SARS-CoV-2 Nsp14 mediates the effects of viral infection on the host cell transcriptome

Michela Zaffagni[1], Jenna M Harris[1], Ines L Patop[1], Nagarjuna Reddy Pamudurti[1], Sinead Nguyen[1], and Sebastian Kadener[1]

[1] Department of Biology, Brandeis University, Waltham, MA, USA

Key words: SARS-CoV-2, Nsp14, IMPDH2, circRNA, CXCL8, transcription, splicing.

SUMMARY

To identify functions of SARS-CoV-2 proteins, we performed transcriptomic analyses of cells expressing individual viral proteins. Expression of Nsp14, a protein involved in viral RNA replication, provoked a dramatic remodeling of the transcriptome that strongly resembled that observed following SARS-CoV-2 infection. Moreover, Nsp14 expression altered the splicing of more than 1000 genes and resulted in a dramatic increase in the number of circRNAs, which are linked to innate immunity. These effects were independent of Nsp14 exonuclease activity and the co-factor Nsp10. Activation of the NFkB pathway and increased expression of *CXCL8* occurred early upon Nsp14 expression. IMPDH2, which catalyzes the rate-limiting step of guanine biosynthesis, was identified as a key mediator of the effect. Nsp14 expression caused an increase in GTP cellular levels, and the effect of Nsp14 was strongly decreased in presence of an IMPDH2 inhibitor. Together, our data demonstrate an unknown role for Nsp14 with implications for therapy.

INTRODUCTION

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the virus responsible for the COVID-19 pandemic that began in 2019. As of early June 2021, COVID-19 has caused 3.1 million deaths worldwide. Coronaviruses are enveloped, relatively small (60-140 nm diameter), positive-stranded RNA viruses belonging to the *Coronaviridae* family. They derive their name from the crown-like appearance (corona means crown in Latin) of these viruses that results from the spike glycoproteins in their envelope (V'kovski *et al.*, 2021). The SARS-CoV-2 RNA genome is 30-kb long and encodes 14 open reading frames (ORFs) that encode 29 proteins (16 non-structural proteins, 4 structural proteins, and 9 accessory factors, Kim *et al.*, 2020). During the first step of the viral infection, the spike glycoprotein on the viral envelope mediates attachment and fusion with the cellular membrane. Once the virus is in the cytoplasm of the host cells, the host ribosome machinery is recruited for the synthesis of viral proteins. Non-structural proteins are required for viral genome replication and are generated by proteolytic cleavage of the polyprotein encoded by ORF1a and ORF1ab. Once the viral genome is replicated, virions are assembled in the host endoplasmic reticulum-Golgi intermediate complex. Finally, new viral particles are incorporated into vesicles and secreted by the host cells (V'kovski *et al.*, 2021).

Viral infection triggers a variety of pathways in the host cells that ultimately lead to the hijacking of the cellular machineries escape from immune surveillance. SARS-CoV2 infection elicits a peculiar gene expression response that first involves activation of interferon pathway (Blanco-Melo *et al.*, 2020; Vanderheiden *et al.*, 2020; Wyler *et al.*, 2021) and then the NFkB pathway (Kircheis *et al.*, 2020; Hariharan *et al.*, 2021; Wyler *et al.*, 2021), as well as expression of specific cytokines such as IL6 and IL8 (Wang *et al.*, 2007; Blanco-Melo *et al.*, 2020; Coperchini *et al.*, 2020; Park and Lee, 2020).

Recent and intense research on SARS-CoV-2 has resulted in characterizations of functions of the viral proteins in replication of this virus; these functions are often conserved across coronaviruses (V'kovski *et al.*, 2021). Less is known about the roles of the individual proteins in modulating host cell pathways (Gordon, Hiatt, *et al.*, 2020; Gordon, Jang, *et al.*, 2020). For instance, recent studies have proposed that Nsp16 is a splicing modulator (Banerjee *et al.*, 2020) and that Nsp1 and Nsp14 are translational repressors (Schubert *et al.*, 2020; Hsu *et al.*, 2021).

Nsp14 is a 60-kDa protein conserved among coronaviruses that is involved both in viral replication and immune surveillance escape (Ogando *et al.*, 2020). The N-terminal region of Nsp14 contains an exonuclease domain (ExoN) that excises mismatched nucleotides to ensure accurate replication of the viral genome (Ogando *et al.*, 2020). As a result of this proofreading mechanism, coronaviruses have a lower mutation rate than other RNA viruses (error rate of 10^6-10^7 versus 10^3-10^5) (Sanjuán *et al.*, 2010; Robson *et al.*, 2020). Loss of function of the ExoN domain results in increased sensitivity to the RNA mutagen 5-fluorouracil (Eckerle *et al.*, 2010) and attenuated virulence (Graham *et al.*, 2012). Furthermore, the interaction with Nsp10 augments Nsp14 ExoN activity up to 35 fold and inhibition of the interaction between Nsp10 and Nsp14 leads to reduced replication fidelity (Ma *et al.*, 2015; Smith *et al.*, 2015). Nsp14 is also involved in assembly the cap at the 5' end of the viral RNA genome, which is crucial for evading immune surveillance. The C-terminal region of Nsp14 functions as an S-adenosyl methionine-dependent guanine-N7 methyl transferase that is independent of the ExoN activity (Chen *et al.*, 2009).

Recent research suggested that Nsp14 might mediate immune surveillance escape by reducing activity of the interferon pathway and activating the pro-inflammatory response through NFkB transcriptional activity and *CXCL8* expression (Yuen *et al.*, 2020; Hsu *et al.*, 2021; Li *et al.*, 2021). However, the biological mechanism behind these events has not fully characterized.

Furthermore, a global interactome study showed that Nsp14 interacts with the cellular enzyme inosine monophosphate dehydrogenase 2 (IMPDH2) and that this interaction is conserved also in SARS-CoV-1 and MERS-CoV viruses (Gordon, Hiatt, *et al.*, 2020; Gordon, Jang, *et al.*, 2020). IMPDH2 catalyzes the conversion of inosine-5'-monophosphate (IMP) to xanthine-5'-monophosphate (XMP) (Hedstrom, 2009), which is the rate-limiting step of de novo guanine nucleotide biosynthesis. Guanosine-5'-triphosphate (GTP) is necessary for DNA replication and transcription and is used as energy source for translation and as mediator of signal transduction (Hesketh and Oliver, 2019). How the physical interaction between Nsp14 and IMPDH2 impacts the host pathways is not completely understood (Li *et al.*, 2021).

Interestingly, a recent study shows that expression of Nsp14 results in global translation inhibition, but it is not clear if this is a direct effect or a downstream consequence of changes that expression of the protein provoke to the cellular environment (Hsu *et al.*, 2021). Indeed, no direct interaction between Nsp14 and ribosomes or a known translational modulator has been reported, suggesting that the potential translational inhibition might be a downstream effect.

Here we undertook transcriptome analyses in cells that express each SARS-CoV-2 protein individually. Expression of Nsp14 altered the expression of about 4,000 genes, mostly involved in splicing, RNA metabolism, and cell-cycle control. Importantly, the effect of Nsp14 on cellular gene expression resembled the transcriptional changes that occur upon SARS-CoV-2 infection and included the activation of the NFkB pathway and the expression of *CXCL8* (encoding IL8), a marker of acute severe respiratory distress syndrome in COVID-19 patients (Adcock *et al.*, 2015; Blanco-Melo *et al.*, 2020; Kircheis *et al.*, 2020). Intriguingly, we also detected an increase in circRNA expression upon Nsp14 expression; recent studies indicate that circRNAs can act as modulators of the innate immune response during viral infections (Li *et al.*, 2017; Liu *et al.*, 2019;

Chen *et al.*, 2020; Yan and Chen, 2020). Moreover, we showed that the transcriptional response induced by Nsp14 expression is partially mediated by the cellular enzyme IMPDH2. We found that *IMPDH2* mRNA is downregulated upon Nsp14 expression and that the cellular GTP concentration is strongly increased, indicating that Nsp14 might activate IMPDH2 enzymatic activity. In accordance with our hypothesis, we showed that treatment with mycophenolic acid (MPA), an IMPDH2 inhibitor (Seungheon Lee *et al.*, 2020), partially rescued the changes in gene expression induced by Nsp14.

RESULTS

Expression of individual SARS-CoV-2 proteins specifically remodels the transcriptome

Infection of cells with SARS-CoV-2 induces strong and specific changes in the transcriptome of host cells and tissues (Blanco-Melo *et al.*, 2020; Wyler *et al.*, 2021). It is assumed that this response, which occurs during most viral infections, results from the hijacking of the cellular systems by the virus as well as defense by the host. To identify unknown functions of the individual SARS-CoV-2 proteins as well as to determine how much each protein contributes to the takeover of cellular systems, we determined how the transcriptome was altered when each viral protein was expressed individually in a human cell line. Specifically, we expressed individual SARS-CoV-2 proteins (Gordon, Jang, *et al.*, 2020) in HEK293T cells, and after 48 hours we purified RNA, generated and sequenced 3'-RNA-seq libraries, and identified differentially expressed genes (DEGs; Figure 1A). Extensive cell death or changes in morphology did not occur upon expression of individual proteins, as assessed visually.

Interestingly, expression of most proteins resulted in modest or no changes in the transcriptome of the HEK293T cells. Indeed, we detected less than 300 DEGs upon expression of 17 of the 25 tested proteins (Table 1 and Figure 1B). Expression of seven viral proteins, M, Nsp9, E, ORF9b, ORF3a, Nsp13, and Nsp1, modestly but robustly altered the transcriptome (between 300 and 700 DEGs). Interestingly, these DEGs tended to be upregulated rather than downregulated (Figure 1B). Striking, Nsp14 altered the expression of more than 4,000 RNAs (1,862 upregulated and 2,161 downregulated; Figure 1B). The profound impact of Nsp14 expression on the transcriptome of HEK293T cells suggests that this protein has roles beyond its known functions in viral genome proofreading and immune system escape.

To identify cellular pathways affected by the expression of the individual SARS-CoV-2 proteins, we performed Gene Ontology (GO) analysis of the DEGs upon expression of the different viral proteins (Table 2 and Figure 1C). Expression of the viral proteins impacted mRNAs encoding proteins related to different aspects of gene expression (regulation of transcription, translation, and RNA metabolism), cell metabolism, cell division, and innate immunity. For example, expression of Nsp7, an accessory protein of the RNA-dependent RNA polymerase (Kirchdoerfer and Ward, 2019; Hillen *et al.*, 2020), deregulated expression of genes encoding proteins involved in DNA and RNA metabolism (Figure 1C). Expression of Nsp9 and Nsp13, known to bind RNA and to regulate RNA metabolism, respectively (Egloff *et al.*, 2004; Shu *et al.*, 2020), resulted in the mis-regulation of mRNAs encoding proteins involved in RNA metabolism and translation (Figure 1C). Furthermore, expression of the protein ORF3b, known to be an interferon inhibitor (Konno *et al.*, 2020), resulted in alteration of genes involved in cell metabolism and immunity. In fact, the interferon response is known to impact several metabolic pathways, including carbohydrate catabolism (reviewed in Konno *et al.*, 2020), and genes involved in these pathways are downregulated upon ORF3b expression (Figure 1C). In general, expression of the viral proteins mainly impacted pathways related to RNA metabolism, translation, and cell metabolism that were previously reported to be affected upon viral infection (Blanco-Melo *et al.*, 2020; Gardinassi *et al.*, 2020; Wyler *et al.*, 2021). Expression of Nsp14 affected the expression of genes involved in RNA splicing, metabolism, and processing, translation, cell-cycle control, and the cytoskeleton. Nsp14 also impacted the expression of genes involved in general metabolism, but we observed a remarkable effect on genes implicated in nucleotide metabolism (Figure 1C).

Nsp14 expression induces transcriptional changes that resemble SARS-CoV-2 infection

We decided to focus on Nsp14, as expression of this protein resulted in significantly more DEGs than any other tested protein. To obtain a more comprehensive view of the transcriptome changes provoked by Nsp14, we complemented the 3'-RNA-seq data with total RNA-seq data generated from cells transfected with Nsp14. The new datasets strongly resembled the one obtained by 3'-RNA-seq (Figure 2A, Table 3, and Figure S1A) but also contained information about the full-length RNAs, non-poly-adenylated RNAs, and pre-mRNAs.

To gain further insights into the mechanism and pathways altered by Nsp14, we performed Gene Set Enrichment Analysis (GSEA) of the total RNA-seq data from Nsp14-expressing cells. Interestingly, we found that for genes upregulated upon expression of Nsp14, three of the four most highly enriched GSEA datasets were those of cells infected with SARS-CoV-2 (Blanco-Melo *et al.*, 2020) (Figure 2B and 2C, Figure S1A). Moreover, we observed a smaller but significant enrichment of genes upregulated upon infection with MERS, a related virus (Blanco-Melo *et al.*, 2020). Further, the mRNAs downregulated upon expression of Nsp14 tended to be strongly enriched for those downregulated following SARS-CoV2 viral infection (Figure 2B). We did not detect enrichment when our dataset were compared to data from cells infected with influenza A virus (IAV) (Blanco-Melo *et al.*, 2020) (Figure 2B). These results demonstrate that expression of a single viral protein can recapitulate a considerable portion of response of a host cell to SARS-CoV-2 infection and highlights the potential importance of Nsp14 in hijacking host gene expression.

Nsp14 alters gene expression mostly at the transcriptional level

To evaluate whether the transcriptome changes were due to alterations at the transcriptional or post-transcriptional level or a combination of both, we utilized intronic signals from the total

RNA-seq experiment as a proxy for transcriptional activity (Stuart Lee *et al.*, 2020). For each DEG in the 3'-RNA-seq dataset, we determined whether the intronic signal in the total RNA-seq dataset was altered in the same direction. We observed a strong correlation (R = 0.73 and p-value = $2.2e^{-16}$) between these two measures (Figure 2D), indicating that a large part of the response is transcriptional. For example, of the 1,862 genes that were upregulated upon Nsp14 expression, 1,006 also had higher intronic signal (at least 20% signal increase), whereas only 57 of the showed decreased intronic signal. Similarly, of the 2,161 genes downregulated upon Nsp14 expression, 1,242 displayed lower intronic signal (at least 20% signal decrease), whereas only 109 showed higher intronic signal (Table 4). We then performed a similar analysis using the exonic signal of DEGs in the same total RNA-seq dataset. We obtained an even stronger correlation between the changes in exonic signal and total RNA signal for each gene (R=0.87 and p-value= $2.2e^{-16}$, Figure 2E). In sum, we observed that more than 50% of the changes in the transcriptome upon Nsp14 expression are transcriptional. Notably, this might be an underestimation, as changes in intronic signal could be lower than the steady-state RNA levels or some introns could be very efficiently and quickly removed, so that splicing intermediates are not detected. Interestingly, we observed that there are more than 1000 genes that display higher intronic signals, suggesting that in addition to the transcriptional effects, Nsp14 might also affect splicing.

Nsp14 expression provokes changes in alternative splicing and circRNAs production

Genes upregulated upon Nsp14 expression are enriched in genes encoding proteins with GO terms related to RNA metabolism and, more specifically, splicing (Figure 1C). Moreover, for more than 1,000 genes, there were significant increases in intron signal upon expression of Nsp14 without changes in mRNA levels, suggesting that Nsp14 alters the splicing of these pre-mRNAs.

We found that expression of Nsp14 strongly altered the inclusion patterns of almost 25,000 exons, with more than 2,000 exons less likely to be included in the mature mRNA and 238 more likely to be included when Nsp14 was expressed (Figure 3A, Table 5, and Figure S2A). Furthermore, we also identified cases where alternative acceptor or donor splice sites were used upon expression of Nsp14 (Figure 3A and Table 5). Moreover, we observed an increase in the retention of more than 2,000 introns following Nsp14 expression (Figure 3B and Figure 3C). Although the effects on exon inclusion and use of alternative splice sites clearly indicate that Nsp14 influences splicing, the increase in intron retention could be secondary to changes in transcription. However, as most of the introns with increased retention are within genes that were not differentially expressed (Figure 3B), we reason that expression of Nsp14 leads to changes in splicing in this subset of genes as well. Moreover, the effect of Nsp14 on alternative splicing appears to be specific to particular introns and exons in each gene as the majority (~62%) of identified genes had a single spicing event altered (Figure 3D and Table 5).

To further characterize the effects of Nsp14 expression on alternative splicing, we looked for genomic features associated with the altered patterns of splicing in the presence of Nsp14. Interestingly, we found that the affected introns tended to be at the 5' end of the transcript and were embedded in genomic regions with higher-than-average GC content (Figure S2B and Table 6). The GC content may be a confounding factor of the location of these introns, as the GC content is higher at the 5' end of genes and around the transcription start sites (Zhang *et al.*, 2004). Therefore, we concluded that Nsp14 expression has a strong and specific effect on splicing efficiency for a subset of genes.

There were 2053 exons with decreased inclusion and 238 with increased inclusion upon Nsp14 expression (Table 5). Most of these exons were constitutively included or excluded in

control samples. These exons are shorter and have higher GC content around their splice sites than a randomized subset of exons not affected by Nsp14 (Figure S2C and Table 6). The higher GC content suggests that stable RNA structures form around the splice sites in these exon, which correlates with alternative splicing propensity (Zhang, Kuo and Chen, 2011; Lin, Taggart and Fairbrother, 2016).

Using the total RNA-seq data, we also demonstrated that there was a strong increase (more than 2 fold) in the total number of circRNA reads upon Nsp14 expression (Figure 3E). Most of the 246 circRNAs that were differentially expressed upon expression of Nsp14 were upregulated (Figure 3F and Table 7). These deregulated circRNAs were encoded by 190 genetic loci, as some loci generate multiple circRNAs. Interestingly, the levels of most of the mRNAs that host these circRNAs were unchanged (Figure 3G, and Table 4). Together, these results suggest that increased circRNA levels are not caused by enhanced transcription of these loci but rather by increased biosynthesis or stability of the circRNAs. Supporting the hypothesis that increased levels of circRNAs are not due to enhanced transcription, as we observed increased levels of intronic sequences in only 20 out of the 190 loci with upregulated circRNAs (Figure 3H and Table 4).

