Washington University in St. Louis
Washington University Open Scholarship

Arts & Sciences Electronic Theses and Dissertations

Arts & Sciences

Winter 12-15-2021

Defining the epigenetic and transcriptional regulatory mechanisms by which hRSV NS1 alters host transcriptional response

Nina Rose Beri
Washington University in St. Louis

Follow this and additional works at: https://openscholarship.wustl.edu/art_sci_etds

 Part of the Medical Immunology Commons

Recommended Citation

Beri, Nina Rose, "Defining the epigenetic and transcriptional regulatory mechanisms by which hRSV NS1 alters host transcriptional response" (2021). *Arts & Sciences Electronic Theses and Dissertations*. 2567.
https://openscholarship.wustl.edu/art_sci_etds/2567

This Dissertation is brought to you for free and open access by the Arts & Sciences at Washington University Open Scholarship. It has been accepted for inclusion in Arts & Sciences Electronic Theses and Dissertations by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.

WASHINGTON UNIVERSITY IN ST. LOUIS

Division of Biology and Biomedical Sciences

Molecular Genetics and Genomics

Dissertation Examination Committee:

Jacqueline E. Payton, Chair
Thomas J. Brett
Cristina de Guzman Strong
Deborah J. Lenschow
Daisy W. Leung

Defining the Epigenetic and Transcriptional Regulatory Mechanisms by which hRSV NS1 Alters Host Transcriptional Response

by

Nina Beri

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

December 2021

St. Louis, Missouri

© 2021, Nina Beri

Contents

List of Figures

List of Tables

Acknowledgements

Consider the Platonic ideal of a gene: a self-contained unit of information. Now consider that unit of information as an infinitesimal fraction of an organism, a community, and the world. The complexity increases infinitely as abstraction yields to form. And the researcher has the Herculean, or perhaps Sisyphean, task of isolating the ideal and unchanging nature of the gene as it exists in an imperfect and chaotic universe.

In this dissertation, I explore the nature of the respiratory syncytial virus gene product, nonstructural protein 1, in a cellular milieu in constant flux. What would be an impossible task for one person has become, through the support of my family, friends, and mentors, a concrete step forward in understanding the nature of reality. Sarah and Jared and Jen, thank you for tolerating every quirk, attitude, and questionable comment I've expressed in the lab. Hannah and Chaz, you are wonderful scientists. Rachel, Ginger, and Brian, your friendship has kept me going through the good times and the bad. You guys are the best. Priya, your friendship has kept me going since long before grad school. Zach, I love you. Mom and Papa, I love you and you inspire me to grow. Gaya and Daisy, thank you for the constant encouragement. Jackie, thank you for, well, everything. The mentorship, the opportunity to work on a wonderful project, and the encouragement to explore my scientific passions in a rigorous and enthusiastic manner. All of you, and many others I don't have the space to thank, form the rich broth in which I have grown (like *E. coli* in log phase).

Nina Beri

Washington University

December 2021

Chapter 1: Introduction

1.1 Overview of respiratory syncytial virus

Human respiratory syncytial virus (hRSV) is a common respiratory pathogen that causes significant morbidity and mortality in infants as well as the elderly and the immunocompromised around the world. Annually 33 million pediatric cases of acute lower respiratory tract infection (LRTI) and over 60,000 deaths can be attributed to hRSV[1–5]. The virus is known to suppress the host immune response, a phenomenon that has been associated with the high rate of reinfection[6–8]. hRSV is also associated with a higher incidence of asthma[1,2,9] and other respiratory inflammatory diseases. Current treatments do not fully address the burden of disease[10,11]. No vaccine is currently available[12–15], although several are in Phase I-III clinical trials[16–18]. This situation calls for continuing research into the basic biology of hRSV to develop new preventive and therapeutic options for treatment.

hRSV is a single-stranded negative-sense RNA virus in the genus *Pneumoviridae* [19]. The ten genes in its genome code for 11 proteins[20]. hRSV performs mRNA transcription using the L, N, P, and M2-1/2 proteins[21]. The L (large polymerase) protein, an RNA-dependent RNA polymerase (RdRP), transcribes RNAs from the genome. In hRSV, there are gene start (GS) and gene end (GE) signals flanking each open reading frame (ORF), which are used in a stop-start mechanism wherein each gene must be fully transcribed (mediated by GE) before the next GS signal is read[22]. The M2-1 protein regulates processivity, wherein the L protein either obeys or bypasses GS and GE factors during genome replication. Replication produces a positive-sense antigenome as a replication intermediate. mRNA transcription begins from the 3’ end of the genome, and due to the stop-start mechanism, the genes at the 3’ end are produced most abundantly. As infection continues, the virus switches from RNA transcription to genome replication in preparation for

release of new virions[23]. M2-2, a processivity factor, accumulates throughout infection. As M2-2 allows the RdRP to continue transcribing the full genome rather than discrete ORFs, this leads to genome replication instead of RNA transcription[24].

