume steady state. In the case of the infected samples, the gene composition from the total RNA levels were used normalize the expression of the old mRNA fraction as follows. The total number of reads in each sample was scaled according to the ratio of cellular mRNA to rRNA and tRNA, as calculated based on sequencing of total RNA-seq without poly-A selection. These normalized ratios were used to normalize the old transcript fractions in each sample.

The half-life of each gene in uninfected and infected cells, was calculated by linear regression of the log values of the calculated old transcript fraction. Estimated variance of the values as calculated by GRAND-SLAM were used as weights in the linear regression. The regression coefficient lambda was converted to half-life as -log(2) / lambda. For further analysis, only genes for which the p-value in the regression was < 0.01 and the adjusted r2 > 0.8 were used.

For analysis of intronic RNA turnover, reads that were aligned to any transcript annotation in Ensembl hg19 annotations were filtered out (may represent exonic reads). The rest of the reads were aligned to hg19 genome and were used as input for GRAND-SLAM using intron annotations based on the known canonical UCSC genes.

Immunofluorescence

Cells were plated on ibidi slides, infected as described above or left uninfected and at the indicated time point washed once with PBS, fixed in 3% paraformaldehyde for 20 minutes, washed in PBS, permeabilized with 0.5% Triton X-100 in PBS for 2 minutes, and then blocked with 2% FBS in PBS for 30 minutes. Immunostaining was performed with rabbit anti-SARS-CoV-2 serum [53] at a 1:200 dilution. Cells were washed and labeled with anti-rabbit Alexafluor 488 conjugated antibody and with DAPI (4=,6-diamidino-2-phenylindole) at a 1:200 dilution. Imaging was performed on a Zeiss AxioObserver Z1 wide-field microscope using a X40 objective and Axiocam 506 mono camera.

Plasmids and cloning

pLVX-EF1alpha-SARS-CoV-2-nsp1-2XStrep-IRES-Puro and pLVX-EF1alpha-SARS-CoV-2-nsp2-2XStrep-IRES-Puro were kindly provided by Nevan Krogan, University of California, San Francisco. mCherry-flag was cloned in to the lentiCRISPR v2 plasmid (addgene #52961) instead of the Cas9 cassette. The viral genomic 5'UTR was constructed based on nucleotides 4-265 of the reported sequence of SARS-CoV-2 isolate Wuhan-Hu-1 (NC_045512.2) by sequential annealing of DNA oligonucleotides (IDT, 5'UTR oligo_1-5 listed in the table below). The coding sequence for the first 12 amino acids (aa) of ORF1a as well as the GFP homology region were added to the 5'UTR by two PCR amplifications. The viral 5'UTR with the 12 aa region was cloned into pAcGFP1-C1 (Takara Biotech) using restriction-free cloning. The entire expression cassette from the promoter to the poly-A site was amplified and cloned into pDecko-mCherry

(Addgene plasmid #78534) using restriction-free cloning. Primers for PCR amplification of fragments were ordered from Sigma-Aldrich. The viral subgenomic 5' leader was amplified from the viral 5'UTR plasmid using primers containing homologous regions to clone the 5' leader back into the plasmid. This was subsequently done using restriction-free cloning. For the human beta-globin (HBB) 5'UTR control plasmid, the HBB 5'UTR with GFP homology was ordered from Sigma-Aldrich as two oligonucleotides. These were used in a self-priming PCR reaction and inserted upstream of GFP in place of the viral 5'UTR using restriction-free cloning. All primers and oligonucleotides used for cloning are listed in the table below.

Primer	Sequence
5'UTR oligo_1	caaGGTACCaaaggtttataccttcccaggtaacaaaccaaccaactttcgatctcttgtagatctgttctctaaacgaa
5'UTR oligo_2	Tgcgtgagtgcactaagcatgcagccgagtgacagccacacagattttaaagttcgtttagagaacagatctacaagaga
5'UTR oligo_3	Catgcttagtgcactcacgcagtataattaataactaattactgtcgttgacaggacacgagtaactcgtctatcttctg
5'UTR oligo_4	Cgaaacctagatgtgctgatgatcggctgcaacacggacgaaaccgtaagcagcctgcagaagatagacgagttactcgt
5'UTR oligo_5	atcatcagcacatctaggtttcgtccgggtgtgaccgaaaggtaagGGATCCaac
UTR_Amplification_round_1_Fwd	GGTCTATATAAGCAGAGCTGGTTTAGTGAACCGattaaaggtttataccttcccaggtaacaaacc
UTR_Amplification_round_1_Rev	GAAACCAGGGACAAGGCTCTCCATcttacctttcggtcacacccggacg
UTR_Amplification_round_2_Fwd	GGTCTATATAAGCAGAGCTGGTTTAGTGAACCG
UTR_Amplification_round_2_Rev	GAACAGCTCGGCGCCCTTGCTCACtgttttctcgttgaaaccagggacaaggctctcc