The effect of Nsp14 on gene expression is independent of the exonuclease activity

Nsp14 is a proofreading exonuclease conserved among coronaviruses (Eckerle *et al.*, 2007, 2010; Denison *et al.*, 2011; Ogando *et al.*, 2020; Robson *et al.*, 2020). To determine whether the exonuclease activity is required for the strong effects of Nsp14 expression on transcription, we performed two different experiments. First, we tested whether co-expression of Nsp10, which dramatically increases the exonuclease activity of Nsp14 (Bouvet *et al.*, 2012, 2014; Ma *et al.*, 2015), resulted in an enhanced effect of Nsp14. Briefly, we transfected HEK293T cells with

plasmids for expressing Nsp14 and/or Nsp10. As a control, we used cells transfected with an empty vector. Co-expression of Nsp14 and Nsp10 did not impact levels of Nsp14 protein (Figure 4A). We then prepared RNA from these samples and quantified the expression of some mRNAs that were altered upon expression of Nsp14 alone. For instance, the levels of *FGF-18* increased upon expression of Nsp14 alone, whereas the levels of *SH2D2A* were significantly downregulated (Figure S3A). Interestingly, expression of Nsp10 alone or in combination with Nsp14 did not further alter the levels of any of the mRNAs tested (Figure S3A). To determine whether the effect is global, we prepared and sequenced total RNA from these samples. Principal component analysis (PCA) showed that data from samples expressing Nsp14 alone separated from the data from the sample transfected with the empty plasmid, whereas that the ones from the samples expressing only Nsp10 almost overlapped with the control. Co-expression of Nsp14 and Nsp10 resulted in transcriptional changes similar to the ones induced by expression of only Nsp14 (Figure S4B). Unlike expression of Nsp14, which provoked a strong remodeling of the mRNA and circRNA populations (Figure 1 and Figure 2), expression of Nsp10 did not considerably alter the gene expression profile (only 88 DEGs, Figure 4B,4C, and Figure S3C). As also suggested by the RT-qPCR experiments (Figure S3A), the gene expression profiles of cells transfected with Nsp14 alone or in combination with Nsp10 are strikingly similar. Indeed, we observed an almost complete overlap between the DEGs resulting from Nsp14 expression and from co-expression of Nsp14 and Nsp10, and there was a very strong correlation between the fold change in DEGs upon Nsp14 expression with and without Nsp10 co-expression (Figure 4B and 4C). Moreover, co-expression of Nsp10 did not result in additional differentially expressed mRNAs (Figure 4B and Figure 4C) and did not alter of the number of circRNAs reads (Figure 4D) compared to expression of Nsp14

alone. These data strongly suggests that the Nsp14 exonuclease activity is not responsible for the dramatic effects of Nsp14 expression on the transcriptome.

To definitely test whether the exonuclease activity is related to the remodeling of gene expression upon Nsp14 expression, we determined the effect of expressing Nsp14 ExoN mutants on gene expression. Briefly, we generated plasmids for expression of Nsp14 with mutations in the ExoN domain: the double mutant D90A/G92A and the single mutant D273A (Figure 4E). These mutations have been previously shown to completely inhibit the exonuclease activity of Nsp14 (Ma *et al.*, 2015; Ogando *et al.*, 2020). We generated and sequenced 3'-RNA-seq libraries from cells that expressed wild-type or ExoN-mutant Nsp14 proteins or eGFP, as a control. PCA indicated that the two mutants have overlapping transcriptomes that differed from the control (Figure S4D). The mutations in the exonuclease domain did not alter the strong effects induced by Nsp14 expression (Figure 4F) and the fold changes in mRNA expression induced upon expression of the wild-type or mutated Nsp14 were also strikingly similar (Figure 4G). Moreover, we did not detect significant differences in the expression of the three tested circRNAs between samples that expressed the wild-type Nsp14 and the ExoN mutants (Figure S3F). Taken together these data indicate that the exonuclease activity of Nsp14 is not required for its effect on the host transcriptome.

Dissecting the remodeling of the transcriptome induced by Nsp14

To gain further insights into the gene expression program orchestrated by Nsp14, we performed a time-course experiment. Briefly, we prepared and sequenced 3'-RNA-seq libraries from cells at 6, 8, 12, 24, and 48 hours after transfection with the plasmid driving expression of Nsp14. As a control, we collected data from cells transfected with a control plasmid (eGFP)

harvested at the same time points. As expected, the biological triplicates mapped together in the PCA (Figure 5A). Data diverged across time and between control and Nsp14 expressing cells. The effect of Nsp14 expression dramatically increased over time with a gradual increase in the number of DEGs (Figure 5B, Figure S4A, and Table 8), as well as an increase in the number of significantly enriched GO terms (Figure 5C and Table 9). For example, expression of genes related to cell metabolism, RNA metabolism, and cell cycle were altered relative to the control only at 24 and 48 hours post transfection (Figure 5C).

We reasoned that the mRNAs that were differentially expressed at early time points could provide cues to the mechanism of action of Nsp14. We observed only 19 DEGs at the 8-hour time point and only 200 DEGs at the 12-h time point (Figure 5B). Interestingly, *CXCL8*, which encodes IL8, was the most highly induced mRNA at the 8-hour time point (fold change =11.5 , p-value $=1.2e^{-5}$, Table 8). Furthermore, we found that *CXCL8* was induced by more than 12 fold consistently between 8 hours and 48 hours post transfection (Figure 5D). It was previously reported that *CXCL8* mRNA is significantly upregulated following infection of cells with SARS-CoV-2 and that the levels of IL8 are increased in COVID-19 patients (Blanco-Melo *et al.*, 2020; Coperchini *et al.*, 2020; Del Valle *et al.*, 2020; Gardinassi *et al.*, 2020; Li *et al.*, 2021).

To investigate whether Nsp14 expression induces the transcription of *CXCL8*, we generated a firefly luciferase reporter under the control of the *CXCL8* promoter (-500 bp to +80 bp around the transcription start site) (Figure S4C related to Figure 5). Expression of Nsp14 led to an almost 4-fold increase in the signal from the luciferase reporter (Figure 5E), indicating that Nsp14 expression induces *CXCL8* transcription.

CXCL8 expression is controlled by several transcription factors, including NFkB, a master regulator of inflammation (Kunsch and Rosen, 1993). Recent reports have shown that NFkB is

activated during infections with SARS-CoV and SARS-CoV-2, although the mechanism and the importance of this activation are not well-established (Liao *et al.*, 2005; Kircheis *et al.*, 2020; Park and Lee, 2020; Hariharan *et al.*, 2021; Hsu *et al.*, 2021; Li *et al.*, 2021). Strikingly, some of the top upregulated RNAs at 8-hours post transfection of Nsp14 were *NFKBIA*, *JUN*, and *ATF3* (Figure S4C related to Figure 5), which encode proteins involved in the activation and modulation of the NFkB pathway (Stein *et al.*, 1993; Baldwin, 1996; Pahl, 1999).

To test the hypothesis that NFkB activation contributes to the early response of cells to Nsp14 expression, we performed an unbiased analysis for transcription factor binding site enrichment on the promoter regions (100-bp regions upstream of the transcription start sites) of the genes with strong upregulation at the 8-hour time point. We found a significant enrichment in those promoters for binding sites recognized by components of the NFkB transcription complex and the NFkB regulatory network (Figure 5F, Table 10, and Table 11, Stein *et al.*, 1993; Baldwin, 1996; Pahl, 1999). We next tested whether expression of Nsp14 activates a luciferase reporter containing a minimal promoter and NFkB sites (Wilson *et al.*, 2013). Indeed, we observed that expression of Nsp14 lead to an approximately 1.6-fold increase in the luciferase signal (Figure S5D), demonstrating that this viral protein activates the NFkB pathway.

We then performed an analysis of extended promoters regions of the mRNAs strongly upregulated early after transfection with the Nsp14-expressing plasmid; we analyzed regions 1,000-bp upstream of the transcription start sites. We found that these regions were enriched for binding sites for factors involved in the interferon response such as EGR2, STAT2, and SP3 (Table 12). Previous work showed that SARS-CoV-2 infection modulates both the interferon and the NFkB pathways (Blanco-Melo *et al.*, 2020; Vanderheiden *et al.*, 2020; Wyler *et al.*, 2021) and that

Nsp14 affects these pathways (Hsu *et al.*, 2021; Li *et al.*, 2021). Our data indicates that the initial response to Nsp14 expression involves modulation of both the NFkB and the interferon pathways.

IMPDH2 partially mediates the transcriptional changes induced by Nsp14 expression

Recent proteomics studies found that Nsp14 interacts with the host protein IMPDH2 (Gordon, Hiatt, *et al.*, 2020; Gordon, Jang, *et al.*, 2020), which catalyzes the rate-limiting step of guanine biosynthesis, but did not reveal whether this interaction results in any metabolic or cellular change (Li *et al.*, 2021). One possibility is that the interaction between Nsp14 and IMPDH2 activates IMPDH2. In support of this, expression of genes involved in purine metabolism were altered upon Nsp14 expression (Fig 1C). Furthermore, *IMPDH2* mRNA expression was significantly downregulated when Nsp14 was expressed (Figure 6A). GTP levels inversely regulate *IMPDH2* expression (Glesne, Collart and Huberman, 1991), suggesting that Nsp14 promotes the activation of IMPDH2, leading to higher GTP levels that negatively regulate *IMPDH2* mRNA expression. To test whether expression of Nsp14 alters IMPDH2 activity, we performed metabolomic analysis of lysates of cells transfected with a plasmid for expression of Nsp14 or a control plasmid. Consistent with modulation of IMPDH2 activity by Nsp14, we detected an approximately 3-fold increase in the levels of xanthine diphosphate and GTP (Figure 6B and Table 13), the downstream products of XMP. Furthermore, we observed increased concentrations of other nucleotides that directly or indirectly derive from IMP (Figure 6B and Table 13), probably due to activated compensatory pathways.

To determine if IMPDH2 is the main pathway by which Nsp14 alters gene expression, we compared the gene expression changes upon Nsp14 expression in the presence of an IMPDH2 inhibitor or the vehicle (DMSO). Briefly, we transfected cells with a Nsp14 or with a control

plasmid (eGFP) and, at 8 hours after transfection, treated the cells with Mycophenolic acid (MPA), a non-isoform-selective pan-IMPDH inhibitor, or DMSO. After 40 hours, cells were harvested. We used a concentration of MPA of 0.5 µM to avoid cytotoxic effects (Qasim et al., 2011). Although we treated the cells with a low concentration of MPA, MPA treatment could reduce the concentration of cellular GTP, which is the primary energy source for mRNA translation. We confirmed that MPA treatment did not alter Nsp14 or IMPDH2 levels by western blot (Figure 6C).

We next performed 3'-RNA-seq on the same samples. Our data indicate that the effects induced by expression of Nsp14 are diminished in presence of MPA (Figure 6D and Figure S5A). We detected only 115 DEGs in cells treated with MPA compared to control cells. In cells expressing Nsp14, we detected 3,761 DEGs. By contrast, there were only 1,245 DEGs in cells expressing Nsp14 that were treated with MPA compared to control cells, demonstrating that a very large fraction of the changes depends on the activity of IMPDH2 (Figure 6E). Approximately 1,000 genes were deregulated in cells expressing Nsp14 with or without MPA (Figure S5B). In addition, the effect of those genes still differentially expressed was milder in presence of MPA. This can be seen by comparing the distribution of fold changes of the DEGs (Figure 6F). When cells expressing Nsp14 were treated with MPA, the distribution of fold change narrowed, indicating that MPA treatment mitigates the effects of Nsp14 even in those mRNAs that are still differentially expressed in the presence of the drug. Moreover, we observed a similar, milder, effect of Nsp14 expression when assaying the levels of a subset of circRNAs (Figure S5C). Together, the data presented here demonstrate that IMPDH2 is a key mediator of the effects of Nsp14 on the transcriptome of the hosting cell.

DISCUSSION

Protein interactome experiments defined the interaction map of SARS-CoV-2 proteins with cellular host proteins (Gordon, Hiatt, *et al.*, 2020; Gordon, Jang, *et al.*, 2020). Further studies showed that Nsp1 blocks cellular translation (Schubert *et al.*, 2020) and RNA immunoprecipitation experiments indicated that Nsp8 and Nsp19 alter protein trafficking (Banerjee *et al.*, 2020), and that Nsp16 alters splicing (Banerjee *et al.*, 2020). Here we used an alternative approach; we performed an expression screen to determine how the individual expression of SARS-CoV-2 viral proteins influenced host cell gene expression. Nsp14 had the most striking effects, deregulating expression of about 4,000 genes. Importantly, expression of Nsp14 had effects on the host cell transcriptome that resembled those of SARS-CoV-2 infection. When expressed individually M, Nsp9, E, ORF9b, ORF3a, Nsp13, and Nsp1 also altered the host cells transcriptome in ways that will require further investigation. Although most of the tested proteins did not dramatically alter gene expression does not rule out roles in remodeling of host cell gene expression. For instance, some may require other viral proteins as co-factors.

Nsp14 is conserved among coronaviruses and has characterized functions in viral replication and immune surveillance escape (Ogando *et al.*, 2020). Here we demonstrated an unexpected function of the SARS-CoV-2 protein Nsp14 in regulation of host cell gene expression. When expressed in HEK293T cells, Nsp14 altered the expression of more than 4,000 genes, strikingly resembling the transcriptional changes induced by SARS-CoV2 infection. *CXCL8* and other NFkB transcriptional targets were upregulated early after induction of Nsp14 expression, suggesting that the activation of the NFkB pathway is one of the first events that contributes to the effects induced by Nsp14. Genes altered later upon expression of Nsp14 are enriched for those encoding proteins involved in control of the cell cycle, nucleotide metabolism, splicing, and

mRNA processing. In accordance with regulation of expression of genes involved in RNA biosynthesis, we detected numerous changes in splicing patterns and increased expression of circRNAs in cells that express Nsp14.

Our data indicate that neither the ExoN activity nor the interaction with Nsp10 are required for the effect of Nsp14 on the host cell transcriptome. However, we do not know if the methylase domain is necessary. Furthermore, Nsp14 contains zinc fingers motifs that can mediate RNA binding, and it would be interesting to determine if Nsp14 interacts directly with cellular nucleic acids.

Our data suggest that Nsp14 mostly alters the host gene expression by acting on transcription. Nevertheless, for a subgroup of genes the response appears to involve alterations of splicing patterns and other post-transcriptional events, and Nsp14 expression clearly alters RNA metabolism. Genes involved in RNA processing and splicing are among the most mis-regulated upon Nsp14 expression, and we observed that Nsp14 expression altered inclusion of thousands of introns and exons. However, unlike splicing modulation by Nsp16 (Banerjee *et al.*, 2020), the effect of Nsp14 seems specific to certain exons and introns as, in most cases, only one splicing event is modulated per gene. In the case of intron retention, the effect could be due to changes in elongation rather than splicing, as we observed that a large fraction of the altered introns are the first introns in the pre-mRNA. It is likely that Nsp16 and Nsp14 influence alternative splicing through different mechanisms, although we could not compare the effects of the two proteins as the plasmid for Nsp16 expression was not in the collection we utilized in our screening (Gordon, Jang, *et al.*, 2020). Whereas Nsp16 directly binds to pre-mRNA (Banerjee *et al.*, 2020) , Nsp14 may trigger a cascade of events that culminate in altered splicing patterns.

Surprisingly, we also detected increased expression of circRNAs upon Nsp14 expression. This increase is not due to enhanced transcription of the hosting RNAs, and our results indicate that Nsp14 influences biosynthesis or stability of some circRNAs. The effects of Nsp14 on the transcriptome may result in inhibition of pathways related to circRNA degradation or in inhibition of cell division, which is known to result in circRNA stabilization (Bachmayr-Heyda *et al.*, 2015). Interestingly, recent studies indicate that some circRNAs modulate immune factors (Yan and Chen, 2020). Under normal conditions, the complex of immune factors NF90 and NF110 promotes circRNAs production (Li *et al.*, 2017). Furthermore, cytoplasmatic circRNA that form double-stranded structures can trap the viral defense mediator PKR in the cytoplasm (Liu *et al.*, 2019). Upon viral infection, the NF90/NF110 complex shuttles from the nucleus to the cytoplasm (Li *et al.*, 2017), and RNaseL degrades circRNAs freeing PKR to translocate to the nucleus where it orchestrates a transcriptional immune response (Liu *et al.*, 2019). As a result, circRNA titer dramatically decreases upon viral infection. The fact that we detected an increase in levels of certain circRNAs upon Nsp14 expression suggest a possible mechanism for the immune evasion orchestrated by Nsp14.

As mentioned above, our analysis of the transcriptomes of HEK293T cells as a function of time after induction of Nsp14 expression revealed that the response is temporarily subdivided into two phases. The early response to Nsp14 expression involves NFkB activation, whereas the later phase involves effects on cellular metabolism, RNA metabolism, and cell-cycle control. Early after transfection with the Nsp14 expression vector, we detected increased *CXCL8* expression, which encodes IL8. IL8 is elevated in plasma of COVID-19 patients (Blanco-Melo *et al.*, 2020; Hariharan *et al.*, 2021; Li *et al.*, 2021) and is a marker of acute severe lung distress (Adcock *et al.*, 2015), an

inflammatory condition that can lead to severe complications or death. Use of a reporter assay showed that Nsp14 influences transcription of *CXCL8*.

We also detected a significant enrichment for NFkB binding sites in the promoters of genes that were differentially expressed at early time points after Nsp14 expression. NFkB is one of the first pathways activated upon SARS-CoV-2 infection (Li *et al.*, 2021; Wyler *et al.*, 2021). One activator of NFkB is PKR (Der and Lau, 1995; Clemens and Elia, 1997). It will be of interest to determine whether Nsp14 expression leads to PKR activation. Previously published RNA-seq data indicate that the interferon response anticipates NFkB activation during SARS-CoV-2 infection (Wyler *et al.*, 2021). In our analysis of the promoters of early upregulated genes, we also detected enrichment for binding sites of transcription factors involved in the interferon response. In support of our findings, recent reports showed that Nsp14 reduces the expression of INFb and ISG luciferase reporters (Yuen *et al.*, 2020), as well as the expression of proteins involved in the NFkB pathway (Hsu *et al.*, 2021; Li *et al.*, 2021).