RSV encodes two unique proteins that perform multiple immune antagonist functions, nonstructural protein 1 (NS1) and nonstructural protein 2 (NS2). Following these genes in the genome are Nucleoprotein, Phosphoprotein, Matrix, Small Hydrophobic Protein, Glycoprotein, Fusion, M2 (with two ORFs 1 and 2) and L, the large RdRp[23].The two transcript variants of M2 control the switch between transcription processivity and RNA replication. N and P are needed for encapsidation, those two plus L for RNA replication, and those three plus M2-1 for transcription processivity. M2-2 downregulates transcription and upregulates RNA replication or in other words, replication of the full negative-sense genome via a sense intermediate.

1.2 hRSV NS1 encodes a unique multifunctional interferon antagonist protein

NS1 and NS2, the first two genes transcribed starting at the 3' end of the genome, are produced abundantly early during infection. In cells infected with recombinant RSV minigenomes with NS1 deleted, NS2 is more abundant than in WT RSV-infected cells[25]; NS1 also inhibits genome replication and mRNA transcription of hRSV[26]. NS1 immunoprecipitates with Matrix protein[27]. The 142 amino acid long sequence of NS1 has an approximate molecular weight of 17 kDa.

Spann, Tran, and Collins showed localization of NS1 and NS2 in both the cytoplasm and nucleus of hRSV-infected cells via Western blot using a polyclonal serum that detects both NS proteins[28]. A similar antiserum has been used as early as 1998 to detect both NS proteins[26]. While designed against the C-terminal 10 amino acids of NS2, the two proteins share the terminal DLNP peptide. This is thought to be the basis for the recognition of both proteins by the antiserum used by Atreya et al[26]. In 2009, Swedan et al[29] performed immunofluorescence on A549s transiently transfected with individual FLAG-tagged NS1 or NS2 protein. They found cytoplasmic localization as well as overlap of both NS1 and NS2 localization with DAPI-stained nuclei.

Figure 1. Image from Chatterjee et al., *Nature Microbiology*, 2017

The crystal structure of hRSV NS1 was recently solved by our collaborators in the Leung group[30] (**Figure 1.1**). The NS1 crystal structure reveals significant structural homology to the NTD of hRSV Matrix[30] while the NS1 structure contains additional features not present in the Matrix structure. Specifically, it has a β5 sheet in the NTD, a reversal in direction of another sheet, and a C-terminal α3 helix that is not present in Matrix. This gene has no homologs in any other viruses outside of the *Orthopneumoviridae* clade, encompassing hRSV and its animal orthologs, which include bovine RSV (bRSV) and pneumonia virus of mice (PVM)[31]. The closest relatives of RSV, the *Metapneumoviridae*, do not encode orthologs to NS1.

Separately from and together with the other non-structural protein NS2, NS1 acts to antagonize immune response to RSV infection. Known activities of NS1 occur in the cytoplasm. Spann and colleagues found that NS1/NS2 deletion RSV has attenuated replication in type I IFN competent cells[32]. They also found that IFNα/β/λ were more highly expressed in A549s and

macrophages in response to infection with single or double NS deletion viruses than to wild type virus. They later showed that NS1 inhibits IRF3 nuclear translocation[28]. Using immunofluorescence to image apoptosis with TUNEL stain, Bitko and colleagues showed that each NS protein has an anti-apoptotic role[33]. They used siRNA against NS1 or NS2 in hRSV-infected cells and demonstrated an increase in apoptosis in hRSV infected cells[33]. Elliott and colleagues further demonstrated that NS1 and NS2 cooperate in the degradation of STAT2[34]. Furthermore, they showed that NS1 has homology to E3 ligase components and can bind CUL2 domains. They found NS1 degrades STAT2 by promoting K48 ubiquitination. Here they used tagged NS1/NS2 & ubiquitination complex components, and siRNA against NS1/2. Finally, they used bioinformatics to identify NS1 regions homologous to ubiquitin complex components. Taken together, these data suggest that NS1 antagonizes interferon responses in RSV infected cells.