pDecko cloning Fwd	ccatttgtctcaagatctagttacgccaagcttcgttacataacttacggtaaatggcccg cctggctg
pDecko cloning Rev	gacccacctcccaaccccgaggggacccagtaagatacattgatgagtttggacaaac cacaac
SARS-CoV-2_Leader Fwd	ggtgggaGGTCTATATAAGCAGAGCTGGTTTAGTGAACCGattaaaggtttat accttcc
SARS-CoV-2_Leader Rev	TCACtgttttctcgttgaaaccagggacaaggctctccATgttcgtttagagaacagatc
HBB-5'UTR Fwd	ggtgggaGGTCTATATAAGCAGAGCTGGTTTAGTGAACCGacatttgcttctg acacaactgtgttcactagcaacctcaaacagacacc
HBB-5'UTR Rev	TCACtgttttctcgttgaaaccagggacaaggctctccATggtgtctgtttgaggttgctagtg aacacagttgtgtcagaagcaaatgt

627

628 Reporter assay

629 293T cells were transfected using JetPEI (Polyplus-transfection) following the manufacturer's
630 instructions. 24 hours post transfection cells were imaged on a Zeiss AxioObserver Z1 wide-field
631 microscope using a X20 objective and Axiocam 506 mono camera and assayed for reporter
632 expression by flow cytometry on a BD Accuri C6 flow cytometer. In parallel cells were assayed
633 for expression of NSP1, NSP2 and mCherry-flag and reporter mRNA levels as detailed below.

634

635 Flow cytometry analysis of strep and flag tags

636 The expression of NSP1, NSP2 and mCherry was verified by staining of the fused tags, strep-tag
637 for NSP1 and NSP2 and flag-tag for mCherry, followed by flow cytometry. Cells were fixed in
638 4% paraformaldehyde, permeabilized in 0.1% Triton X100, and stained using either Strep-
639 Tactin®XT DY-649 (IBA-lifesciences) or Alexa Fluor®647 anti-DYKDDDDK Tag Antibody
640 (BioLegend). Flow cytometry analysis was performed on BD Accuri C6 and analyzed on
641 FlowJo. Normalization to mode is presented in the histograms.

Quantitative real-time PCR analysis

Total RNA was extracted using Direct-zol RNA Miniprep Kit (Zymo Research) following the manufacturer's instructions. cDNA was prepared using qScript FLEX cDNA Synthesis Kit with random primers (Quanta Biosciences) following the manufacturer's instructions. Real time PCR was performed using the SYBR Green PCR master-mix (ABI) on the QuantStudio 12K Flex (ABI) with the following primers (forward, reverse):

GFP (TGACCCTGAAGTTCATCTGC, GAAGTCGTGCTGCTTCATGT)

mCherry (ACCGCCAAGCTGAAGGTGAC, GACCTCAGCGTCGTAGTGGC)

18S (CTCAACACGGGAAACCTCAC, CGCTCCACCAACTAAGAACG)

GFP and mCherry mRNA levels were calculated relative to 18S rRNA.

Data availability

All next-generation sequencing data files were deposited in Gene Expression Omnibus under accession number GSE162323. Reviewer password : mlcjmqgqljglzsz

References:

1. Zhu, N. *et al.* A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* **382**, 727–733 (2020).
2. Zhou, P. *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273 (2020).
3. Kamitani, W., Huang, C., Narayanan, K., Lokugamage, K. G. & Makino, S. A two-pronged strategy to suppress host protein synthesis by SARS coronavirus Nsp1 protein. *Nat. Struct. Mol. Biol.* **16**, 1134–1140 (2009).

4. Lokugamage, K. G. *et al.* Middle East Respiratory Syndrome Coronavirus nsp1 Inhibits Host Gene Expression by Selectively Targeting mRNAs Transcribed in the Nucleus while Sparing mRNAs of Cytoplasmic Origin. *J. Virol.* **89**, 10970–10981 (2015).

5. Stern-Ginossar, N., Thompson, S. R., Mathews, M. B. & Mohr, I. Translational control in virus-infected cells. *Cold Spring Harb. Perspect. Biol.* **11**, a033001 (2019).

6. Abernathy, E. & Glaunsinger, B. Emerging roles for RNA degradation in viral replication and antiviral defense. *Virology* **479**–**480**, 600–608 (2015).

7. Thoms, M. *et al. Structural basis for translational shutdown and immune evasion by the Nsp1 protein of SARS-CoV-2*. *Science* **369**, (2020).

8. Lapointe, C. P. *et al.* Dynamic competition between SARS-CoV-2 NSP1 and mRNA on the human ribosome inhibits translation initiation. doi:10.1073/pnas.2017715118/-/DCSupplemental

9. Schubert, K. *et al.* SARS-CoV-2 Nsp1 binds the ribosomal mRNA channel to inhibit translation. *Nat. Struct. Mol. Biol.* **27**, 959–966 (2020).

10. Shi, M. *et al.* SARS-CoV-2 Nsp1 suppresses host but not viral translation through a bipartite mechanism. *bioRxiv* 2020.09.18.302901 (2020). doi:10.1101/2020.09.18.302901

11. Banerjee, A. K. *et al.* SARS-CoV-2 Disrupts Splicing, Translation, and Protein Trafficking to Suppress Host Defenses. *Cell* **183**, 1–15 (2020).