It is unlikely thatNsp14 works as a transcription or spicing modulator factor directly, asit localizes to the cytoplasm (Gordon, Hiatt, *et al.*, 2020). Therefore, the effects of Nsp14 on the transcriptome are probably mediated by cellular proteins. Here we reported altered expression of genes involved in purine synthesis, including a decrease in *IMPDH2* mRNA levels. We also detected elevated cellular GTP concentrations upon Nsp14 expression. Furthermore, pharmacological inhibition of IMPDH2 activity with MPA mitigated the effects of Nsp14 expression. Previous studies showed that Nsp14 interacts with IMPDH2 (Gordon, Hiatt, *et al.*, 2020; Gordon, Jang, *et al.*, 2020). Intriguingly, MPA is a known immunosuppressant that has been used as antiviral agent (Pan *et al.*, 2012; To *et al.*, 2016; Dang *et al.*, 2017). It was also show to

inhibit SARS-CoV-2 propagation (Kato *et al.*, 2020). Therefore, IMPDH2 and Nsp14 are appealing targets for drug development.

Together, our data indicate that Nsp14 activates IMPDH2. It is unknown whether the effect is triggered by the direct interaction between the protein or if it requires a co-factor. Intriguingly, IMPDH2 binds nucleic acids (McLean *et al.*, 2004; Mortimer and Hedstrom, 2005) and acts as a transcription factor in *Drosophila* (Kozhevnikova *et al.*, 2012). However, although IMPDH2 protein is conserved (Hedstrom, 2009), the transcription factor activity of IMPDH2 has not been demonstrated in mammalian cells. Furthermore, IMPDH2 can associate with polyribosomes in mammalian cells and modulates the translation of specific mRNAs (Hsu *et al.*, 2021).

In the context of SARS-CoV-2 infection, a recent study showed that the activation of *CXCL8* expression and of NFkB transcriptional reporters induced by Nsp14 expression is mediated by IMPDH2 (Li *et al.*, 2021). Furthermore, another report indicates that Nsp14 blocks the interferon response, which is upstream of the NFkB pathway, by shutting down global translation (Hsu *et al.*, 2021). However, our data indicate that the transcriptional activation of the NFkB pathways occurs very early after transfection and that treatment with an IMPDH inhibitor partially rescues the long-term effects of Nsp14 expression in HEK293T cells. Therefore, we speculate that the Nsp14 can alter both gene expression and translation. In fact, the described translational effect is dependent on the ExoN activity of Nsp14, while our profound effect on gene expression is not, suggesting that our study tackles a different aspect of Nsp14 function. Moreover, we found that the transcription and the steady state mRNAs correlate very strongly upon Nsp14 expression, suggesting that the changes in the mRNA levels are consequence of transcriptional changes. Furthermore, we detected alterations on splicing and circRNAs levels, events which are not necessarily downstream to translational inhibition. Moreover, we showed that the effects on the

transcriptome are triggered soon after Nsp14 expression, as treatment with MPA 8 hours after transfection partially recovered the gene expression profile of the cells, suggesting that some of the gene expression changes might precede the global translation shutoff. Another possibility is that the transcriptome and translational effects depends on the levels of Nsp14 expression, that cannot be compared between the two studies. Finally, a compelling model would be that both effects are mediated by IMPDH2, considering our results on global transcriptome in cells expressing Nsp14 and treated with MPA, and the ability of IMPDH2 to associate with polyribosomes.

Methods

Cell Culture and transient plasmid transfection

HEK293T cells were a kind gift from Dr. Marr at Brandeis University. HEK293T were maintained in Dulbecco's Modified Eagle's Medium (Genesee Scientific, 25-500), supplemented with 10% serum (FetalPlex, Gemini Bio/Fisher, # 100-602), and 1X Penicillin-Streptomycin (Penicillin-Streptomycin 100x, Genesee Scientific, 25-512).

Plasmids pLVX-EF1alpha-SARS-CoV-2-proteins-2xStrep-IRES-Puro proteins are a gift from Nevan Krogan (Addgene)(Gordon, Jang, *et al.*, 2020) are listed in Table 14 and purified by Midi-prep using Invitrogen kit. The day before transfection, $0.35*10^6$ cells per well were plated in 12 well-plates. Then, cells were transfected with PEI (Polyethylenimine 25kD linear from Polysciences, cat# 23966-2). Briefly, 1500ng of total DNA (1200 ng of pLVX-EF1alpha-SARS-CoV-2-proteins-2xStrep-IRES-Pur + 300ng of pLVX-EF1alpha-eGFP-2xStrep-IRES-Puro) were incubated at RT for 20 min with PEI, with a ratio 3:1=PEI:DNA in 100ul of serum free medium, and the mixture was added to the cells. 48hrs after transfection, the medium was removed, and cells were harvested for following analysis.

MPA treatment

Mycophenolic acid (MPA) (Sigma Aldrich) was a kind gift from Hedstrom lab at Brandeis University and was reconstituted in DMSO (Sigma-Aldrich) to obtain 10 mM stock solutions. Aliquots were stored at −20 °C. MPA was freshly diluted into cell media. 8h after transfection, cells were treated with 50uM of MPA or with an equivalent amount of DMSO.

RNA extraction and qPCR

RNA was extracted with TriZol (Invitrogen) reagent following the manufacturer instructions and then treated with DnaseI (NEB). For the qPCRs showed in Figure S3A and Figure S3F, cDNA was prepared with 1000ng of DnaseI treated RNA (iScript Select cDNA synthesis Kit, BioRad. Random Priming, following manufacturer's instructions). The produced cDNA was utilized as a template for quantitative real-time PCR performed with the C1000 Thermal Cycler Bio-Rad. The PCR mixture contained Taq polymerase (SYBR green Bio-Rad). Cycling parameters were: 95°C for 3 min, followed by 40 cycles of 95°C for 10 s, 55°C for 10 s, and 72°C for 30 s. Fluorescence intensities were plotted versus the number of cycles by using an algorithm provided by the manufacturer. For the qPCRs showed in Figure S5A and Figure S5B cDNA was synthetized using 1000ng of RNA as input, 500ng of oligodT or random hexamers (for circRNA validation), 50 mM Tris-HCl pH 8.3, 50 mM KCl, 3 mM $MgCl_2$, 10 mM DTT, 400 µM dATP, 400 µM dCTP, 400 µM dGTP, 400 µM dTTP, and homemade purified MLV-RT (200 units). The produced cDNA was utilized as a template for quantitative real-time PCR performed with the C1000 Thermal Cycler Bio-Rad. The PCR mixture contained homemade purified Taq polymerase, 20mM Tris-HCl pH 8.8, 10mM $(NH_4)_2SO_4$, 10mM KCl, 2mM MgSo4, 0.1% Triton X-100, 1X SYBR, 1.075 M Ethylene glycol. Cycling parameters were: 95°C for 3 min, followed by 40 cycles of 95°C for 10 s, 55°C for 30 s, and 72°C for 30 s. Fluorescence intensities were plotted versus the number of cycles by using an algorithm provided by the manufacturer. All the primers used in this assay are listed in Table 15.

3' RNA sequencing libraries

75ng of RNA were used as input for preparing 3'RNA sequencing libraries following CelSeq2 (Hashimshony *et al.*, 2016) protocol, changing the UMI to 6 bases. Sequencing was performed on Illumina NextSeq 500 system.

Total RNA seq libraries

40ng of rRNA depleted RNA was used as a starting material for library preparation. NEBNext rRNA Depletion kit(NEB #E7405) was used to deplete ribosomal RNA. Libraries were prepared using the NEXTFLEX® Rapid Directional RNA-Seq Kit 2.0 by PerkinElmer, following the manufacturer's instructions. The samples were sequenced by Novogene (Novogene Corporation Inc. 8801 Folsom BLVD, Suite 290, Sacramento, CA 95826) with Hiseq-4000.

Metabolomic analysis

Cells were transfected as aforementioned, harvested with Trypsin-EDTA 0.25% 1X phenol red, Genesee Scientific (25-510) and washed 3 times with cold PBS. PBS was removed and samples were snap-frozen in liquid nitrogen. Samples were sent to Creative Proteomics 45-1 Ramsey Road, Shirley, NY 11967, USA for metabolomic analysis. Briefly, cell pellet in each sample tube was made into 200 μL of 80% methanol. Cells were lysed on a MM 400 mill mixer with the aid of two metal balls at 30 Hz for 1 min three times, followed by sonication for 3 min in an ice-water bath. The samples were placed at - 20° C for 1 hours before centrifugal clarification at 21,000 g and 5° C for 15 min. The clear supernatant was collected for LC-MS and the precipitated pellet was used for protein assay using a standardized Bradford procedure. Standard solutions containing all targeted nucleotides were prepared, in a concentration range of 0.0002 to 20 nmol/mL, in an internal standard solution of 13C5-AMP, 13C10-ATP and 13C10- GTP. 20 μL of the clear supernatant of each sample was mixed with 180 μL of the internal standard solution. 10-μL aliquots of resultant sample solutions and standard solutions were injected to a C18 column (2.1x150 mm, 1.7 μm) to run UPLC-MRM/MS on a Waters Acquity UPLC system coupled to a Sciex QTRAP 6500 Plus mass spectrometer operated in the negative-ion ESI mode. The mobile

phase was an acyl amine buffer (A) and acetonitrile/methanol (B) for binary gradient elution (5% to 65% B in 16 min), at 0.25 mL/min and 55° C. Concentrations of the detected analytes were calculated with internal-standard calibration by interpolating the constructed linear-regression curves of individual compounds, with the analyte-to-internal standard peak ratios measured from sample solutions.

Protein Extraction and Western Blot

Cells were lysed in RIPA buffer (150mM NaCl, 1% TritonX-100, 0.5% SDS, 50mM Tris-HCl pH 8.0) supplemented with EDTA free protease inhibitor (Pierce, PIA32955) and centrifuged at 4°C 21.1 x1000 g for 30 min. Lysate was quantified by BCA Assay (ThermoFisher, 23225). Lysates were run on 4-20% gradient polyAcrylamide gels (BioRad, 4568094) with Tris-Glycine SDS Running buffer. Transfer was performed in a Tris-Glycine buffer supplemented with 20% Methanol. Membranes were blocked with 3% BSA in TBS-T Buffer pH 7.6 (Tris base 0.2M, NaCl 1.5 M, 0.1% Tween-20). Membranes were incubated with the following primary antibodies: mouse anti-Actin (Cell sig, 3700), rabbit anti-IMPDH2 (Proteintec, 12948-1-AP), mouse anti-Strep-tag (Quiagen, 34850). Secondary antibodies: rabbit anti mouse - HRP conjugated (Millipore), mouse anti rabbit- HRP conjugated (Millipore). Detection was performed using ECL (Pierce, 32106).

Cloning of Nsp14 ExoN mutants

<u>pLXV-EIF1 alpha-2xStrep-SARS-CoV-2-nsp14-D90A/G92A-IRES-Puro</u>

pLXV-EF1alpha-2xStrep-SARS-CoV-2-nsp14-IRES-Puro was opened with BsrGI-HF (NEB) and EcoRI-HF (NEB). The removed cassette was replaced by ligation with a Gblock ordered from GenScript (GenScript UAS Inc. Piscataway, NJ) previously amplified with specific primers.

Gblock:

Gaattcgccgccaccatgtggtcccatccgcagtttgagaagggtggtggttcaggcggaggctccgggggcgggagctggtctcaccc
gcaatttgaaaaaggcgctgcggctgctgaaaatgtaacgggcttgtttaaagactgtagtaaagtgatcacaggactccaccccacacaag
cacctacccacctttccgtagatacgaagttcaaaacggaaggattgtgtgtggatataccagggataccaaaggatatgacgtaccgaagg
ctgatttccatgatgggttttaagatgaattaccaagttaatggctaccccaacatgttcatcaccagggaggaggcaattagacacgtaagag
cctggataggcttcGCCgttGCCggttgccatgcaacccgggaagccgtaggcacaaaccttccgttgcagcttggcttttccacgggc
gtcaacctcgttgccgtaccgactggctatgttgacacgccgaacaacaccgatttctctcgcgtaagtgctaagcctcctccgggagatcaa
ttcaagcatcttatacctctcatgtaca

Primers:

D90AG92A_F: cacacGaattcgccgccac

D90AG92A _R: cacacTGTACATGAGAGGTATAAGA

<u>pLXV-EIF1 alpha-2xStrep-SARS-CoV-2-nsp14-D273A-IRES-Puro</u>

pLXV-EF1alpha-2xStrep-SARS-CoV-2-nsp14-IRES-Puro was opened with BstBI (NEB) and AfeI (NEB). The removed cassette was replaced by ligation with a Gblock ordered from GenScript previously amplified with specific primers.

G-block:

CCACACttcgaacttacttctatgaaatactttgtaaaaatcgggcccgagcggacatgttgtttgtgcgaccgacgagctacttgttttag
cacagcatctgacacctacgcatgctggcaccacagtataggcttcgattacgtctacaatccctttatgatagacgtacaacaatggggcttt
acgggtaacttgcagagtaatcacgatctttactgccaagttcatgggaacgcacacgtggcctcctgcGCCgcgataatgacgaggtgc
ttggccgtgcacgagtgctttgttaagcgggtcgattggactatagagtatcccataatcggtgacgaacttaaaattaatgctgcatgcagga
aagtgcagcacatggtagtaaaagcagcgct.

Primers:

D273A_F: CCACACTTCGAACTTACTTCTATG

D273A_R: cacacagcgctgcttttactacc

Cloning of CXCL8-Firefly reporter

pGL_RSV_RF_BG (a modification of the pGL plasmid) was a kind gift from Dr. Marr at Brandeis University. pGL_RSV_RF_BG was opened with XhoI and NcoI. DNA was extracted from HEK293T cells with DNeasy Blood & Tissue Kit Quiagen (ID: 69504) according to manufacturer's instructions. The CXCL8 promoter was amplified by PCR with Phusion® High-Fidelity PCR Master Mix with HF Buffer (NEB) according to the manufacturer's instruction with the following primers:

CXCL8_promoter_R:

CACCACCATGGTGGCTAGCAGCTGGTACCCAGCTTGGACCTGGCTCTTGTCCTA

GAAGCT

CXCL8_promoter_F: CACCACTCGAGACAGTTCCTAGAAACTCTCTAAAATGCTTAGAA

The purified insert was digested with XhoI and NcoI and inserted into the plasmid by ligation with T7 DNA ligase (NEB) following the manufacturer's instructions to generate the pGL_CXCL8_promoter_Firefly.

Luciferase Assay

The day before transfection 0.18*106 cells per well were plated into 24 well plates. Cells were transfected as aforementioned using 750ng of total DNA. Cells were transfected with 75ng of

Firefly reporter, 75ng of Renilla reporter, and 600ng of pLVX-EF1alpha-SARS-CoV-2-Nsp14-2xStrep-IRES-Puro or 600ng of empty vector. We used the following plasmids as reporters:

pGL_CXCL8_promoter_Firefly, generated as indicated above.

pHAGE NFkB-TA-LUC-UBC-GFP-W was a gift from Darrell Kotton (Addgene plasmid # 49343 ; http://n2t.net/addgene:49343 ; RRID:Addgene_49343)(Wilson *et al.*, 2013)

7xE-Box::Renilla was a gift from Koen Venken (Addgene plasmid # 124532 ; http://n2t.net/addgene:124532 ; RRID:Addgene_124532, Sarrion-Perdigones et al., 2020)

At the indicated time point, cells were lysate in lysis buffer (25 mM Tris-phosphate at pH 7.8, 10% glycerol, 1% Triton X-100, 1 mg/ml of bovine serum albumin (BSA), 2 mM cyclohexylene diamin tetraacetate, and 2 mM DTT). An aliquot of the lysate was added to Firefly Buffer 1X (75 mM Tris pH 8.0,100 μM EDTA, 530 μM ATP, 5mM $MgSO_4$) freshly supplemented with 0.1 μM D-luciferin (Goldbio.com) and 100 μM Coenzyme-A (Fisher, #BP2510250), and luminescence was measured. Immediately after, an equal amount of Renilla Buffer 1X supplemented with 10 μM Coelenterazine (Goldbio.com) was added to the sample and luminescence was measured again. Non-transfected cell lysate was used to measure the background signal, that was subtracted before calculating the firefly/renilla ratios.

Bioinformatic analyses

Raw reads were aligned to the human genome (hg38) using STAR (Dobin *et al.*, 2013). Mapped reads in the 3' UTR were quantified using End Sequence Analysis Toolkit (Gohr and Irimia, 2019)for 3'RNA libraries. FeatureCounts (Gohr and Irimia, 2019) was used to quantify mapped transcripts from total RNA-seq libraries. DEGs were found using DESeq2 as previously described

(Sarrion-Perdigones, Gonzalez and Venken, 2020). DEGs with |L2FC| > 1, p-adjusted value <0.05 were considered significant and used as input in downstream Gene Ontology (GO) analysis (DAVID, v 6.8). For the analysis of the timepoint 3'RNA seq experiment, we used |L2FC| > 0.5, p-adjusted value <0.05. HEK293T cell transcriptome was used as a reference to query for Enriched GO terms from up- and down- regulated DEG lists. Gene set enrichment analysis (GSEA) was performed using the fgsea package in R (Sergushichev *et al.*, 2016). Gene rank was determined prior to GSEA by calculating -log(p-value)* sign(log2 fold change)(Sergushichev *et al.*, 2016). Gene sets used for GSEA were downloaded from Molecular Signatures Database (http://www.msigdb.org). To compare our RNA-seq results with other studies we used a gene set from MSigDB containing DEGs from published RNA-seq data. A curated list of terms from GSEA were plotted to assess the transcriptional response to viral protein expression. Data visualization was carried out using ggplot2 in R. Non-omics statistical analysis and data visualization was completed using Prism.

circRNA detection and differential expression analysis

Computational analysis of circRNAs was performed on total RNA-seq data. circRNAs were detected by searching unaligned reads for head-to-tail splice junctions as previously described (Memczak *et al.*, 2013) . Differentially expressed circRNAs were found by DESeq2 (Love, Huber and Anders, 2014) . circRNA reads were normalized using size factors computed by DESeq2 together with all mapped reads.