1.3 Experimental evidence indicates a likely nuclear role for hRSV NS1

This section of the background will highlight major points from a previous publication by our collaborators, the Leung group, with relevance to my dissertation research. As previously mentioned, the NS1 protein consists of a core connected to an alpha helix (α3) by a flexible linker domain (**Figure 1.1**). The core contains two alpha helices and two beta sheets. The primary form of NS1 is thought to be the monomer. Alanine mutants of residues throughout the protein were generated based on the crystal structure. Within the α3 helix, mutants were generated at Y125 and L132/L133. A truncation mutant of NS1 was generated as well, NS1 1-118; this mutant excludes the alpha-three helix encompassing residues 119-142. hRSV carrying either wild-type or mutant NS1 was used to infect A549s, after which RNA was harvested at 24,48, and 96 hours post infection (h.p.i). Alanine mutated residues in the core did not significantly affect the transcriptional

response to RSV. However, mutations in the α3 helix, including NS1 Y125A, resulted in significantly different transcriptional profiles compared to WT NS1 infection (**Figure 3.6)**. While there are some transcriptomic differences between mock, 24 and 48 h.p.i for α3 helix mutant NS1 RSV infections, they are less profound than WT or core residue mutant NS1 infections[30]. Moreover, at 96 h.p.i., transcription levels in α3 helix mutant NS1 infections are essentially equivalent to mock infection (**Figure 3.6**). This suggests that the transcriptional response to RSV was altered by mutation of the α3 helix residues but not of the core residues. I interpret this observation as indicating that wild-type NS1 enables hRSV to persist in the cell, inducing an antiviral response longer than α3 helix mutant NS1.

Much of the transcriptional response to RSV infection involves interferon signaling. The *IFN-β* promoter is a primary target of the type I interferon induction pathway. The Leung group tested the impact of NS1 on IRF3-driven transcription using an *IFN-β* promoter reporter [30]. That NS1 mutants have a significant effect on gene transcription throughout infection suggested that NS1 might be active in the nucleus. However, NS1 is known to restrict IRF3 nuclear entry[28,35]. The Leung group bypassed this issue by using a mutant of IRF3 that was constitutively active and nuclear-localized: IRF3-5D[36,37]. Five residues were mutated to aspartic acid resulting in a constitutively active phosphomimetic. Activated IRF3, whether phosphorylated WT or IRF3-5D, dimerizes, translocates to the nucleus, and binds regulatory elements.

A plasmid constitutively expressing IRF3-5D with or without a plasmid constitutively expressing NS1 was transfected into Sendai virus infected A549 cells to further determine whether gene expression under control of the *IFN-β* promoter was altered in the presence of NS1[30]. The Leung group then carried out luciferase reporter assay under the control of the *IFN-β* promoter. The *IFN-β* promoter, the target of the luciferase reporter assay, is one of the best characterized