12. Addetia, A. *et al.* SARS-CoV-2 ORF6 disrupts nucleocytoplasmic transport through interactions with Running title: SARS-CoV-2 ORF6 disrupts mRNA nuclear export. *bioRxiv* 2020.08.03.234559 (2020). doi:10.1101/2020.08.03.234559

13. Xia, H. *et al.* Evasion of Type I Interferon by SARS-CoV-2. *Cell Rep.* **33**, 108234 (2020).

14. V'kovski, P., Kratzel, A., Steiner, S., Stalder, H. & Thiel, V. Coronavirus biology and replication: implications for SARS-CoV-2. *Nat. Rev. Microbiol.* 1–16 (2020). doi:10.1038/s41579-020-00468-6

15. Bojkova, D. *et al.* Proteomics of SARS-CoV-2-infected host cells reveals therapy targets. *Nature* **583**, 469–472 (2020).

16. Davidson, A. D. *et al.* Characterisation of the transcriptome and proteome of SARS-CoV-

2 reveals a cell passage induced in-frame deletion of the furin-like cleavage site from the spike glycoprotein. *Genome Med.* **12**, 68 (2020).

17. Finkel, Y. *et al.* The coding capacity of SARS-CoV-2. *Nature* **589**, 125–130 (2021).

18. Blanco-Melo, D. *et al.* Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19. *Cell* **181**, 1036 (2020).

19. Wyler, E. *et al.* Bulk and single-cell gene expression profiling of SARS-CoV-2 infected human cell lines identifies molecular targets for therapeutic intervention. *bioRxiv* 2020.05.05.079194 (2020). doi:10.1101/2020.05.05.079194

20. Ribero, M. S., Jouvenet, N., Dreux, M. & Nisole, S. Interplay between SARS-CoV-2 and the type I interferon response. *PLoS Pathogens* **16**, e1008737 (2020).

21. Meffre, E. & Iwasaki, A. Interferon deficiency can lead to severe COVID. *Nature* (2020). doi:10.1038/d41586-020-03070-1

22. Hadjadj, J. *et al.* Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. *Science (80-.).* **369**, 718–724 (2020).

23. Bastard, P. *et al.* Autoantibodies against type I IFNs in patients with life-threatening COVID-19. *Science* **370**, (2020).

24. Zhang, Q. *et al.* Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. *Science* **370**, (2020).

25. Ingolia, N. T., Lareau, L. F. & Weissman, J. S. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* **147**, 789–802 (2011).

26. Stern-Ginossar, N. & Ingolia, N. T. Ribosome Profiling as a Tool to Decipher Viral Complexity. *Annu. Rev. Virol.* **2**, 335–349 (2015).

27. Irigoyen, N. *et al.* High-Resolution Analysis of Coronavirus Gene Expression by RNA Sequencing and Ribosome Profiling. *PLoS Pathog.* **12**, 1005473 (2016).

28. Finkel, Y. *et al.* Comprehensive annotations of human herpesvirus 6A and 6B genomes reveal novel and conserved genomic features. *Elife* **9**, e50960 (2020).

29. Wolff, G. *et al.* A molecular pore spans the double membrane of the coronavirus replication organelle. *Science (80-.).* **369**, 1395–1398 (2020).

30. Liu, J., Xu, Y., Stoleru, D. & Salic, A. Imaging protein synthesis in cells and tissues with an alkyne analog of puromycin. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 413–418 (2012).

31. Del Valle, D. M. *et al.* An inflammatory cytokine signature predicts COVID-19 severity and survival. *Nat. Med.* **26**, 1636–1643 (2020).

32. Herzog, V. A. *et al.* Thiol-linked alkylation of RNA to assess expression dynamics. *Nat. Methods* **14**, 1198–1204 (2017).

33. Jürges, C., Dölken, L. & Erhard, F. Dissecting newly transcribed and old RNA using GRAND-SLAM. *Bioinformatics* **34**, i218–i226 (2018).

34. Zuckerman, B., Ron, M., Mikl, M., Segal, E. & Ulitsky, I. Gene Architecture and Sequence Composition Underpin Selective Dependency of Nuclear Export of Long RNAs on NXF1 and the TREX Complex. *Mol. Cell* **79**, 251-267.e6 (2020).

35. Nakagawa, K., Lokugamage, K. G. & Makino, S. Viral and Cellular mRNA Translation in Coronavirus-Infected Cells. in *Advances in Virus Research* **96**, 165–192 (Academic Press Inc., 2016).

36. Kamitani, W. *et al.* Severe acute respiratory syndrome coronavirus nsp1 protein suppresses host gene expression by promoting host mRNA degradation. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 12885–12890 (2006).

37. Narayanan, K. *et al.* Severe Acute Respiratory Syndrome Coronavirus nsp1 Suppresses Host Gene Expression, Including That of Type I Interferon, in Infected Cells. *J. Virol.* **82**, 4471–4479 (2008).