Intronic reads quantification and analysis

Reads were aligned with STAR-aligner(Wilson *et al.*, 2013) to the human genome and transcriptome version Hg38. Intronic reads were extracted and counted using Featurecounts function in R with an intronic region reference.

Alternative splicing analysis

Percentage of inclusion (PSI) quantification was done with Vast-Tools (Tapial *et al.*, 2017). Delta PSI was then calculated from the mean of each condition. To determine events changing between Nsp14 and transfection control we chose a minimum of 15% difference in mean PSI and no overlapping between replicas. Introns and exons feature analysis was done using Matt (Gohr and Irimia, 2019)

Motif enrichment analysis

We extracted the promoters' sequences of upregulated genes with a fold change of at least 1.75 (log2FC > 0.8 and adjusted pvalue > 0.05) using windows of 100, 500 and 1000 bases upstream of each annotated transcription start site for each gene. We then used the MEMEsuite (Bailey *et al.*, 2009) using the tool AME(McLeay and Bailey, 2010) for motif enrichment analysis. We used the motif data base HOmo sapiens COmprehensive MOdel COllection (HOCOMOCO) v11 transcription factor (TF) binding models (binding profiles or binding motifs) for human transcription factors. As control for the input sequences, we used the tool provided for scrambled sequences option. Finally, we used MEMEsuit tool FIMO (Grant, Bailey and Noble, 2011) to identify the position of these motifs in the analyzed 100bp promoters.We extracted the promoters' sequences of upregulated genes with a fold change of at least 1.75 (log2FC > 0.8 and adjusted pvalue > 0.05) using windows of 100, 500 and 1000 bases upstream of each annotated transcription start site for each gene. We then used the MEMEsuite tool AME(McLeay and Bailey, 2010) for motif enrichment analysis using. We used the motif data base HOmo sapiens COmprehensive MOdel COllection (HOCOMOCO) v11 transcription factor (TF) binding models (binding profiles or binding motifs) for human transcription factors. As control for the input sequences, we used the tool provided scrambled sequences option. Finally. we used MEMEsuit

tool FIMO(Grant, Bailey and Noble, 2011) to identify the position of these motifs in the analyzed 100bp promoters.

List of the tables

Table 1: List of DEGs in the screening described in Figure 1. 3'RNA seq.

Table 2: List of GO terms described in Figure 1. 3' RNA seq.

Table 3: GSEA analysis described in Figure 2. Total RNA seq.

Table 4: Differential gene expression results. Described in figure 2.

Table 5: Alternative Splicing analysis described in Figure 3.

Table 6: Feature analysis with MATT for upregulated intron retentions and downregulated exon. inclusion events described in figure Figure S2.

Table 7: Deregulated circRNAs described in Figure 3.

Table 8: DEG at different time points described in Figure 5. 3' RNA seq.

Table 9: GO described in Figure 5. 3'RNA seq.

Table 10: Ame analysis with 100bp of promoter described in Figure 5.

Table 11: Fimo analysis with 100bp promoter described in Figure 5.

Table 12: Ame anslysis with 1000bp promoter.

Table 13: Concentrations of cellular metabolites detected by HPLC/MS.

Table 14: list of plasmids used for SARS-CoV-2 protein expression in the initial screening described in Figure 1.

Table 15: list of primers used for the qPCR assays.

Figure Legends

Figure 1: SARS-CoV-2 proteins alter gene expression distinctively. A) Scheme of the experimental approach. **B)** Heatmap showing the number of DEGs detected in 3' RNA sequencing for each expressed SARS-CoV-2 protein. "Up" stands for upregulated genes, "Down" for downregulated genes. lfc<0.5, corrected p-value<0.05. C**)** Heatmap showing the Gene Ontology analysis.

Figure 2: Overexpression of Nsp14 induces transcriptional changes like SARS-CoV-2 infection. A) Volcano plot showing the fold change of expression in samples overexpressing Nsp14 compared to control detected in the Total RNA sequencing. In red significantly downregulated genes, in blue upregulated genes, and in grey non significantly deregulated genes. **B)** Table reporting the GSEA terms, adjusted p-value, and NES when comparing our Total RNA sequencing data with previously published datasets. **C)** Example of GSEA analysis. **D)** Nsp14 expression vs control Fold Change of intronic signal from total RNA seq vs 3' RNA seq signal in logarithmic scale for each detected gene. Colored dots represent significantly changing genes (Fold change = 2, adjusted p-value = 0.05, N=3). **E)** Nsp14 expression vs control Fold Change of intronic signal from total RNA seq vs exonic signal from total RNA seq in logarithmic scale for each detected gene. Colored dots represent significantly changing genes (Fold change = 2, adjusted p-value = 0.05, N=3).

Figure 3: Nsp14 expression alters the splicing of a subgroup of genes and increases circRNAs expression. A) Table summarizing splicing analysis comparison between Nsp14 expression and control. Thresholds used: ΔPSI > 15% and a non-overlapping distribution with minimum of 5% difference (N=3). **B)** Fold Change vs Expression in logarithmic scale for the genes with upregulated intron retention. In red genes with increased expression and in blue the ones with

downregulated expression. (Foldchange = 2, adjusted p-value = 0.05, N=3). **C)** Representative IGV alignment tracks of on gene (PAXIP1) with intronic events differentially changing between conditions (Control and Nsp14expression). The box marks the changing event. On the right, quantification of PSI. **D)** Pie chart representing number of alternative splicing events deregulated upon Nsp14 expression by gene. 1772 genes have only one alternative splicing event changing between conditions, 615 has 2 events and 243 genes have 3 alternative splicing events changing. **E)** Number of circRNAs reads detected in the Total RNA sequencing experiment. Data represented as mean ± SEM, N=3, t-test, *** p-value<0.0005 **F)** Fold Change vs Expression in logarithmic scale for circRNAs in Nsp14 expression vs control. In red upregulated genes and in blue downregulated genes (Foldchange = 2, adjusted p-value = 0.05, N=3). **G)** Plot of Fold Change vs Expression in logarithmic scale for exonic signal detected in the Total RNA sequencing dataset in Nsp14 vs control for genes with upregulated circRNA expression. In red genes with increased expression and (Foldchange = 2, adjusted p-value = 0.05, N=3). **H)** Plot of Fold Change vs Expression in logarithmic scale for intronic signal detected in the Total RNA sequencing dataset in Nsp14 vs control for genes with upregulated circRNA expression. In red genes with increased expression and in grey non-significant ones (Foldchange = 2, adjusted p-value = 0.05, N=3).

Figure 4: The co-factor Nsp10 and the ExoN activity of Nsp14 are not required for changing gene expression. A) Western Blot in cells transfected with an empty plasmid (control), Nsp10, Nsp14, or co-expressing Nsp10 and Nsp14 (Nsp10+Nsp14). Nsp10 and Nsp14 were detected through the Strep tag. Actin was used as loading control. **B)** Volcano plots showing the expression fold change in between the indicated conditions in the Total RNA seq dataset. Significantly upregulated genes in blue, downregulated in red, and not significantly deregulated in grey. **C)** Plot showing the Fold change of deregulated genes in samples co-expressing Nsp10 and Nsp14 vs

control (on the y-axis) and Nsp14 vs control (on the x-axis). **D)** Number of circRNAs reads detected in each indicated condition. Data represented as mean ± SEM, N=3. **E)** Schematic of Nsp14 ExoN and N7-MTase domains and the tested mutations. **F)** Volcano plots showing the expression fold change in between the indicated conditions in 3' RNA seq. Significantly upregulated genes in blue, downregulated in red, and not significantly deregulated in grey. **G)** Plots showing the Fold change of deregulated genes in the indicated condition vs control.

Figure 5: NFkB pathway is activated soon after Nsp14 transfection. A) Principal component analysis of the 3' RNA seq of the time course experiment. Arrows indicate how samples separates according to the time point or condition (Nsp14 expression or control). **B)** Heatmap showing increasing number of up and downregulated genes at different time points after Nsp14 overexpression. **C)** Heatmap representing the Gene Ontology Analysis result at the indicated timepoints. **D)** Expression of CXCL8 across the indicated time points in the 3'RNA seq data. Data represented as mean ± SEM, N =3,t-test, ** p-value<0.005, * p-value<0.05 **E)** Luciferase assay showing that CXCL8 is transcriptionally activated after Nsp14expression. Firefly expression is controlled by CXCL8 promoter, whereas renilla is under the control of a constitutive promoter. Data represented as mean ± SEM, N =6, t-test, **p-value<0.005. **F)** Table showing the enrichment for specific TF binding sites in the promoter (100 bp upstream the transcription starting site of the upregulated genes (lfc >0.8, adjusted p-value < 0.05) at 8h after Nsp14 transfection.

Figure 6: Pharmacological inhibition of IMPDH2 partially reverts the transcriptional changes induced by Nsp14. A) Expression of IMPDH2 is reduced upon Nsp14expression. Data from the Total RNA seq experiment. Data represented as mean ± SEM, N=3, t-test, ***p-value<0.0005. **B)** In the upper panel, scheme reporting some of the tested metabolites deriving from IMP metabolism. IMPDH2 (highlighted in light blue) catalyzes the conversion of IMP to

XMP, precursor of GTP. Significantly upregulated metabolites are highlighted in red. In the lower panel, GTP cellular concentration significantly increases in Nsp14 overexpressing cells. Data represented as mean ± SEM, N=3, t-test, **p-value<0.005. **C)** Western Blot showing that MPA treatment does not alter Nsp14 (detected through the Strep tag) or IMPDH2 expression. Actin used as loading control. **D)** Principal component Analysis of the 3' RNA seq library of the indicated samples. **E)** Table reporting the number of upregulated and downregulated genes in the indicated comparisons. **F)** Plot showing the distribution of Fold changes of all genes detected in the 3' RNA seq in the indicated conditions.

Figure S1: A) Plot showing the overlap between the 3'RNA Seq dataset and the Total RNA seq dataset. **B)** GSEA analysis showing that Nsp14 expression resemble SARS-CoV-2 infection.

Figure S2: A) Representative IGV alignment tracks of on gene (HJURP) with multiple exonic events differentially changing between conditions (Control and Nsp14expression). The box marks the changing event. On the right, quantification of PSI. **B)** Kernel density plot for 6 intronic features that are significantly changing (Mann Whitney U test, adjusted p-value < 0.05, N=3) in upregulated intron retention upon Nsp14expression. Thresholds used: ΔPSI > 15% and a non-overlapping distribution with minimum of 5% difference (n=3). **C)** Kernel density plot of intron GC content. In grey the distribution for internal introns (never in first position). In blue introns that appear 10% of the times at first position and in red introns that are in first position in more than 10% of the transcripts. **D)** Average gene expression in Nsp14 expression vs control in logarithmic scale (N=3). In grey genes with no intron retained, in color genes with upregulated intron retention. Color represents the proportion of transcripts in which the intron retained is in first position. **E)** Kernel density plot for 6 exonic features that are significantly changing (Mann Whitney U test,

adjusted p-value < 0.05, N=3) in downregulated exon inclusion upon Nsp14expression. Thresholds used: ΔPSI > 15% and a non-overlapping distribution with minimum of 5% difference (N=3).

Figure S3: A) qPCR showing the expression of FGF-18 and SH2D2A in the indicated conditions. Data represented as mean ± SEM, N=3. **B)** PCA for the 3' RNA seq of the indicated samples. **C)** Plot showing the Fold change of deregulated genes in samples expressing Nsp10 vs control (on the y-axis) and Nsp14 vs control (on the x-axis). **D)** PCA for the 3' RNA seq of the indicated samples. **E)** Volcano plots showing the expression fold change in between the indicated conditions in 3' RNA seq. Significantly upregulated genes in blue, downregulated in red, and not significantly deregulated in grey. **F)** qPCR showing the expression of circ-CDK1, circMARCHF7, and circVKR1 in the indicated conditions. Data represented as mean ± SEM, N=3.

Figure S4: A) Volcano plots showing the expression fold change in between the indicated conditions in 3' RNA seq. Significantly upregulated genes in blue, downregulated in red, and not significantly deregulated in grey. **B)** Expression of selected genes across the indicated time points from the 3'RNA seq dataset. Data represented as mean ± SEM, N=3, error.**C)** Scheme of the CXCL8 promoter region used for the firefly reporter. **D)** luciferase assay. Firefly expression is controlled by a minimal transcriptional activator recognized by NFkB, whereas renilla is under the control of a ubiquitous promoter. Data represented as mean ± SEM, N=6, t-test, **p-value<0.005.

Figure S5: A) Expression of IMPDH2, FGF-18, SH2D2A, and CXCL8 in the 3' RNA seq dataset in the indicated conditions. **B)** Venn diagram showing the common genes upregulated (left) or downregulated (right) in samples expressing Nsp14 and treated with or without MPA.

C) qPCR showing the expression of circ-CDK1, circMARCHF7, and circVKR1 in the indicated conditions. Data represented as mean ± SEM, N=3.

Competing interests

The authors declare no competing interests.

Founding

This work was supported by the NIH R01 Grant R01AG057700 to SK and NIH NIGMS P41 GM132087 to NP. NP is an investigator of HHMI.

Acknowledgments

We thank the Hedstrom Lab, the Marr lab, and the Rosbash Lab at Brandeis University for sharing reagents, instruments, and for helpful discussion. Thanks Dr. tenOever at Icahn School of Medicine at Mount Sinai for critical discussion of the data. Thanks to Dr. Aarti Jajoo for the initial steps of the screening data analysis. We warmly thank all the members of Kadener lab for helpful and critical discussion.

Authors' contribution

M.Z. performed most of the experiments, designed the experiments, and wrote the manuscript. J.M.H. and I.L.P. performed the bioinformatic analysis. I.L.P. helped drafting the manuscript. N.R.P. performed most of the RNA extractions, qPCRs, and prepared the total RNA libraries. S.N. prepared most of the 3'RNA libraries. S.K. wrote the manuscript, designed the experiments, and supervised the project.

References

Adcock, I. M. *et al.* (2015) 'Severity of acute respiratory distress syndrome resulting from tuberculosis correlates with bronchoalveolar lavage CXCL-8 expression', *International Journal of Mycobacteriology*, 4, p. 71. doi: https://doi.org/10.1016/j.ijmyco.2014.11.039.

Bachmayr-Heyda, A. *et al.* (2015) 'Correlation of circular RNA abundance with proliferation – exemplified with colorectal and ovarian cancer, idiopathic lung fibrosis and normal human tissues', *Scientific Reports*, 5(1), p. 8057. doi: 10.1038/srep08057.

Bailey, T. L. *et al.* (2009) 'MEME SUITE: tools for motif discovery and searching.', *Nucleic acids research*, 37(Web Server issue), pp. W202-8. doi: 10.1093/nar/gkp335.

Baldwin, A. S. J. (1996) 'The NF-kappa B and I kappa B proteins: new discoveries and insights.', *Annual review of immunology*, 14, pp. 649–683. doi: 10.1146/annurev.immunol.14.1.649.

Banerjee, A. K. *et al.* (2020) 'SARS-CoV-2 Disrupts Splicing, Translation, and Protein Trafficking to Suppress Host Defenses', *Cell*, 183(5), pp. 1325-1339.e21. doi: https://doi.org/10.1016/j.cell.2020.10.004.

Blanco-Melo, D. *et al.* (2020) 'Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19.', *Cell*, 181(5), pp. 1036-1045.e9. doi: 10.1016/j.cell.2020.04.026.

Bouvet, M. *et al.* (2012) 'RNA 3'-end mismatch excision by the severe acute respiratory syndrome coronavirus nonstructural protein nsp10/nsp14 exoribonuclease complex.', *Proceedings of the National Academy of Sciences of the United States of America*, 109(24), pp. 9372–9377. doi: 10.1073/pnas.1201130109.

Bouvet, M. *et al.* (2014) 'Coronavirus Nsp10, a critical co-factor for activation of multiple replicative enzymes.', *The Journal of biological chemistry*, 289(37), pp. 25783–25796. doi: 10.1074/jbc.M114.577353.

Chen, T.-C. *et al.* (2020) 'Host-derived circular RNAs display proviral activities in Hepatitis C virus-infected cells.', *PLoS pathogens*, 16(8), p. e1008346. doi: 10.1371/journal.ppat.1008346.

Chen, Y. *et al.* (2009) 'Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase', *Proceedings of the National Academy of Sciences*, 106(9), pp. 3484 LP – 3489. doi: 10.1073/pnas.0808790106.

Clemens, M. J. and Elia, A. (1997) 'The double-stranded RNA-dependent protein kinase PKR: structure and function.', *Journal of interferon & cytokine research : the official journal of the International Society for Interferon and Cytokine Research*, 17(9), pp. 503–524. doi: 10.1089/jir.1997.17.503.

Coperchini, F. *et al.* (2020) 'The cytokine storm in COVID-19: An overview of the involvement of the chemokine/chemokine-receptor system.', *Cytokine & growth factor reviews*, 53, pp. 25–32. doi: 10.1016/j.cytogfr.2020.05.003.

Dang, W. *et al.* (2017) 'Inhibition of Calcineurin or IMP Dehydrogenase Exerts Moderate to Potent Antiviral Activity against Norovirus Replication.', *Antimicrobial agents and chemotherapy*, 61(11). doi: 10.1128/AAC.01095-17.

Denison, M. R. *et al.* (2011) 'Coronaviruses: an RNA proofreading machine regulates replication fidelity and diversity.', *RNA biology*, 8(2), pp. 270–279. doi: 10.4161/rna.8.2.15013.

Der, S. D. and Lau, A. S. (1995) 'Involvement of the double-stranded-RNA-dependent kinase PKR in interferon expression and interferon-mediated antiviral activity.', *Proceedings of the National Academy of Sciences of the United States of America*, 92(19), pp. 8841–8845. doi: 10.1073/pnas.92.19.8841.