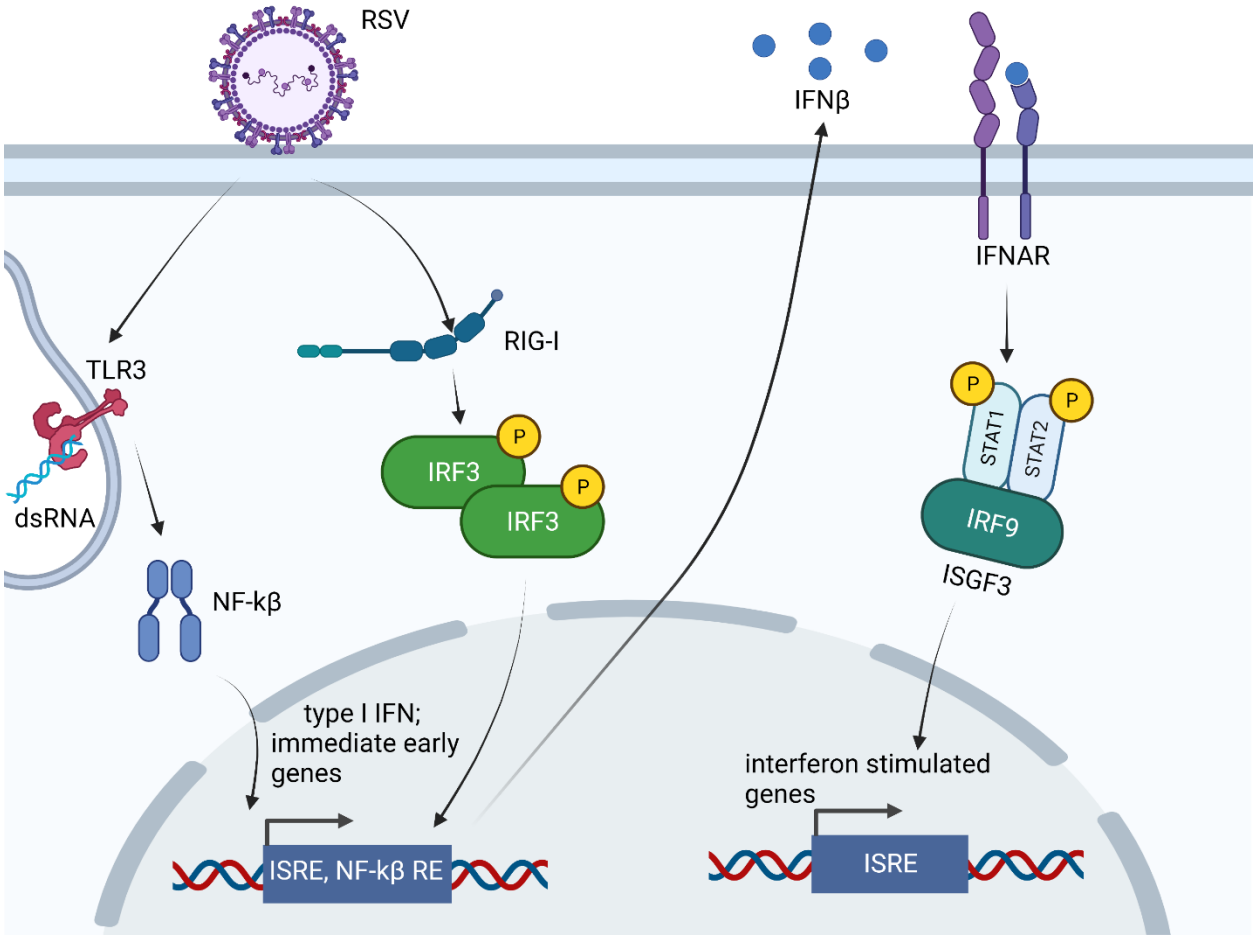


Figure 1.2. Schematic of the type I interferon induction and response pathways. Innate immune response depends on sensing pathogen associated molecular patterns, or non-self molecules associated with infection. Several families of pattern recognition receptors (PRRs) identify different features of PAMPs. For the purposes of this talk, I am focusing on two PRRs that detect RNA virus patterns: RIG-I, a cytoplasmic PRR; and TLR-3, an endosomal PRR, both of which recognize double stranded RNA. Once activated upon detection of double stranded RNA, RIG-I initiates signaling that leads to the phosphorylation, dimerization, and nuclear translocation of IRF3. Once localized in the nucleus, IRF3 binds sequence specific motifs at the regulatory elements of innate immune response genes. Downstream of TLR3 activation, NF-kB dimers are released from inhibition by inhibitor of kappa B kinases A and B (IKKα/β), allowing it to translocate to the nucleus and bind its own sequence specific regulatory elements.

Among the genes targeted by the IRF3 and NF-kB transcription factors is interferon-β, a type I interferon that can act in an autocrine or paracrine manner to induce the cell in which it is detected to adopt an antiviral state. Type I IFN receptors on the cell membrane bind IFN-β, leading to the formation of the ISGF3 complex from IRF9 and phosphorylated STAT1 and STAT2. ISGF3 then translocates to the nucleus, where it binds interferon stimulated regulatory element motifs – as does IRF3 – at a wide selection of interferon stimulated genes, which when expressed put the cell into the antiviral state.

promoters for which expression may be induced by IRF3. Reporter expression was strongly

upregulated when IRF3-5D was added but partially reduced when NS1 is added. This suggests that a cytoplasmic role of NS1 is not sufficient to explain its ability to reduce gene expression and supports a nuclear role for NS1 to modulate reporter gene expression downstream of the IRF3 activation recapitulated by IRF3-5D. Taken together, these findings suggest a role for NS1 to play in the transcriptional regulation of innate immune response and interferon signaling genes in the nucleus.