38. Huang, C. *et al.* SARS coronavirus nsp1 protein induces template-dependent endonucleolytic cleavage of mRNAs: Viral mRNAs are resistant to nsp1-induced RNA cleavage. *PLoS Pathog.* **7**, e1002433 (2011).

39. Rao, S. *et al.* Genes with 5’ terminal oligopyrimidine tracts preferentially escape global suppression of translation by the SARS-CoV-2 NSP1 protein. *bioRxiv Prepr. Serv. Biol.* 2020.09.13.295493 (2020). doi:10.1101/2020.09.13.295493

40. TIDU, A. *et al.* The viral protein NSP1 acts as a ribosome gatekeeper for shutting down host translation and fostering SARS-CoV-2 translation. *RNA* rna.078121.120 (2020). doi:10.1261/rna.078121.120

41. Kim, D. *et al.* The architecture of SARS-CoV-2 transcriptome. *Cell* **S0092-8674**, 30406–2 (2020).

42. Gordon, D. E. *et al.* A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **583**, 459–468 (2020).

43. Carlevaro-Fita, J. & Johnson, R. Global Positioning System: Understanding Long Noncoding RNAs through Subcellular Localization. *Molecular Cell* **73**, 869–883 (2019).

44. Firth, A. A putative new SARS-CoV protein, 3c, encoded in an ORF overlapping ORF3a. *J. Gen. Virol.* jgv001469 (2020). doi:10.1099/jgv.0.001469

45. Jungreis, I., Sealfon, R. & Kellis, M. Sarbecovirus comparative genomics elucidates gene content of SARS-CoV-2 and functional impact of COVID-19 pandemic mutations. *bioRxiv* 2020.06.02.130955 (2020). doi:10.1101/2020.06.02.130955

46. Cagliani, R., Forni, D., Clerici, M. & Sironi, M. Coding potential and sequence conservation of SARS-CoV-2 and related animal viruses. *Infect. Genet. Evol.* **83**, (2020).

47. Hinnebusch, A. G. Molecular Mechanism of Scanning and Start Codon Selection in Eukaryotes. *Microbiol. Mol. Biol. Rev.* **75**, 434–467 (2011).

48. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).

49. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

50. Fields, A. P. *et al.* A Regression-Based Analysis of Ribosome-Profiling Data Reveals a Conserved Complexity to Mammalian Translation. *Mol. Cell* **60**, 816–827 (2015).

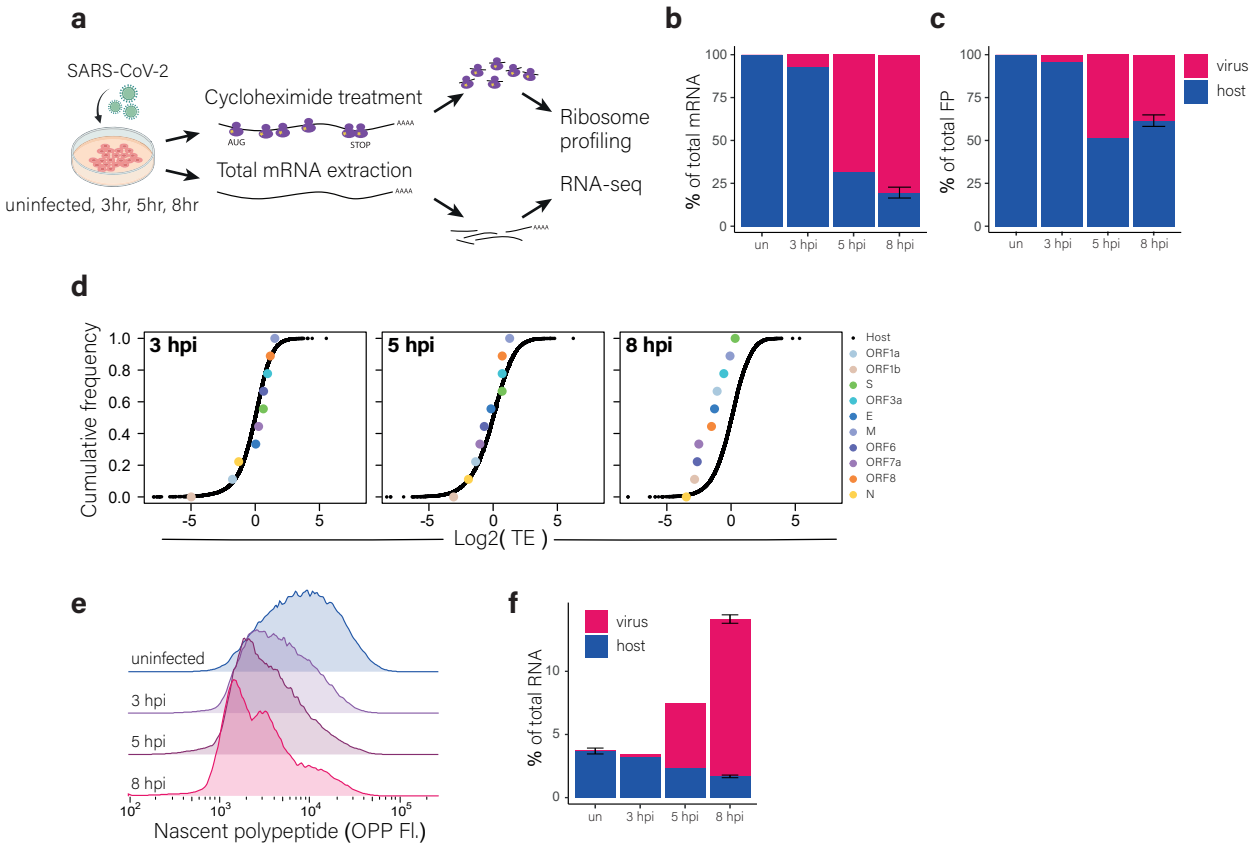
51. Mi, H. & Thomas, P. PANTHER pathway: an ontology-based pathway database coupled with data analysis tools. *Methods Mol. Biol.* **563**, 123–140 (2009).