Dobin, A. *et al.* (2013) 'STAR: ultrafast universal RNA-seq aligner.', *Bioinformatics (Oxford,*

England), 29(1), pp. 15–21. doi: 10.1093/bioinformatics/bts635.
Eckerle, L. D. *et al.* (2007) 'High Fidelity of Murine Hepatitis Virus Replication Is Decreased in nsp14 Exoribonuclease Mutants', *Journal of Virology*, 81(22), pp. 12135–12144. doi: 10.1128/JVI.01296-07.
Eckerle, L. D. *et al.* (2010) 'Infidelity of SARS-CoV Nsp14-exonuclease mutant virus replication is revealed by complete genome sequencing.', *PLoS pathogens*, 6(5), p. e1000896. doi: 10.1371/journal.ppat.1000896.
Egloff, M.-P. *et al.* (2004) 'The severe acute respiratory syndrome-coronavirus replicative protein nsp9 is a single-stranded RNA-binding subunit unique in the RNA virus world', *Proceedings of the National Academy of Sciences of the United States of America*, 101(11), pp. 3792 LP – 3796. doi: 10.1073/pnas.0307877101.
Gardinassi, L. G. *et al.* (2020) 'Immune and Metabolic Signatures of COVID-19 Revealed by Transcriptomics Data Reuse', *Frontiers in Immunology*, 11, p. 1636. doi: 10.3389/fimmu.2020.01636.
Glesne, D. A., Collart, F. R. and Huberman, E. (1991) 'Regulation of IMP dehydrogenase gene expression by its end products, guanine nucleotides', *Molecular and Cellular Biology*, 11(11), pp. 5417–5425. doi: 10.1128/mcb.11.11.5417-5425.1991.
Gohr, A. and Irimia, M. (2019) 'Matt: Unix tools for alternative splicing analysis.', *Bioinformatics (Oxford, England)*, 35(1), pp. 130–132. doi: 10.1093/bioinformatics/bty606.
Gordon, D. E., Jang, G. M., *et al.* (2020) 'A SARS-CoV-2 protein interaction map reveals targets for drug repurposing.', *Nature*, 583(7816), pp. 459–468. doi: 10.1038/s41586-020-2286-9.
Gordon, D. E., Hiatt, J., *et al.* (2020) 'Comparative host-coronavirus protein interaction networks reveal pan-viral disease mechanisms', *Science*, 370(6521), p. eabe9403. doi: 10.1126/science.abe9403.
Graham, R. L. *et al.* (2012) 'A live, impaired-fidelity coronavirus vaccine protects in an aged, immunocompromised mouse model of lethal disease.', *Nature medicine*, 18(12), pp. 1820–1826. doi: 10.1038/nm.2972.
Grant, C. E., Bailey, T. L. and Noble, W. S. (2011) 'FIMO: scanning for occurrences of a given motif.', *Bioinformatics (Oxford, England)*, 27(7), pp. 1017–1018. doi: 10.1093/bioinformatics/btr064.
Hariharan, A. *et al.* (2021) 'The Role and Therapeutic Potential of NF-kappa-B Pathway in Severe COVID-19 Patients.', *Inflammopharmacology*, 29(1), pp. 91–100. doi: 10.1007/s10787-020-00773-9.
Hashimshony, T. *et al.* (2016) 'CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq', *Genome Biology*, 17(1), p. 77. doi: 10.1186/s13059-016-0938-8.
Hedstrom, L. (2009) 'IMP dehydrogenase: structure, mechanism, and inhibition.', *Chemical reviews*, 109(7), pp. 2903–2928. doi: 10.1021/cr900021w.
Hesketh, A. and Oliver, S. G. (2019) 'High-energy guanine nucleotides as a signal capable of linking growth to cellular energy status via the control of gene transcription', *Current Genetics*, 65(4), pp. 893–897. doi: 10.1007/s00294-019-00963-1.
Hillen, H. S. *et al.* (2020) 'Structure of replicating SARS-CoV-2 polymerase', *Nature*, 584(7819), pp. 154–156. doi: 10.1038/s41586-020-2368-8.
Hsu, J. C.-C. *et al.* (2021) 'Translational shutdown and evasion of the innate immune response by SARS-CoV-2 NSP14 protein', *Proceedings of the National Academy of Sciences*, 118(24), p. e2101161118. doi: 10.1073/pnas.2101161118.
Kato, F. *et al.* (2020) 'Antiviral activities of mycophenolic acid and IMD-0354 against SARS-

CoV-2', *Microbiology and Immunology*, 64(9), pp. 635–639. doi: https://doi.org/10.1111/1348-0421.12828.
Kim, D. *et al.* (2020) 'The Architecture of SARS-CoV-2 Transcriptome.', *Cell*, 181(4), pp. 914-921.e10. doi: 10.1016/j.cell.2020.04.011.
Kirchdoerfer, R. N. and Ward, A. B. (2019) 'Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors', *Nature Communications*, 10(1), p. 2342. doi: 10.1038/s41467-019-10280-3.
Kircheis, R. *et al.* (2020) 'NF-κB Pathway as a Potential Target for Treatment of Critical Stage COVID-19 Patients', *Frontiers in Immunology*, 11, p. 3446. doi: 10.3389/fimmu.2020.598444.
Konno, Y. *et al.* (2020) 'SARS-CoV-2 ORF3b Is a Potent Interferon Antagonist Whose Activity Is Increased by a Naturally Occurring Elongation Variant', *Cell Reports*, 32(12), p. 108185. doi: https://doi.org/10.1016/j.celrep.2020.108185.
Kozhevnikova, E. N. *et al.* (2012) 'Metabolic enzyme IMPDH is also a transcription factor regulated by cellular state.', *Molecular cell*, 47(1), pp. 133–139. doi: 10.1016/j.molcel.2012.04.030.
Kunsch, C. and Rosen, C. A. (1993) 'NF-kappa B subunit-specific regulation of the interleukin-8 promoter.', *Molecular and cellular biology*, 13(10), pp. 6137–6146. doi: 10.1128/mcb.13.10.6137-6146.1993.
Lee, Stuart *et al.* (2020) 'Covering all your bases: incorporating intron signal from RNA-seq data', *NAR Genomics and Bioinformatics*, 2(3). doi: 10.1093/nargab/lqaa073.
Lee, Seungheon *et al.* (2020) 'Mycophenolic anilides as broad specificity inosine-5'-monophosphate dehydrogenase (IMPDH) inhibitors.', *Bioorganic & medicinal chemistry letters*, 30(24), p. 127543. doi: 10.1016/j.bmcl.2020.127543.
Li, T. *et al.* (2021) 'SARS-CoV-2 Nsp14 activates NF-κB signaling and induces IL-8 upregulation', *bioRxiv*. doi: 10.1101/2021.05.26.445787.
Li, X. *et al.* (2017) 'Coordinated circRNA Biogenesis and Function with NF90/NF110 in Viral Infection.', *Molecular cell*, 67(2), pp. 214-227.e7. doi: 10.1016/j.molcel.2017.05.023.
Liao, Q.-J. *et al.* (2005) 'Activation of NF-kappaB by the full-length nucleocapsid protein of the SARS coronavirus.', *Acta biochimica et biophysica Sinica*, 37(9), pp. 607–612. doi: 10.1111/j.1745-7270.2005.00082.x.
Lin, C.-L., Taggart, A. J. and Fairbrother, W. G. (2016) 'RNA structure in splicing: An evolutionary perspective.', *RNA biology*, 13(9), pp. 766–771. doi: 10.1080/15476286.2016.1208893.
Liu, C.-X. *et al.* (2019) 'Structure and Degradation of Circular RNAs Regulate PKR Activation in Innate Immunity', *Cell*, 177(4), pp. 865-880.e21. doi: https://doi.org/10.1016/j.cell.2019.03.046.
Love, M. I., Huber, W. and Anders, S. (2014) 'Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.', *Genome biology*, 15(12), p. 550. doi: 10.1186/s13059-014-0550-8.
Ma, Y. *et al.* (2015) 'Structural basis and functional analysis of the SARS coronavirus nsp14-nsp10 complex.', *Proceedings of the National Academy of Sciences of the United States of America*, 112(30), pp. 9436–9441. doi: 10.1073/pnas.1508686112.
McLean, J. E. *et al.* (2004) 'Inosine 5'-monophosphate dehydrogenase binds nucleic acids in vitro and in vivo.', *The Biochemical journal*, 379(Pt 2), pp. 243–251. doi: 10.1042/BJ20031585.
McLeay, R. C. and Bailey, T. L. (2010) 'Motif Enrichment Analysis: a unified framework and an evaluation on ChIP data.', *BMC bioinformatics*, 11, p. 165. doi: 10.1186/1471-2105-11-165.

Memczak, S. *et al.* (2013) 'Circular RNAs are a large class of animal RNAs with regulatory potency.', *Nature*, 495(7441), pp. 333–338. doi: 10.1038/nature11928.
Mortimer, S. E. and Hedstrom, L. (2005) 'Autosomal dominant retinitis pigmentosa mutations in inosine 5'-monophosphate dehydrogenase type I disrupt nucleic acid binding.', *The Biochemical journal*, 390(Pt 1), pp. 41–47. doi: 10.1042/BJ20042051.
Ogando, N. S. *et al.* (2020) 'The Enzymatic Activity of the nsp14 Exoribonuclease Is Critical for Replication of MERS-CoV and SARS-CoV-2.', *Journal of virology*, 94(23). doi: 10.1128/JVI.01246-20.
Pahl, H. L. (1999) 'Activators and target genes of Rel/NF-kappaB transcription factors.', *Oncogene*, 18(49), pp. 6853–6866. doi: 10.1038/sj.onc.1203239.
Pan, Q. *et al.* (2012) 'Mycophenolic acid augments interferon-stimulated gene expression and inhibits hepatitis C Virus infection in vitro and in vivo.', *Hepatology (Baltimore, Md.)*, 55(6), pp. 1673–1683. doi: 10.1002/hep.25562.
Park, J. H. and Lee, H. K. (2020) 'Re-analysis of Single Cell Transcriptome Reveals That the NR3C1-CXCL8-Neutrophil Axis Determines the Severity of COVID-19', *Frontiers in Immunology*, 11, p. 2145. doi: 10.3389/fimmu.2020.02145.
Robson, F. *et al.* (2020) 'Coronavirus RNA Proofreading: Molecular Basis and Therapeutic Targeting', *Molecular Cell*, 79(5), pp. 710–727. doi: https://doi.org/10.1016/j.molcel.2020.07.027.
Sanjuán, R. *et al.* (2010) 'Viral Mutation Rates', *Journal of Virology*, 84(19), pp. 9733–9748. doi: 10.1128/JVI.00694-10.
Sarrion-Perdigones, A., Gonzalez, Y. and Venken, K. J. T. (2020) 'Rapid and Efficient Synthetic Assembly of Multiplex Luciferase Reporter Plasmids for the Simultaneous Monitoring of Up to Six Cellular Signaling Pathways.', *Current protocols in molecular biology*, 131(1), p. e121. doi: 10.1002/cpmb.121.
Schubert, K. *et al.* (2020) 'SARS-CoV-2 Nsp1 binds the ribosomal mRNA channel to inhibit translation', *Nature Structural & Molecular Biology*, 27(10), pp. 959–966. doi: 10.1038/s41594-020-0511-8.
Sergushichev, A. A. *et al.* (2016) 'GAM: a web-service for integrated transcriptional and metabolic network analysis.', *Nucleic acids research*, 44(W1), pp. W194-200. doi: 10.1093/nar/gkw266.
Shu, T. *et al.* (2020) 'SARS-Coronavirus-2 Nsp13 Possesses NTPase and RNA Helicase Activities That Can Be Inhibited by Bismuth Salts.', *Virologica Sinica*, 35(3), pp. 321–329. doi: 10.1007/s12250-020-00242-1.
Smith, E. C. *et al.* (2015) 'Mutations in Coronavirus Nonstructural Protein 10 Decrease Virus Replication Fidelity', *Journal of Virology*, 89(12), pp. 6418–6426. doi: 10.1128/JVI.00110-15.
Stein, B. *et al.* (1993) 'Cross-coupling of the NF-kappa B p65 and Fos/Jun transcription factors produces potentiated biological function.', *The EMBO journal*, 12(10), pp. 3879–3891.
Tapial, J. *et al.* (2017) 'An atlas of alternative splicing profiles and functional associations reveals new regulatory programs and genes that simultaneously express multiple major isoforms.', *Genome research*, 27(10), pp. 1759–1768. doi: 10.1101/gr.220962.117.
To, K. K. W. *et al.* (2016) 'Mycophenolic acid, an immunomodulator, has potent and broad-spectrum in vitro antiviral activity against pandemic, seasonal and avian influenza viruses affecting humans', *Journal of General Virology*, 97(8), pp. 1807–1817. doi: https://doi.org/10.1099/jgv.0.000512.
V'kovski, P. *et al.* (2021) 'Coronavirus biology and replication: implications for SARS-CoV-2',

Nature Reviews Microbiology, 19(3), pp. 155–170. doi: 10.1038/s41579-020-00468-6.
Del Valle, D. M. *et al.* (2020) 'An inflammatory cytokine signature predicts COVID-19 severity and survival.', *Nature medicine*, 26(10), pp. 1636–1643. doi: 10.1038/s41591-020-1051-9.
Vanderheiden, A. *et al.* (2020) 'Type I and Type III Interferons Restrict SARS-CoV-2 Infection of Human Airway Epithelial Cultures.', *Journal of virology*, 94(19). doi: 10.1128/JVI.00985-20.
Wang, W. *et al.* (2007) 'Up-regulation of IL-6 and TNF-alpha induced by SARS-coronavirus spike protein in murine macrophages via NF-kappaB pathway.', *Virus research*, 128(1–2), pp. 1–8. doi: 10.1016/j.virusres.2007.02.007.
Wilson, A. A. *et al.* (2013) 'Lentiviral delivery of RNAi for in vivo lineage-specific modulation of gene expression in mouse lung macrophages.', *Molecular therapy : the journal of the American Society of Gene Therapy*, 21(4), pp. 825–833. doi: 10.1038/mt.2013.19.
Wyler, E. *et al.* (2021) 'Transcriptomic profiling of SARS-CoV-2 infected human cell lines identifies HSP90 as target for COVID-19 therapy.', *iScience*, 24(3), p. 102151. doi: 10.1016/j.isci.2021.102151.
Yan, L. and Chen, Y. G. (2020) 'Circular RNAs in Immune Response and Viral Infection', *Trends in Biochemical Sciences*, 45(12), pp. 1022–1034. doi: https://doi.org/10.1016/j.tibs.2020.08.006.
Yuen, C.-K. *et al.* (2020) 'SARS-CoV-2 nsp13, nsp14, nsp15 and orf6 function as potent interferon antagonists.', *Emerging microbes & infections*, 9(1), pp. 1418–1428. doi: 10.1080/22221751.2020.1780953.
Zhang, J., Kuo, C. C. J. and Chen, L. (2011) 'GC content around splice sites affects splicing through pre-mRNA secondary structures.', *BMC genomics*, 12, p. 90. doi: 10.1186/1471-2164-12-90.
Zhang, L. *et al.* (2004) 'GC/AT-content spikes as genomic punctuation marks.', *Proceedings of the National Academy of Sciences of the United States of America*, 101(48), pp. 16855–16860. doi: 10.1073/pnas.0407821101.

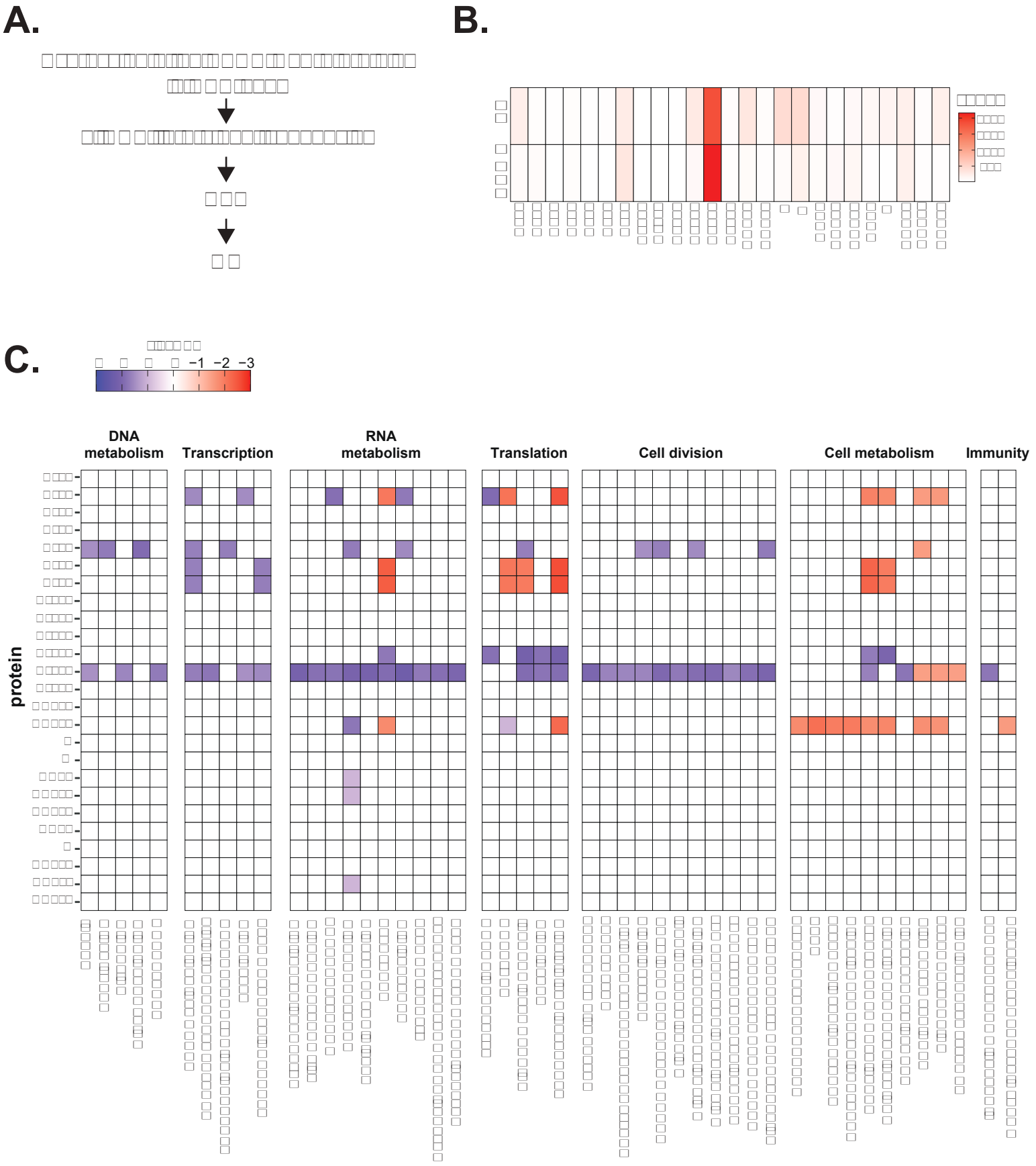
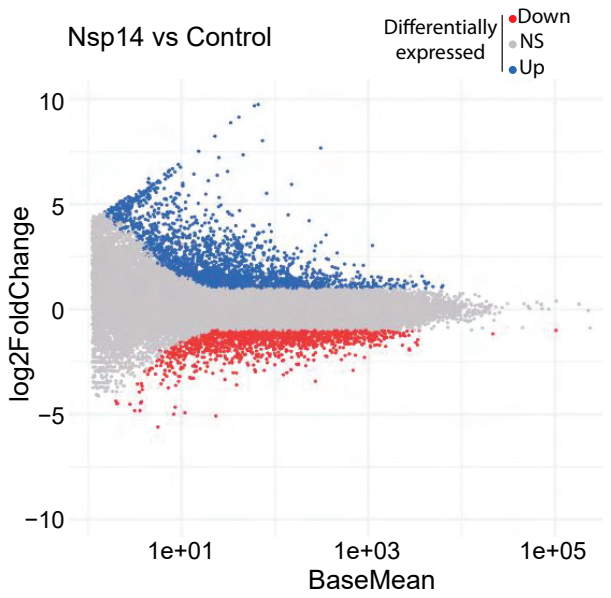