1.4 NS1 immunoprecipitates with several components of the Mediator transcriptional regulatory complex

Affinity precipitation followed by mass spectrometry (AP-MS) experiments with NS1 as bait pulled down many subunits of the Mediator complex[38,39]. As discussed earlier, RSV replicates its genome and transcribes mRNAs in the cytoplasm. Mediator is a chromatin-associated nuclear-localized transcriptional regulatory complex[40], found at active promoters and enhancers, bridging the two into closer proximity and enhancing the activity of RNA Pol II[40]. Mediator also has other activities[41,42]. It consists of three evolutionarily conserved modules: head, middle, and tail[43] and for which the majority of constituent polypeptides are conserved between yeast and human encompassing all three modules. The Head module interacts with RNA Pol II and the pre-initiation complex (PIC) at the core promoter elements of transcriptionally active genes. Several subunits of Mediator interact with the PIC component TFIIH[44]. The MED14 subunit of Mediator connects the head, middle, and tail[44]. The most evolutionarily diverged polypeptides of Mediator are those in the Tail module. This region of Mediator interacts with tissue specific transcription factors found at the enhancer regions of actively transcribed genes[41–46].

1.5 NS2 is the other unique interferon antagonist of hRSV

Aside from NS1, hRSV encodes a second unique interferon antagonist[23]. Like NS1, NS2 has a unique primary sequence excepting the C-terminal DLNP peptide. With the discovery of the crystal structure of NS2[47], the Leung group demonstrated that NS2 binds an inactive form of the RIG-I and MDA-5 signaling molecules, with the N-terminal residues of NS2 directly involved in the binding interaction. This agrees with previous work that showed that NS2 interacts with the N-terminal caspase activation and recruitment domain (CARD) of RIG-I, disrupting its association with MAVS[48]. This interaction is a key step in the RIG-I signaling pathway, and the disruption prevents accumulation of IRF3 in the nucleus.

Deletion of NS2 alone or with NS1 attenuates infectivity of hRSV in interferon competent cells[28,32,47], indicating that NS2 plays an interferon-dependent role in infectivity. Beyond the interaction with RIG-I and MDA-5 described above, several groups have described additional interferon antagonist roles for NS2. Lo and colleagues demonstrated that NS2 leads to a decrease in STAT2 and responsiveness to the type I IFN signaling pathway[49]. Spann et al. showed that NS2 partitions predominantly to the cytoplasm and also to the nucleus of infected cells[28]. Elliott and colleagues showed that NS1 and NS2 contribute to the proteasomal degradation of STAT2[34]. Thus, NS2 is a second multifunctional interferon antagonist protein encoded by hRSV.

1.6 hRSV Matrix is a structural protein with structural homology to NS1

hRSV NS1 shares significant structural homology with the hRSV Matrix protein[30]. Matrix is one of four proteins, along with F, N, and P, required for hRSV virus-like particle formation and budding[50]. As a structural protein, Matrix forms a shell around the nucleocapsid and interacts with

the viral envelope[23]. It has a short hydrophobic region, but not a full-fledged transmembrane domain, via which it transiently interacts with the host cell membrane[51].

Early during the infectious life cycle of hRSV, Matrix localizes to the nucleus using its bipartite nuclear localization (bpNLS) sequence in its interaction with IMPβ1-family importin[22,50,52,53]. Inhibition of nuclear export via IMPβ1-family inhibition using verdinexor restricts the budding of hRSV[54]. Importin family nuclear export receptor CRM1[55], an IMPβ1- related exportin, is responsible for nuclear egress of Matrix[56]. This suggests that hRSV Matrix must enter and then exit the nucleus for the virus to bud.

Furthermore, hRSV Matrix is implicated in transcriptional repression while it is nuclear localized[52]. However, I raise an issue with this conclusion. The experiments in the relevant paper observed host transcriptional repression in the presence of nuclear extract from hRSV-infected cells but did not confirm using isolated Matrix. Their conclusion assumed that Matrix was the only hRSV protein found to be in the nucleus and therefore was the causative agent[52]. However, our results demonstrating nuclear localization of NS1[39] suggest a potential alternative explanation for this phenomenon.

hRSV Matrix has sequence and structural homology to many other NNSV Matrix proteins and shares many functions with them[57,58]. *Paramyxovirus* Matrix proteins share a nuclear export signal (NES) and a bipartite nuclear localization signal (bpNLS) with a required ubiquitinatable lysine[57,59,60]. Matrix of Newcastle disease virus (NDV) can also localize to the nucleus via a different mechanism[60,61]; henipavirus (Nipah, Hendra) Matrix proteins do as well[59,62]; measles virus Matrix has been shown to inhibit host cell transcription[63].