52. Mi, H., Muruganujan, A., Ebert, D., Huang, X. & Thomas, P. D. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis

tools. *Nucleic Acids Res.* **47**, 419–426 (2018).

53. Yahalom-Ronen, Y. *et al.* A single dose of recombinant VSV-ΔG-spike vaccine provides protection against SARS-CoV-2 challenge. *bioRxiv* 2020.06.18.160655 (2020). doi:10.1101/2020.06.18.160655

Figure 1



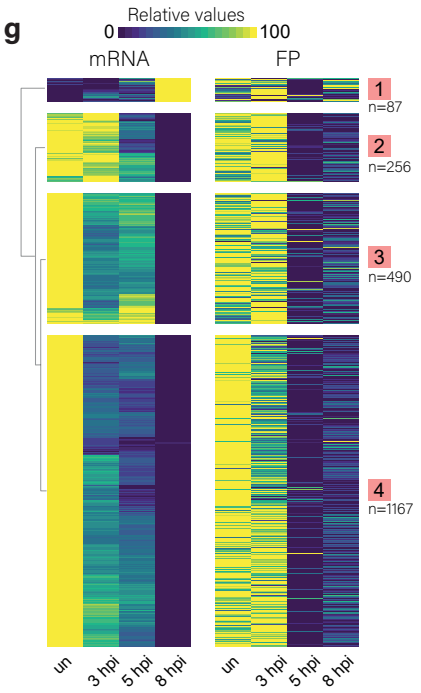


Figure 1: Global reduction of translation and of cellular mRNA levels along SARS-CoV-2 Infection

(A) Calu3 cells were left uninfected or infected with SARS-CoV-2 (MOI=3) for 3, 5 or 8 hours and harvested for RNA-seq and for Ribo-seq. **(B and C)** Percentage of reads that aligned to human or viral coding regions from the sum of aligned reads shown for mRNAs **(B)** and footprints **(C)** in uninfected cells and in cells harvested at 3, 5 and 8 hpi. **(D)** Cumulative frequency of well-expressed human genes (with more than zero reads in each sample, black points) and viral genes (colored points) according to their relative TE at 3, 5 and 8 hpi. **(E)** Protein synthesis measurement by flow cytometry of Calu3 cells infected with SARS-CoV-2 (MOI = 3) for 3, 5 and 8 hpi or an uninfected control following O-Propargyl Puromycin (OPP) incorporation and fluorescent labelling. **(F)** Percent of reads that aligned to the human or viral transcripts from the sum of total RNA reads in uninfected cells and in cells harvested at, 3, 5 and 8 hpi. **(G)** Heat map presenting relative mRNA and footprint levels of well-expressed human transcripts that showed the strongest changes in their mRNA levels at 8 hpi relative to uninfected, across time points during SARS-CoV-2 infection. Shown are expression levels scaled by gene after partitioning clustering. Four main clusters are marked on the right.