A.
B.
C.
−1
−2
−3
DNA metabolism
Transcription
RNA metabolism
Translation
Cell division
Cell metabolism
Immunity
protein

Figure 1

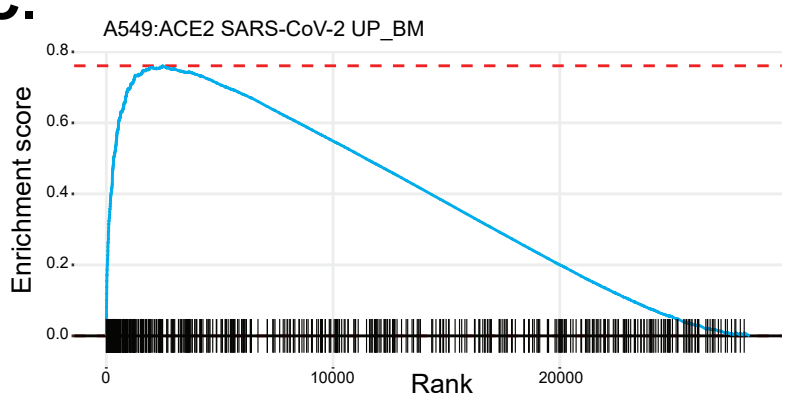
A.



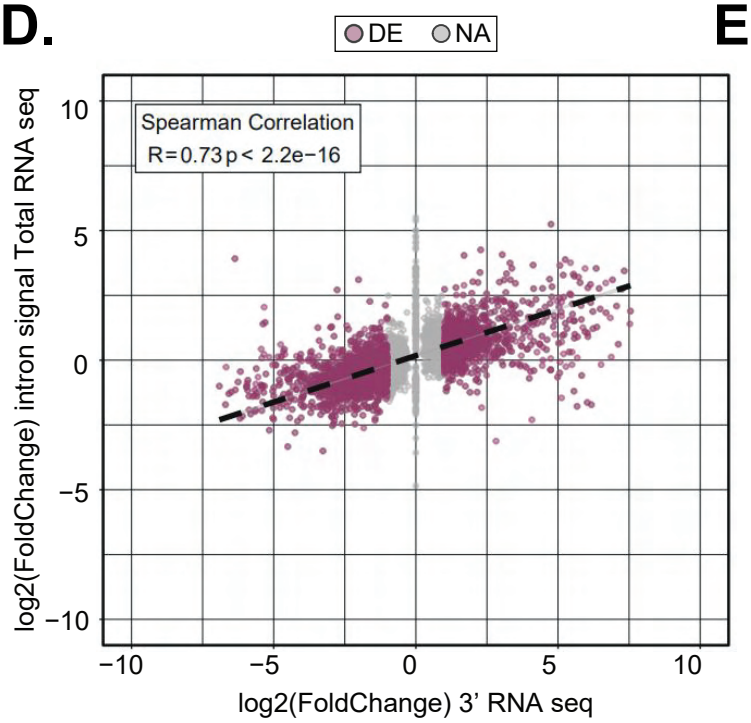
B.

GSEA Term	padj	NES
A549:ACE2 SARS-CoV-2 (RUX) UP_BM	0.0047	2.36
A549:ACE2 SARS-CoV-2 UP_BM	0.0048	2.32
HOXA5 targets 9h UP_CHEN	0.0044	2.31
A549:ACE2 SARS-CoV-2 low MOI UP_BM	0.0044	2.15
MCR5 cells MERS UP_BM	0.0045	1.92
A549:ACE2 SARS-CoV-2 DOWN_BM	0.0071	-1.75
A549:ACE2 Influenza-A UP_BM	0.2700	1.29
A549:ACE2 Influenza-A DOWN_BM	0.7100	-0.88

C.



D.



E.

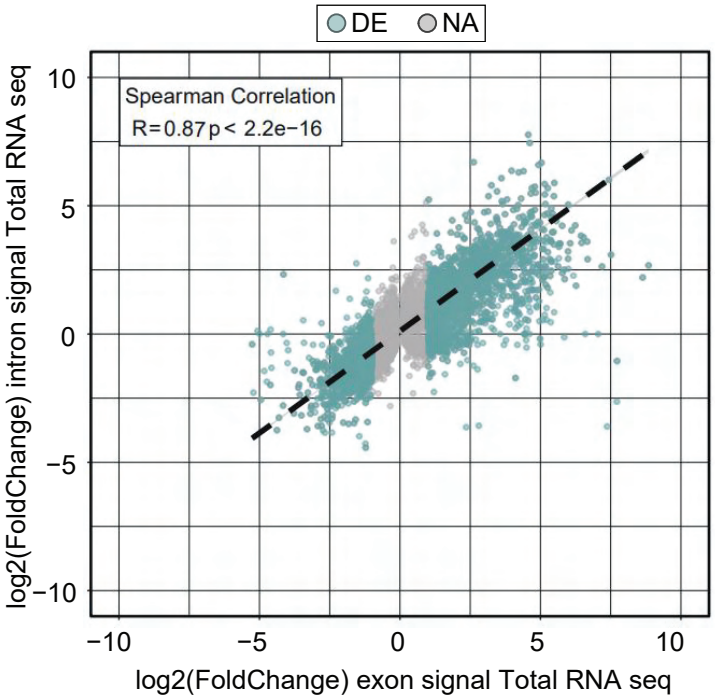
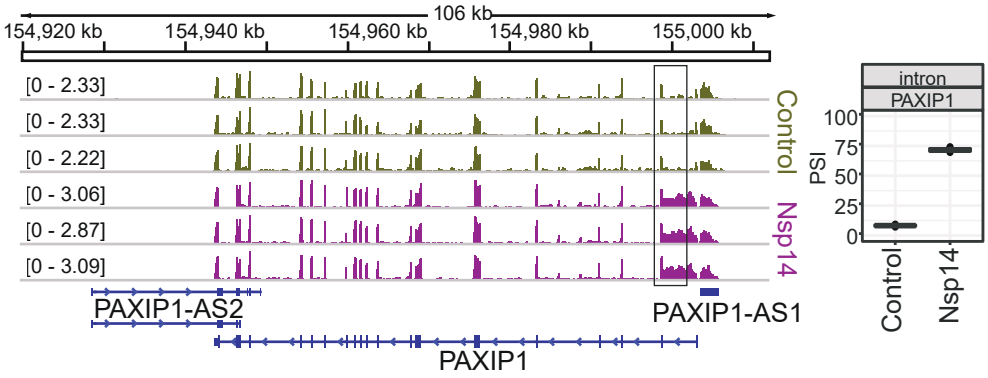
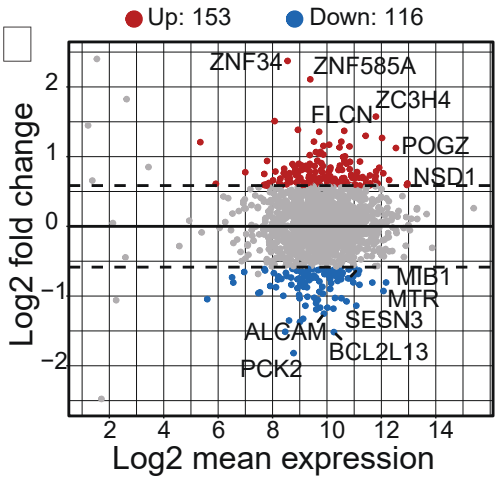
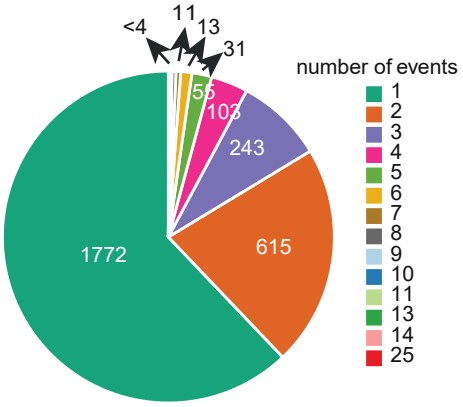


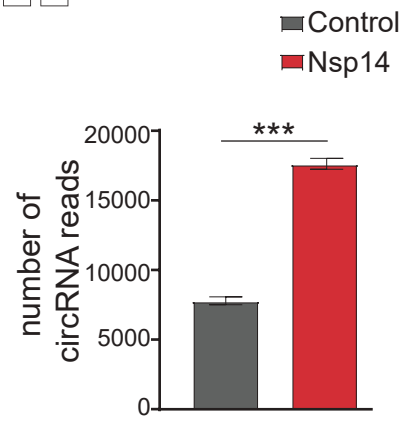
Figure 2

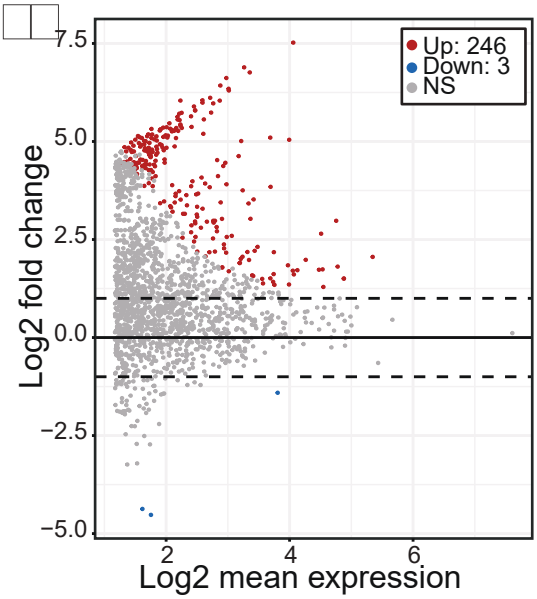
	□□	□□□ □
exon inclusion	238	2053
intron retention	2139	98
ALT Acceptor	65	70
ALT Donor	112	130

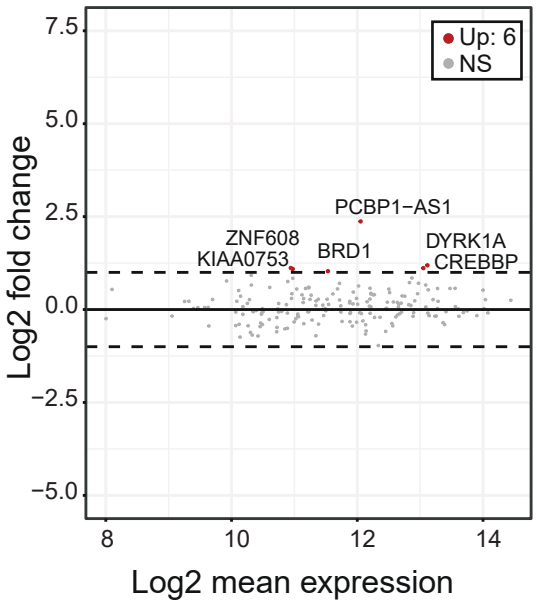


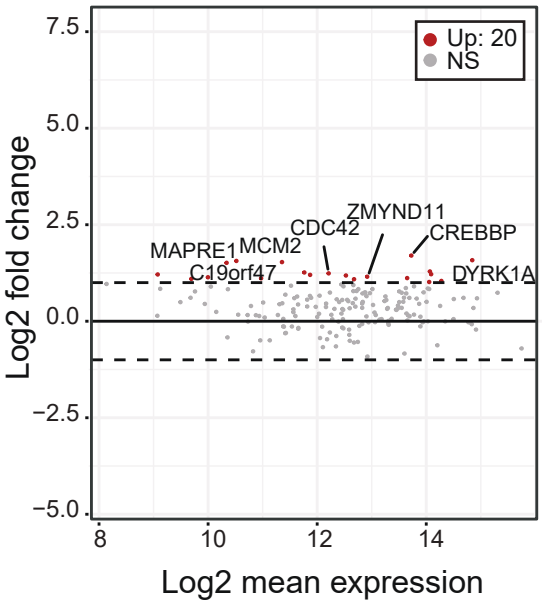












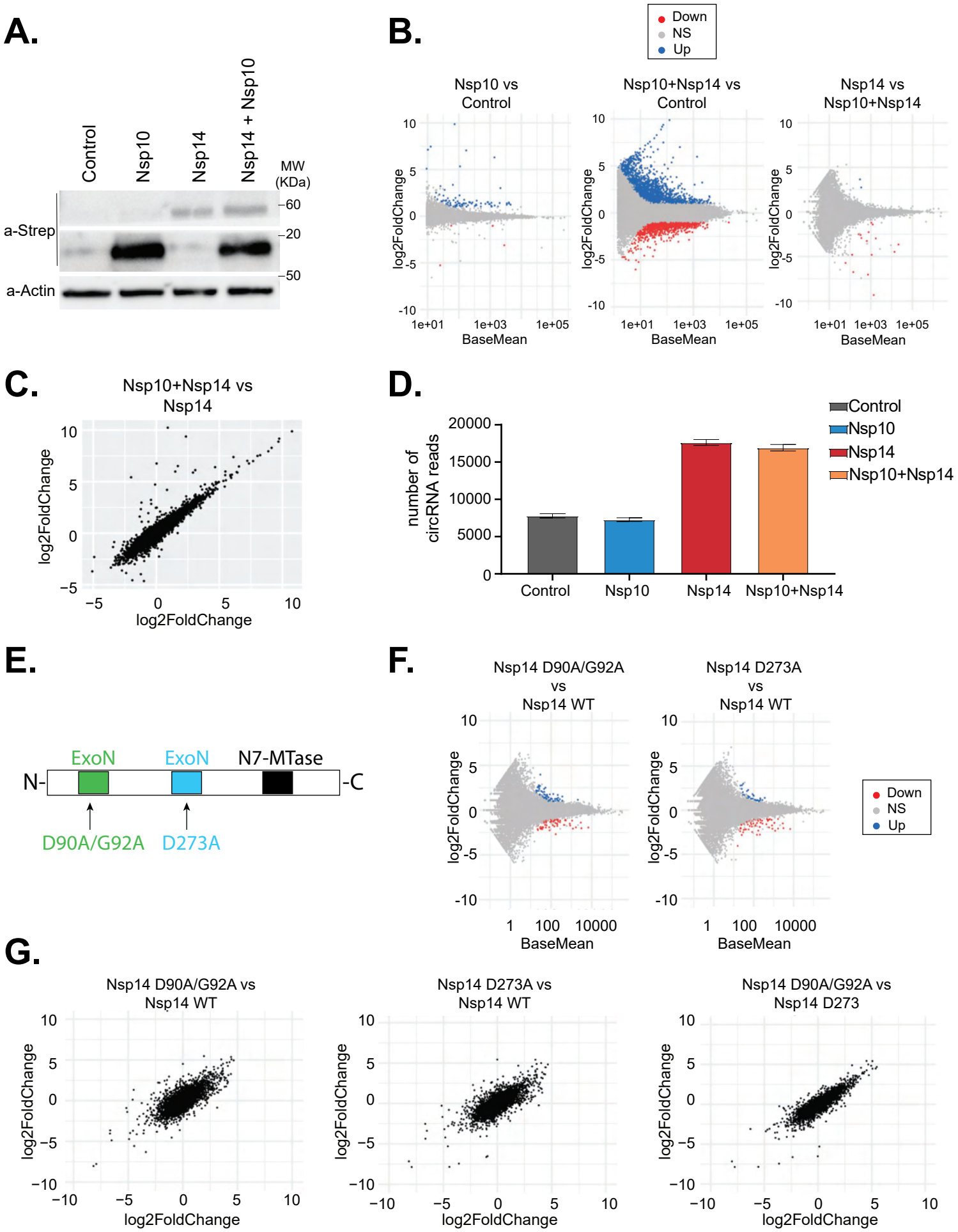
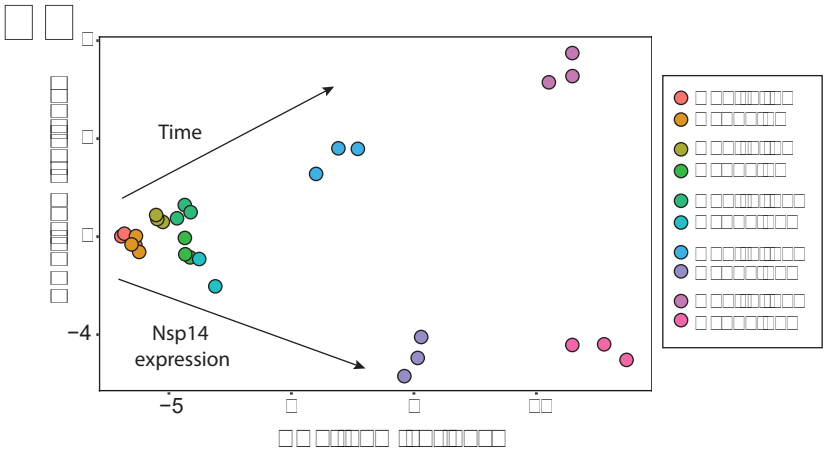
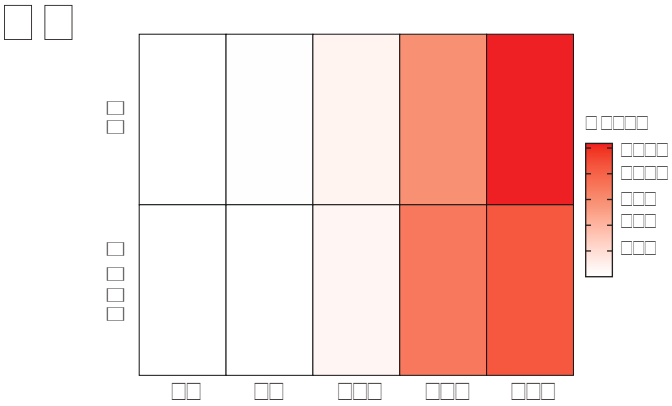
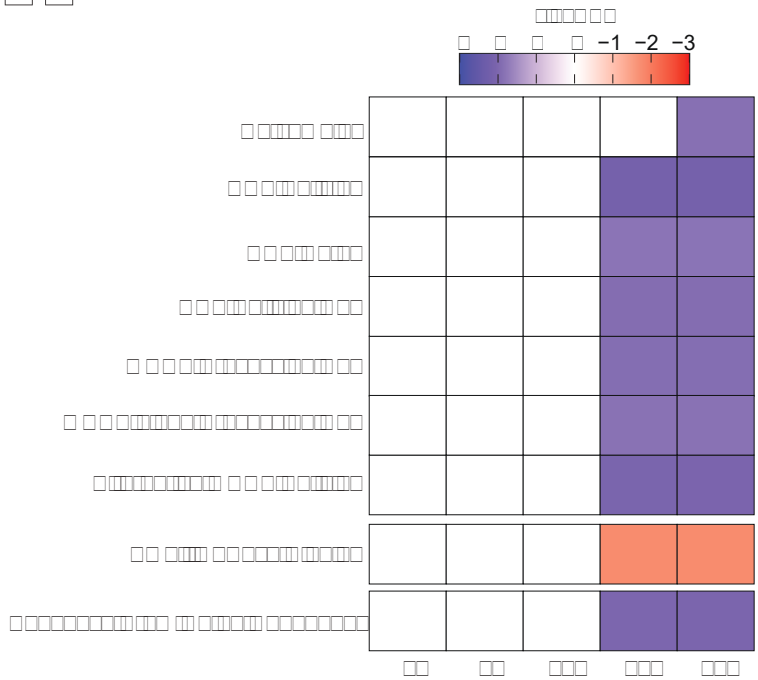
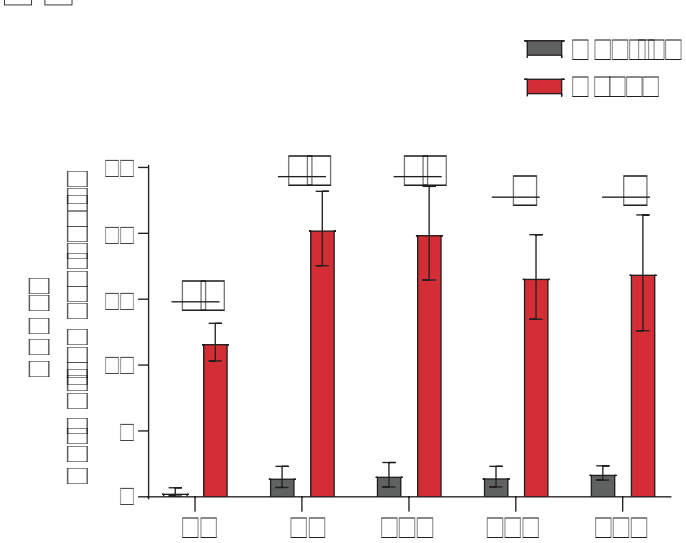