1.7 Host epithelial cells activate innate immune signaling pathways upon recognition of infection

hRSV infects respiratory epithelium throughout the respiratory tract[1,23,64]. In healthy adults, RSV infects tissue of the nasopharynx, or upper respiratory tract leading to a mild infection with cold-like symptoms. However, in high-risk groups the immune response is lacking, which allows hRSV to infect tissues throughout the respiratory tract up to and including the lung tissue.

The respiratory tract is lined with epithelial cells[65]. This tissue constantly interacts with the environment and functions as an interface between the lungs and inhaled environmental air. It is therefore one of the first lines of defense against airborne pathogens[65,66]. Respiratory tract epithelial cells are part of the innate immune system. Surface molecules and other receptors throughout the cell[67,68] known as pattern recognition receptors (PRRs) allow type II lung epithelial cells (ciliated alveolar epithelial cells) and other innate immune cells to recognize infectious material. The PRRs consist of three families of receptors (RIG-I like receptors [RLRs]; Toll-like receptors [TLRs]; and Nod-like receptors [NLRs]) that recognize pathogen-associated molecular patterns (PAMPs), usually molecules in forms that do not occur in mammalian cells.

PRR mediated signaling pathways generally lead to the induction of interferon signaling, which allows the cell and those around it to adopt an "antiviral state" that is less sensitive to infection[69,70]. Interferon signaling, which potentiates this state, leads to transcriptional induction of a suite of interferon stimulated genes[64,71–73]. This state precipitates the secretion of inflammatory mediators that recruit the other types of cells within the immune system to respond to infection and induces paracrine signaling that allows nearby cells to adopt an antiviral state[64,71,74]. The RIG-like receptors (RLRs), RIG-I, MDA-5 are the major recognizers of hRSV infection[47,48,75,76]. Activation of RIG-I or MDA-5 leads to a signaling cascade that results in interferon (IFN) upregulation and the production of IFN-stimulated genes (ISGs, also termed IFN-regulated genes,

IRGs). Within the toll-like receptors, TLR3 is endosomal and is known to sense hRSV RNA as well[22,68,77].

The immune response can be divided into the innate and adaptive phases[8,67,78–80]. During the innate immune response, host cells rapidly detect and respond to molecules associated with infection[64]. In contrast, the adaptive immune response depends on antigen recognition and happens slowly after the first encounter with a pathogen; dedicated cell types including B cells, T cells, and dendritic cells carry out adaptive immune functions. The innate immune response is important because it happens quickly, within hours to days of infection. The rapid replication of microbial pathogens requires a fast-acting response.

PAMPs can be grouped into several types of pathogenic molecules common to a class of infectious pathogens and that have physical attributes not seen in mammalian cells. Examples include peptidoglycan, lipopolysaccharide, and non 5' capped single stranded RNA[67,78]. Common PAMPs include double-stranded RNA lacking canonical post-translational modifications such as a 5' cap. DAMPs, or damage associated molecular patterns, are also recognized by PRRs. These are host molecules usually detectable by PRRs when localized inappropriately[81]. RIG-I recognizes short, unmodified dsRNA, while MDA-5 recognizes long dsRNAs[67,78]. Eukaryotic mRNA is usually modified with a 5' cap and 3' poly-adenylation, but the mRNA of viral pathogens is not[76,82].

To persist long enough to generate virions that will go on to enter new nearby cells, an infectious pathogen must combat the innate immune response of the infected host cell[64]. Successful pathogens have evolved a range of biochemical processes that allow them to evade these responses. For example in the case of the uncapped mRNA, some viruses have the ability to "snatch" the 5'ppp cap from mammalian RNAs[83] so they are not recognized as foreign by RIG-I. These mRNAs

can then be translated into functional proteins, undisturbed by the innate immune machinery within the infected cell.