Figure 2

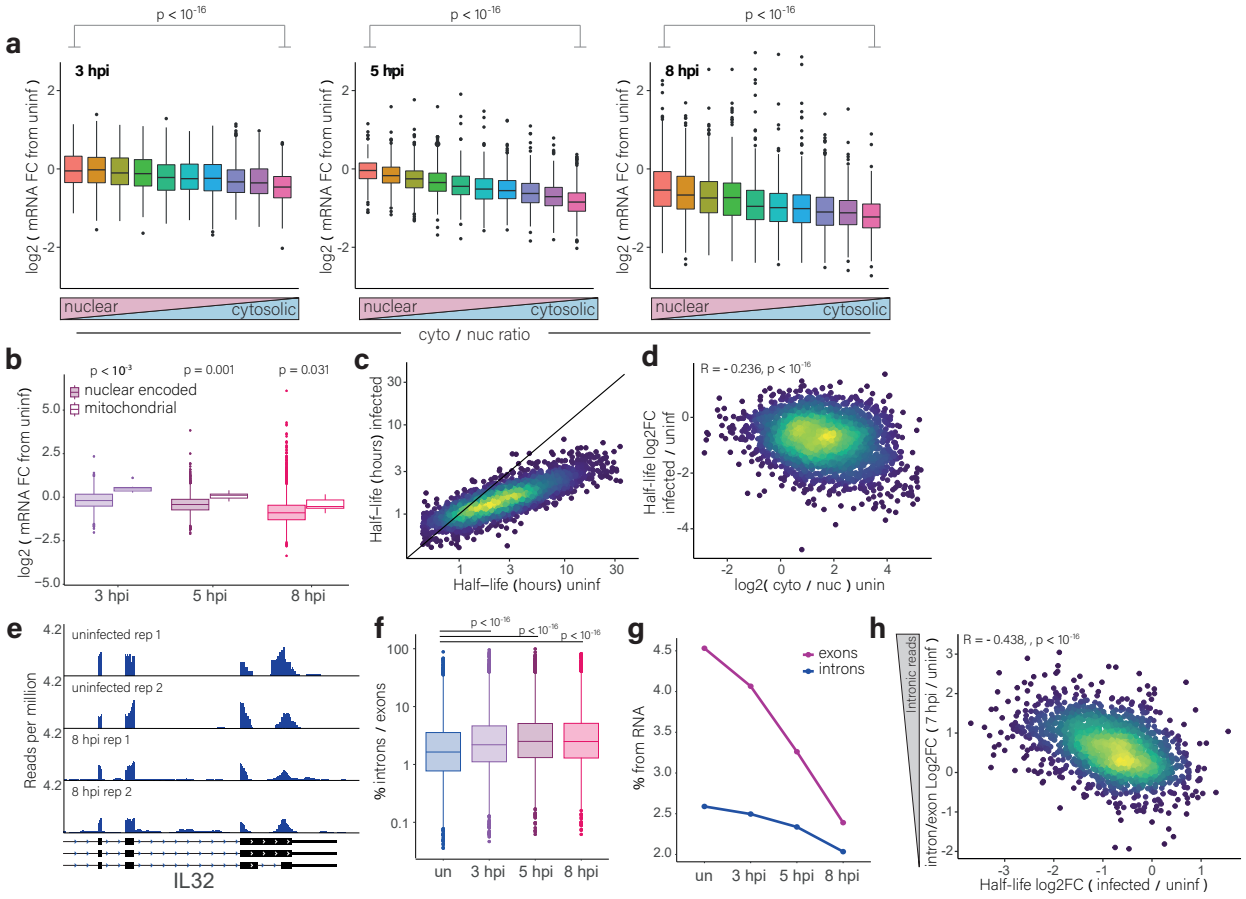


Figure 2: Cytosolic cellular RNAs are degraded during SARS-CoV-2 infection

(A) RNA level fold change (FC) of cellular RNAs at different time points after infection relative to uninfected cells. RNAs were grouped to ten bins based on their cytosol to nucleus localization ratio in uninfected Calu3 cells. p-values calculated using t-test comparing the first and last bins in each time point, fold-change between first and last bins mean values: 3 hpi 0.4-fold, 5 hpi 0.8-fold and 8 hpi 0.7-fold. **(B)** The change in RNA levels of nuclear encoded or mitochondrial encoded RNAs at different time points after infection relative to uninfected cells. p-values calculated using Wilcoxon tests, fold-change between mitochondrial and nuclear genes mean values: 3 hpi 0.7-fold, 5 hpi 0.5-fold and 8 hpi 0.4-fold. **(C)** Scatter plot of mRNA half-life in SARS-CoV-2 infected cells (MOI=3) relative to uninfected cells as was calculated from SLAM-seq measurements **(D)** Scatter plot depicting changes in transcript half-life between uninfected and infected cells relative to cytosol/nuclear ratio of cellular transcripts. Pearson's R and p-value are shown. **(E)** RNA reads on exons and introns of the end of IL-32 gene from uninfected cells and at 8 hpi. **(F)** Box plots presenting the ratio of intronic to exonic reads for each gene in uninfected cells and at the different time points along SARS-CoV-2 infection. p-values calculated using t-test on log values. Fold-change between uninfected and infected mean values: 3 hpi 0.8-fold, 5 hpi 1.2-fold and 8 hpi 1.2-fold. **(G)** The % of reads that align to exonic or intronic regions relative to rRNA abundance along SARS-CoV-2 infection. **(H)** Scatter plot showing the change in the ratio of intronic to exonic reads of cellular genes at 7hpi relative to uninfected cells relative to changes in transcript half-life between infected and uninfected cells. Pearson's R and p-value are shown.