A.
Control
Nsp10
Nsp14
Nsp14 + Nsp10
MW (KDa)
a-Strep
a-Actin
60
20
50
B.
Down
NS
Up
Nsp10 vs Control
Nsp10+Nsp14 vs Control
Nsp14 vs Nsp10+Nsp14
log2FoldChange
BaseMean
C.
Nsp10+Nsp14 vs Nsp14
log2FoldChange
D.
number of circRNA reads
Control
Nsp10
Nsp14
Nsp10+Nsp14
E.
N-
-C
ExoN
ExoN
N7-MTase
D90A/G92A
D273A
F.
Nsp14 D90A/G92A vs Nsp14 WT
Nsp14 D273A vs Nsp14 WT
log2FoldChange
BaseMean
Down
NS
Up
G.
Nsp14 D90A/G92A vs Nsp14 WT
Nsp14 D273A vs Nsp14 WT
Nsp14 D90A/G92A vs Nsp14 D273
log2FoldChange

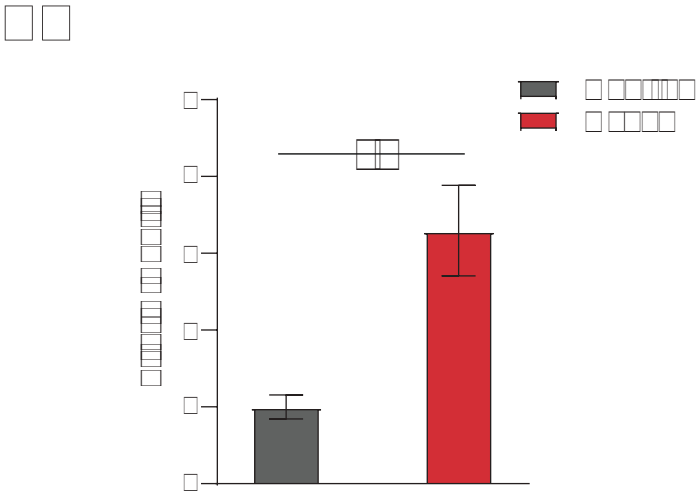
Figure 4

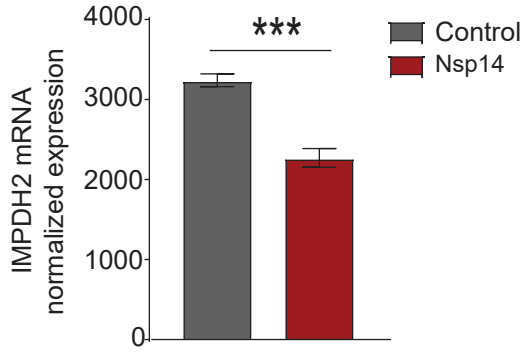

Time
Nsp14
expression
−4
−5

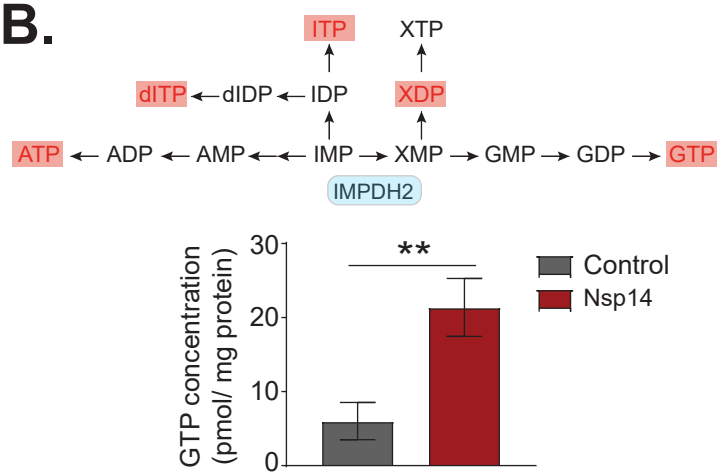


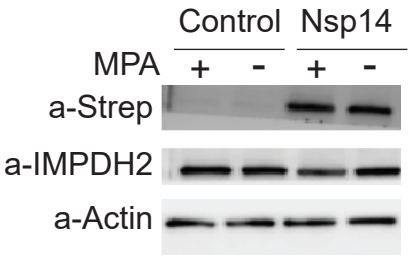

−1
−2
−3

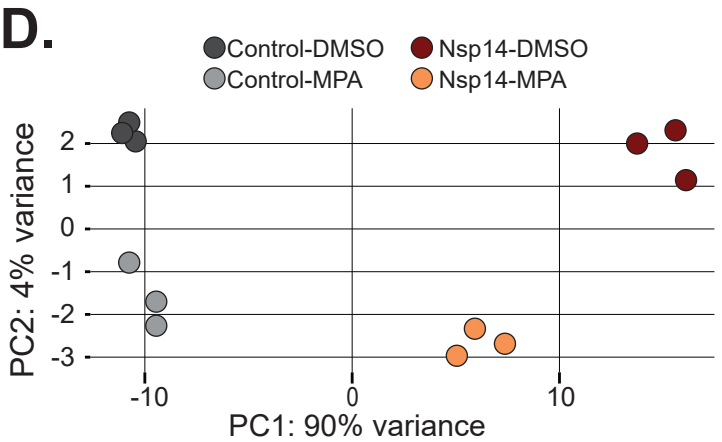












E.

Comparison	Number of Up DEGs	Number of Down DEGs
Nsp14-DMSO vs Control-DMSO	1942	1819
Nsp14-MPA vs Control-MPA	569	676
Control-MPA vs Control-DMSO	27	88

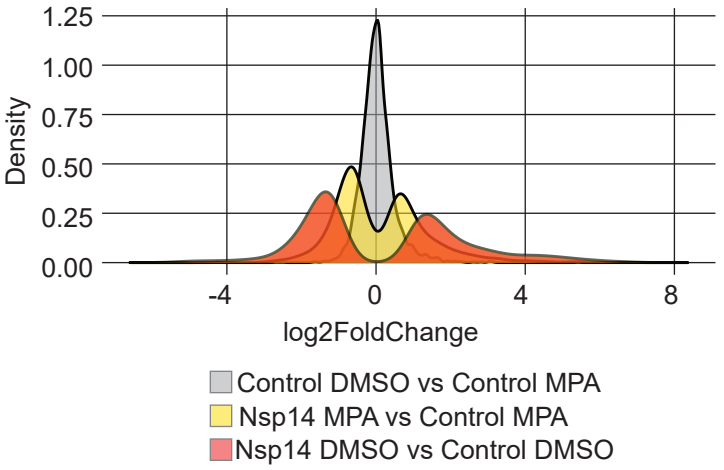
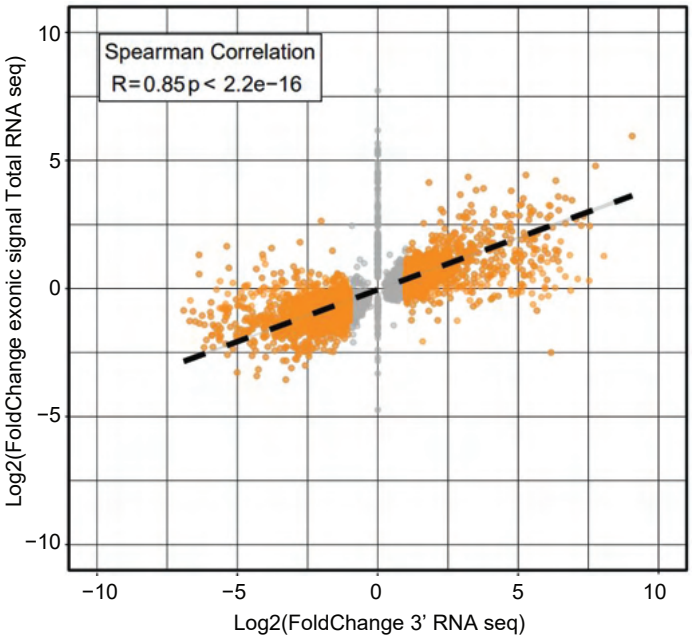


Figure 6

A.



B.

Dataset	Gene ranks	NES	pval	padj
BLANCO_MELO_COVID19_SARS_COV_2_ INFECTION_A594_ACE2_EXPRESSING_CELLS_ RUXOLITINIB_UP		2.36	2.7e−04	4.7e−03
BLANCO_MELO_COVID19_SARS_COV_2_ INFECTION_A594_ACE2_EXPRESSING_CELLS_UP		2.32	2.8e−04	4.8e−03
CHEN_HOXA5_TARGETS_9HR_UP		2.34	2.4e−04	4.4e−03
BLANCO_MELO_COVID19_SARS_COV_2_ LOW_MOI_INFECTION_ A594_ACE2_EXPRESSING_CELLS_UP		2.15	2.4e−04	4.4e−03
BLANCO_MELO_MERS_COV_ INFECTION_MCR5_CELLS_UP		1.92	2.5e−04	4.5e−03
BLANCO_MELO_COVID19_SARS_COV_2_ INFECTION_A594_ACE2_EXPRESSING_CELLS_DN		−1.75	5.3e−04	7.1e−03
BLANCO_MELO_INFLUENZA_A_ INFECTION_A594_CELLS_UP		1.29	1.3e−01	2.7e−01
BLANCO_MELO_INFLUENZA_A_ INFECTION_A594_CELLS_DN		−0.88	7.1e−01	7.7e−01

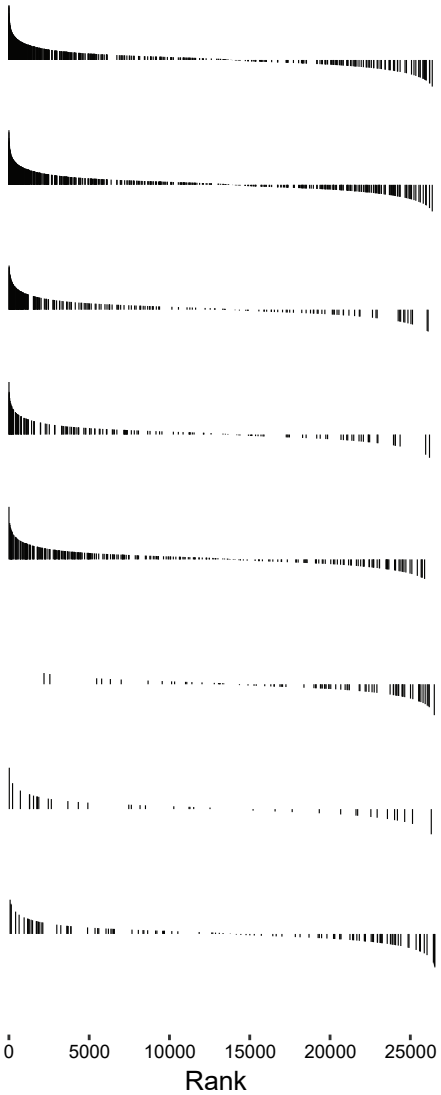
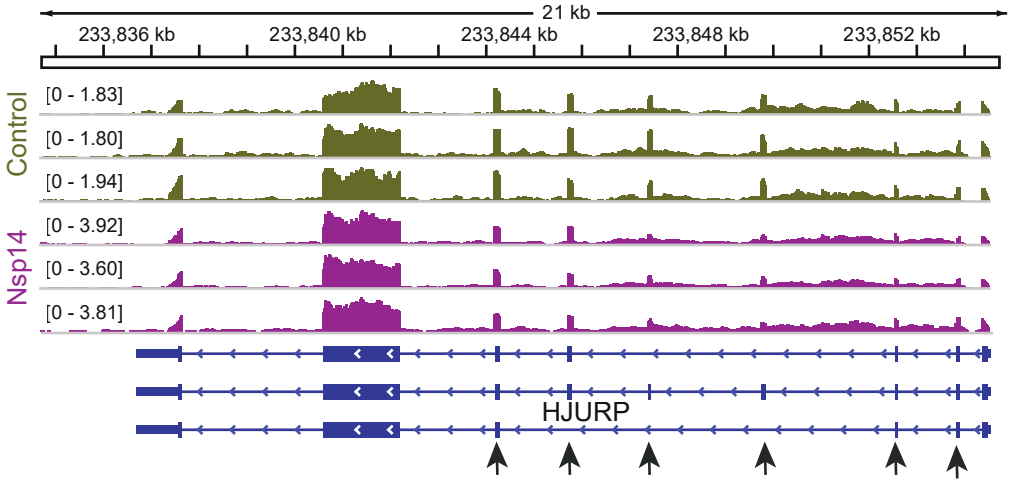
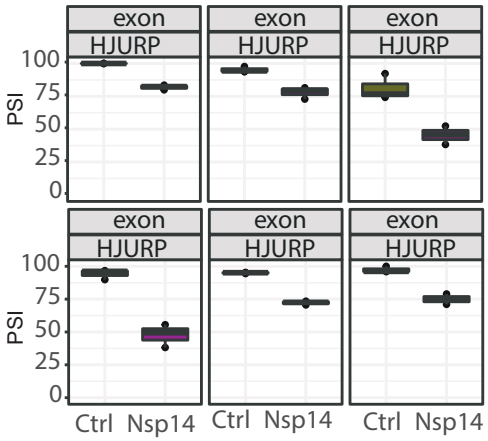
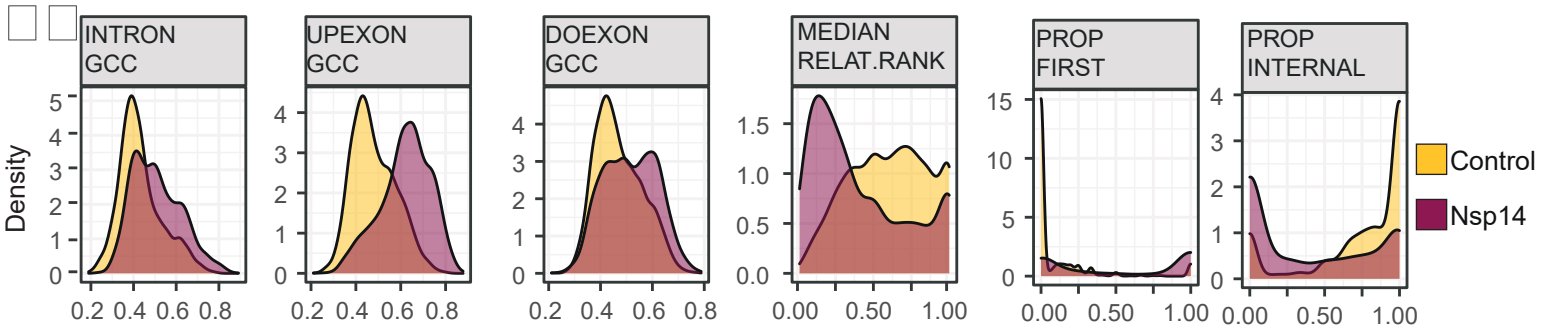
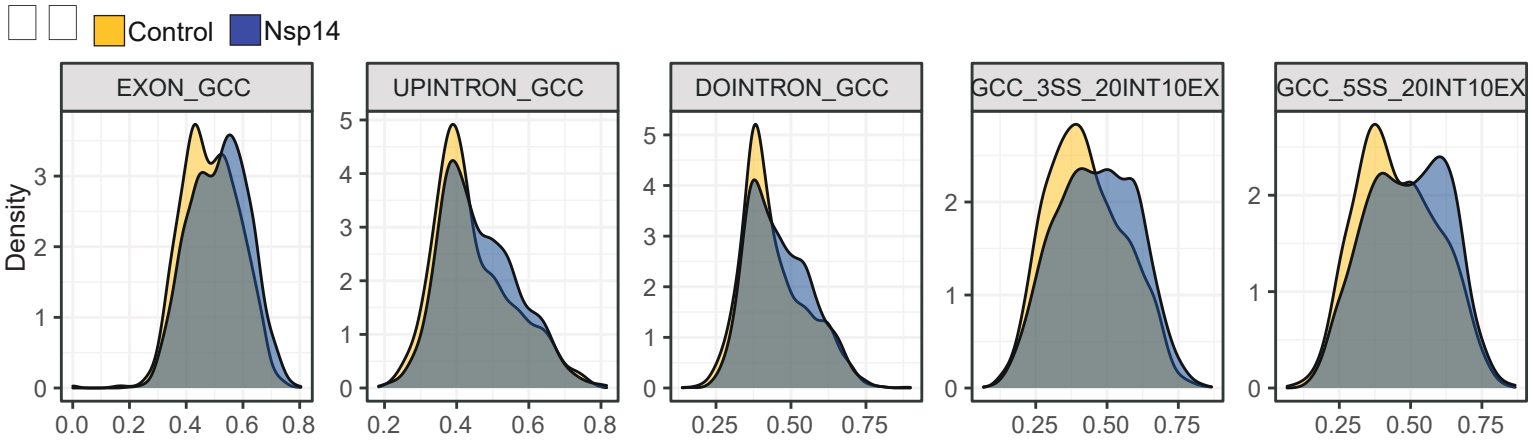


Figure S1


21 kb
233,836 kb
233,840 kb
233,844 kb
233,848 kb
233,852 kb
Control
[0 - 1.83]
[0 - 1.80]
[0 - 1.94]
Nsp14
[0 - 3.92]
[0 - 3.60]
[0 - 3.81]
HJURP


exon
HJURP
PSI
100
75
50
25
0
Ctrl
Nsp14


INTRON GCC
UPEXON GCC
DOEXON GCC
MEDIAN RELAT.RANK
PROP FIRST
PROP INTERNAL
Density
Control
Nsp14


Control
Nsp14
EXON_GCC
UPINTRON_GCC
DOINTRON_GCC
GCC_3SS_20INT10EX
GCC_5SS_20INT10EX
Density

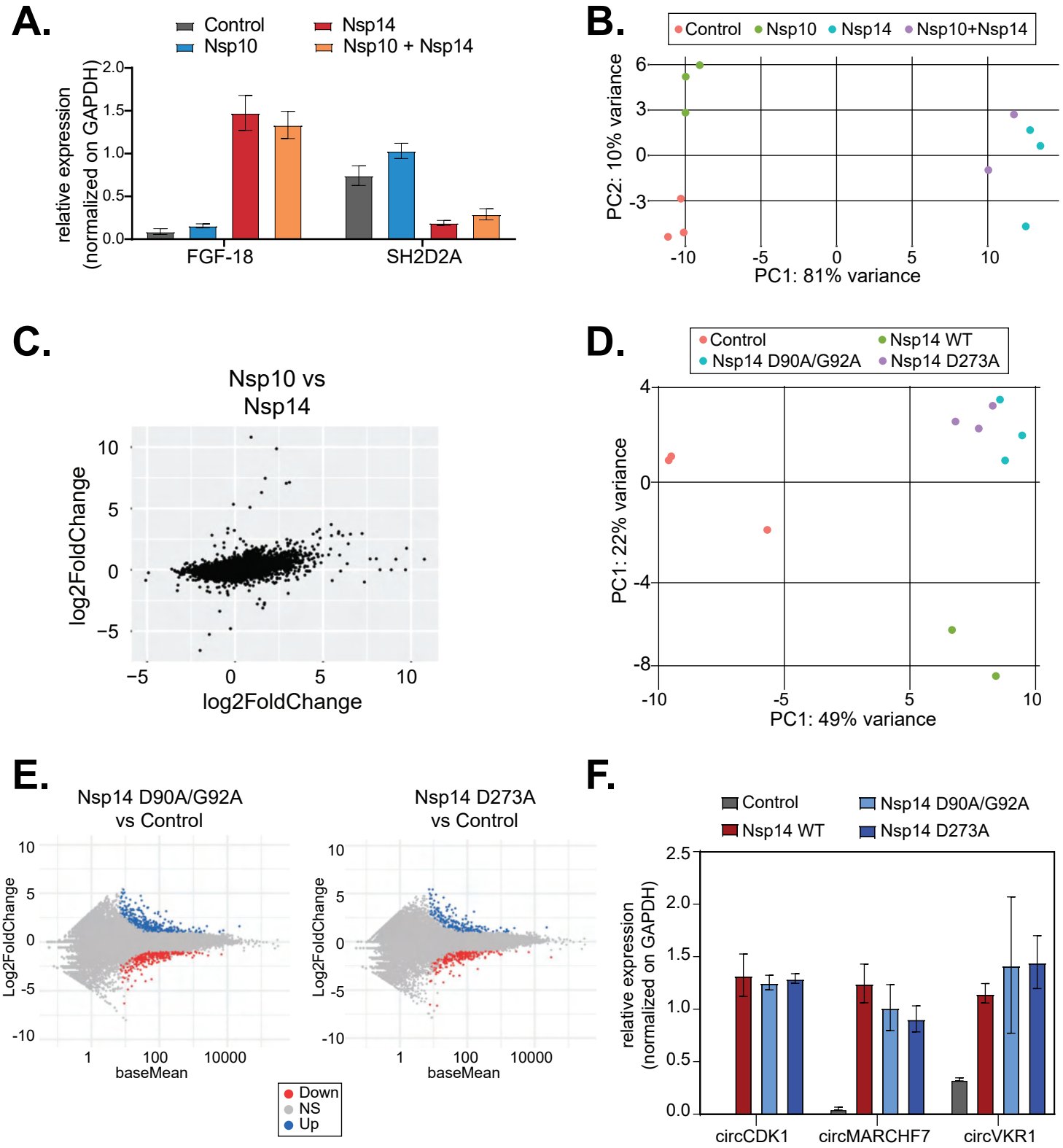

A.
Control
Nsp10
Nsp14
Nsp10 + Nsp14
relative expression
(normalized on GAPDH)
2.0
1.5
1.0
0.5
0.0
FGF-18
SH2D2A
B.
Control
Nsp10
Nsp14
Nsp10+Nsp14
PC2: 10% variance
6
3
0
-3
-10
-5
0
5
10
PC1: 81% variance
C.
Nsp10 vs
Nsp14
log2FoldChange
10
5
0
−5
−5
0
5
10
log2FoldChange
D.
Control
Nsp14 WT
Nsp14 D90A/G92A
Nsp14 D273A
PC1: 22% variance
4
0
-4
-8
-10
-5
5
10
PC1: 49% variance
E.
Nsp14 D90A/G92A
vs Control
Nsp14 D273A
vs Control
Log2FoldChange
10
5
0
-5
-10
1
100
10000
baseMean
Down
NS
Up
F.
Control
Nsp14 D90A/G92A
Nsp14 WT
Nsp14 D273A
relative expression
(normalized on GAPDH)
2.5
2.0
1.5
1.0
0.5
0.0
circCDK1
circMARCHF7
circVKR1

Figure S3

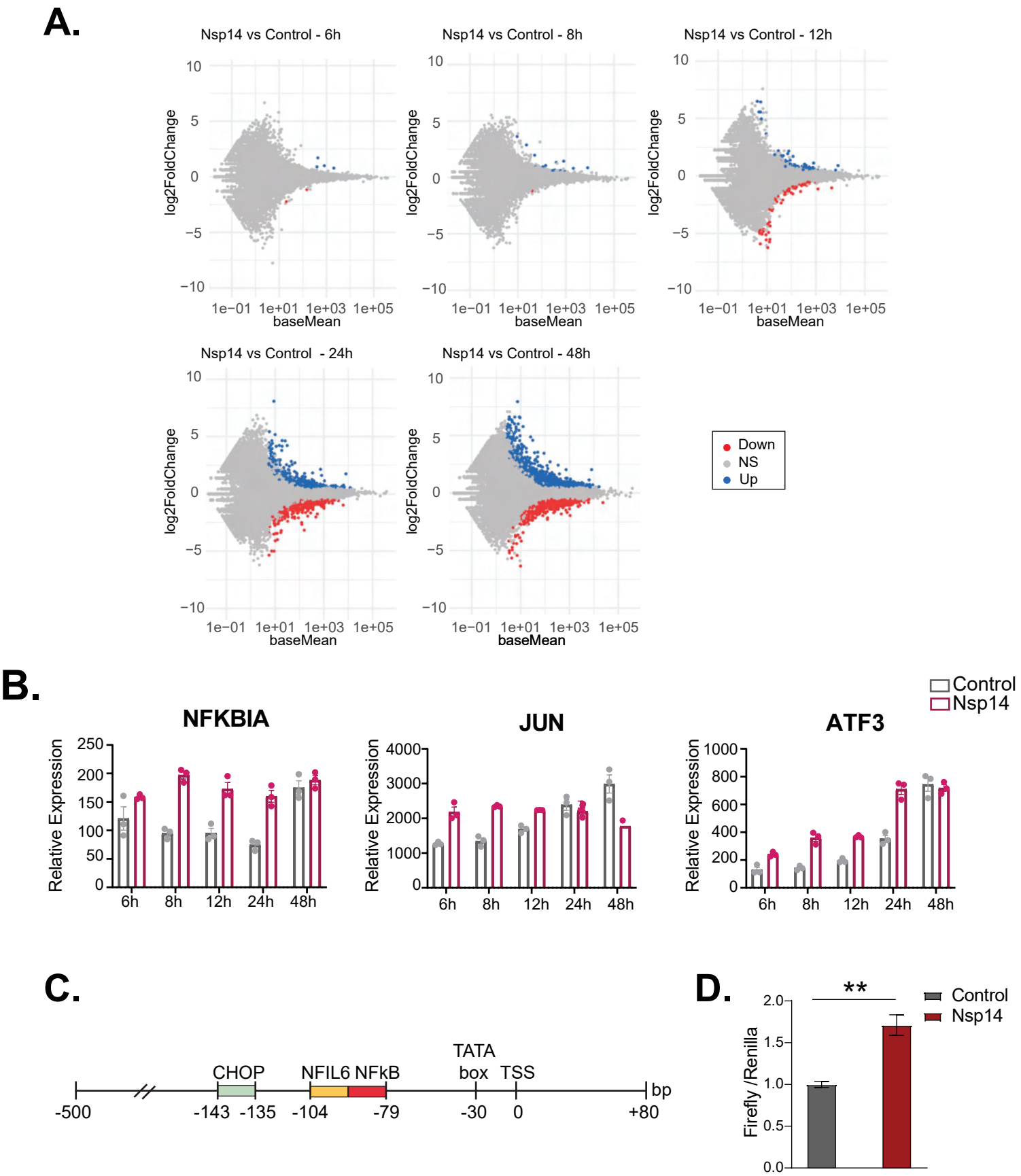
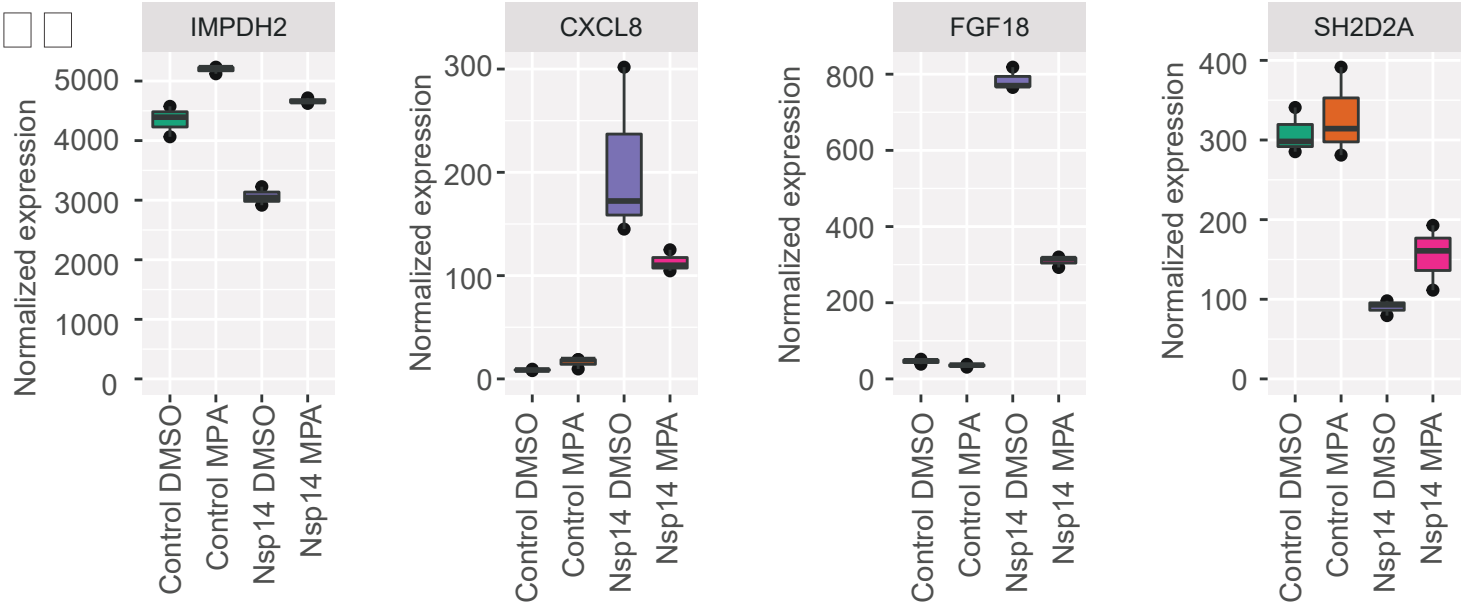
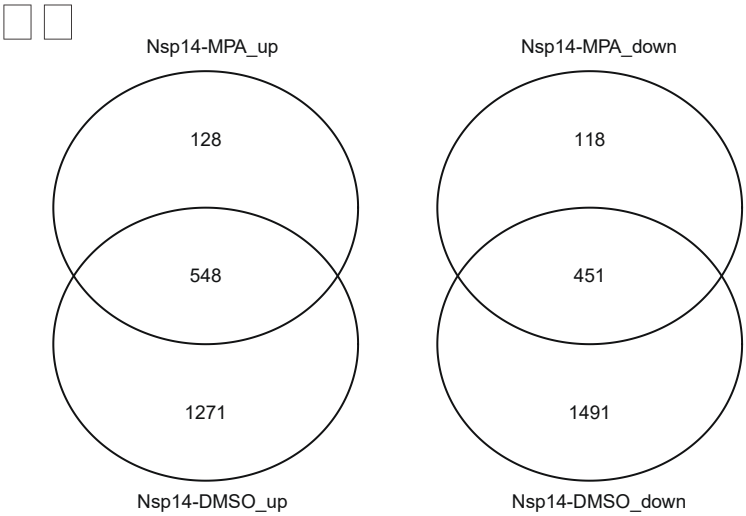
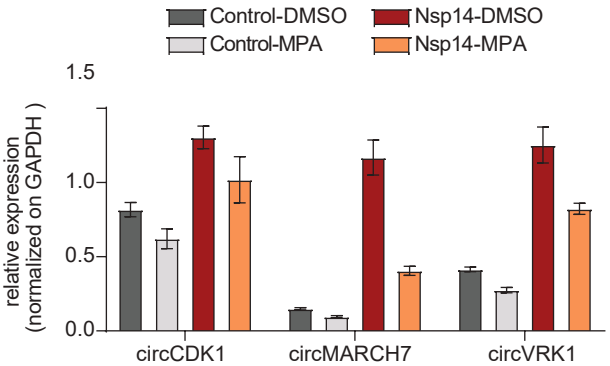

A.
Nsp14 vs Control - 6h
Nsp14 vs Control - 8h
Nsp14 vs Control - 12h
Nsp14 vs Control - 24h
Nsp14 vs Control - 48h
log2FoldChange
baseMean
Down
NS
Up
B.
Control
Nsp14
NFKBIA
JUN
ATF3
Relative Expression
6h
8h
12h
24h
48h
C.
CHOP
NFIL6
NFkB
TATA box
TSS
bp
-500
-143
-135
-104
-79
-30
0
+80
D.
**
Control
Nsp14
Firefly /Renilla

Figure S4


IMPDH2
CXCL8
FGF18
SH2D2A
Normalized expression
Control DMSO
Control MPA
Nsp14 DMSO
Nsp14 MPA


Nsp14-MPA_up
Nsp14-MPA_down
128
548
1271
118
451
1491
Nsp14-DMSO_up
Nsp14-DMSO_down


Control-DMSO
Control-MPA
Nsp14-DMSO
Nsp14-MPA
relative expression
(normalized on GAPDH)
circCDK1
circMARCH7
circVRK1