Figure 3

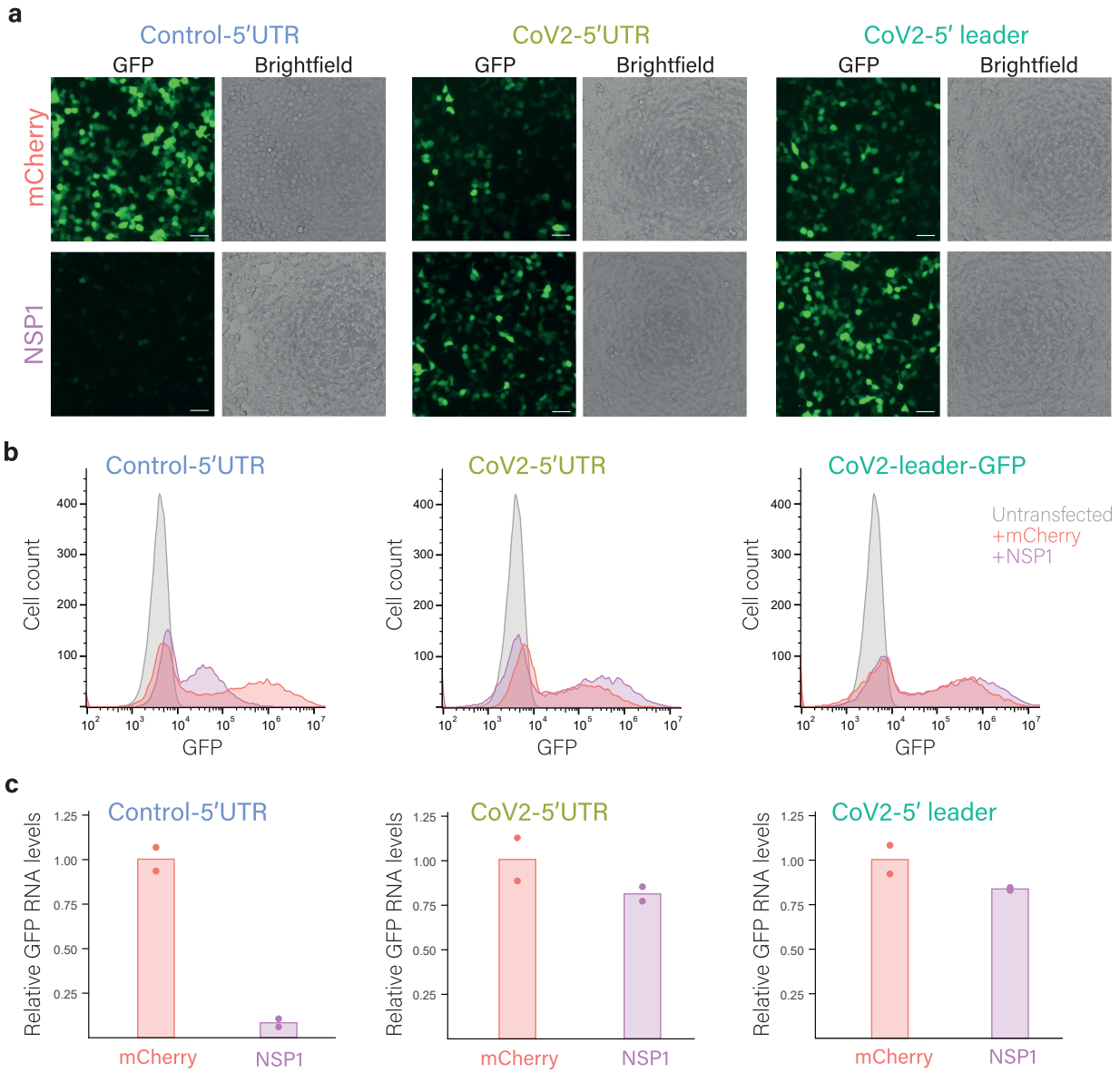


Figure 3: SARS-CoV-2 5′-leader protects mRNA from NSP1-mediated degradation

(A) Microscopy images of 293T cells co-transfected with mCherry (top) or NSP1 (bottom) together with a GFP reporter that includes the human beta-globin as a control (Control-5'UTR), the viral genome 5'UTR (CoV2-5'UTR) or the viral 5' leader (CoV2-5'leader). Scale bars are 100µm. **(B)** Flow cytometry analysis of GFP levels in untransfected cells or cells co-transfected with NSP1 or mCherry together with Control-5'UTR, CoV2-5'UTR, or CoV2-5'leader. **(C)** Relative GFP RNA levels from Control-5'UTR, CoV2-5'UTR, or CoV2-5'leader in cells expressing NSP1 or mCherry as measured by quantitative RT-PCR. Data points show measurement of biological replicates.

Figure 4

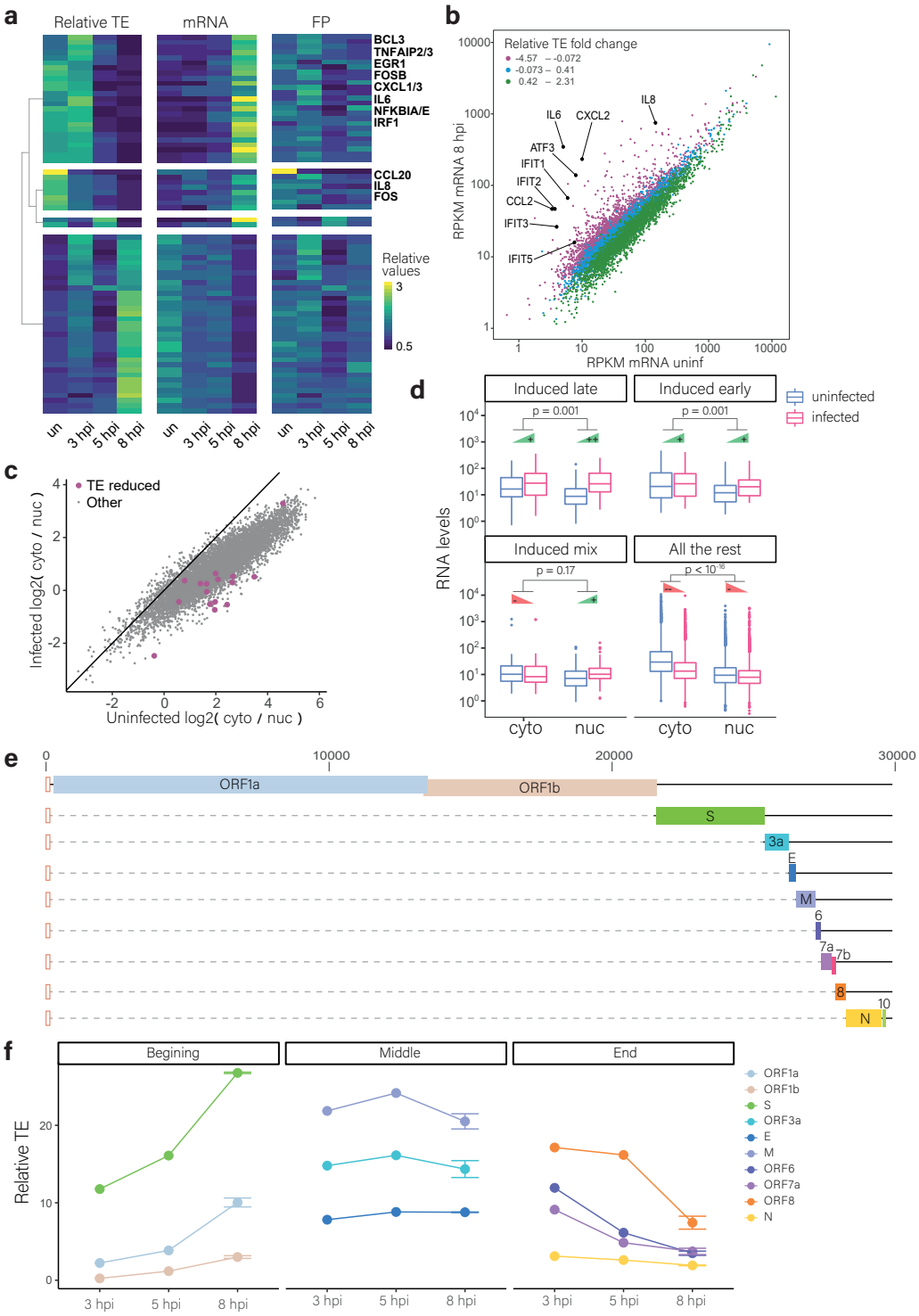


Figure 4: The translation of induced transcripts is impaired during infection

(A) Heat map presenting relative TE, mRNA and footprints (FP) of human genes that showed the strongest changes in their relative TE along SARS-CoV-2 infection. Shown are relative expression ratios after partitioning clustering based on changes in relative TE values. **(B)** Scatter plot presenting cellular transcript levels in uninfected cells compared to 8hpi. Genes are colored based on the relative change in their TE between uninfected and 8hpi. Central cytokines and IFN stimulated genes are labeled. **(C)** Scatter plot depicting Cyto/Nuc ratio in infected (7hpi, MOI=3) and uninfected cells. Genes with reduced TE (decreased TE clusters) are shown in purple **(D)** Effects of infection on cytosolic and nuclear normalized RNA abundance. Transcripts are divided according to the clusters shown in Extended data Fig. 4a (representing pattern of induced transcripts along infection) and all the rest of the transcripts (ATR). Colored rectangles represent the media fold-change between infected and uninfected samples (red decreasing, green increasing, "--" fold-change < -2, "-" -2< fold-change< 0, "+" 0< fold-change <1, "++" fold-change > 2), p-values were calculated from the interaction term in a linear model. Log10 fold-change between infected and uninfected mean values in each cluster are: Induced late cyto 0.2, nuc 0.5; Induced early cyto 0.02, nuc 0.2; Induced mix cyto 0.04, nuc 0.1, All the rest cyto 0.5, nuc 0.08. **(E)** Schematic presentation of the SARS-CoV-2 genome organization, the subgenomic mRNAs and the main ORFs. **(F)** Relative translation efficiency of each canonical viral ORF along infection. Genes are divided to three groups based on their physical location along the genome. **(G)** A model of how SARS-CoV-2 suppresses host gene expression through multi-pronged approach: 1. Global translation reduction; 2. Degradation of cytosolic cellular mRNAs; 3. Specific translation inhibition of newly synthesized cellular mRNAs, likely through inhibition of nuclear mRNA export.