Viruses also commonly encode proteins with immune antagonist functions. Many non-structural (NS) proteins, so named because they do not form part of the infectious virion, serve this function across a range of viruses. Influenza virus NS1[84], unrelated to hRSV NS1, is an antagonist, as is Ebolavirus VP35[85]. From the ssnsRNA viruses, these proteins tend to function in the cytoplasm or at the mitochondrial membrane blocking upstream steps of type I IFN induction via binding PRRs or their adaptors and blocking signal transduction.

The genome of RSV is a negative sense RNA genome that can immediately produce sense mRNA transcripts[21–23]. However, the RNA transcripts of hRSV are the ligand for RIG-I, activating the innate immune response to viral infection. RIG-I ligand binding activates a signaling pathway that leads to the transcription of IFN[86] and a subset of other genes activated directly downstream of RIG-I and TLR3 signaling (**Figure 1.2**). Upon activation of RIG-I, nuclear translocation of IFN regulatory factor family transcription factors then drives transcription. The IRF3 transcription factor is constitutively expressed but inactive pending post-translational modification. RIG-I signaling[76,82] leads to phosphorylation, dimerization and nuclear translocation of IRF3[36,37,86–91]. IRF3 dimers and IRF3/7 dimers are the major transcription factor complexes responsible for expression of genes including type I IFN. TLR3 similarly removes inhibition of NF-kB which permits it to translocate to the nucleus and bind its cognate regulatory elements[92–95].

Once type I IFN (e.g. IFN-β) has been produced, it can act in an autocrine or paracrine manner[71,74,96,97]. Type I IFN engages the type I IFN receptor, IFNAR[64] (**Figure 1.2**). Downstream of this receptor signaling is the JAK/STAT pathway[64]. JAK/STAT signaling culminates in the phosphorylation of STAT1/2 and their complexing with each other and IRF9, forming the

interferon stimulated growth factor 3 (ISGF3)[71,74,98,98]. ISGF3 then translocates to the nucleus where it stimulates the transcription of several hundred interferon stimulated genes (ISGs).

Three major classes of IFN are recognized, types I, II, and III[99,100]. The major class of IFNs expressed throughout the body are type I IFN[71,74,98,101], which include the several variants of IFN-α and the singular IFN-β. IFN-λ is the only type II IFN discovered. IFN-γ is a type III IFN and its expression is restricted to specific cell types[32,102] The activation of type I IFN in the polarized type II lung epithelial cells[65] infected by RSV leads to a signaling cascade that catalyzes the expression of a wide array of ISGs. IFN expression leads to activation and nuclear translocation of factors including IRF-3 and NF-kB, which activate ISG expression[36,37,86,89,91,103,104]. The NS1 protein of hRSV sequesters IRF-3 in the cytoplasm[28,32], which prohibits IRF-3 from acting as a transcription factor to stimulate ISG expression. However, even in the presence of a constitutively active, nuclear-localized IRF-3 (IRF3-5D, discussed in detail below), IFN-β promoter driven reporter expression is still lower in the presence of NS1[30], indicating that the cytoplasmic activities of NS1 do not fully explain its immune antagonist activities.

Chapter 2: Methods

Methods section adapted from Pei, Beri, et al. [39]

Cells

Cell lines: Adenocarcinomic human alveolar basal epithelial (A549) and human embryonic kidney epithelial (293T) cell lines were obtained from ATCC (CCL-185 and CRL-3216; Manassas, VA, USA) and were maintained in complete Dulbecco's modified Eagle's medium (DMEM;

ThemoFisher, 11965) supplemented with 10% heat-inactivated fetal bovine serum (FBS; Sigma-Aldrich, F4135). Cells were maintained at 37 °C with 5% CO_2.

METHOD DETAILS

Plasmids. RSV NS1 was synthesized and subcloned into a pCAGGS vector containing either an N-terminal Flag- or GFP- tag. Keap1[105] was subcloned similarly. pCAGGS containing 3HA-FLAG tagged NS1 (WT or Y125A) was used for chromatin immunoprecipitation and luciferase reporter assays.

Viral infection. RSV strain A2 used in the study was obtained from ATCC (VR-1540; Manassas, VA, USA). To determine the virus titers, cells cultivated in 24-well plates were inoculated with 10-fold serial dilutions of virus and incubated in 10% DMEM with methylcellulose at 37 °C for 7 d. Cells were fixed with cold methanol at −80 °C for 1 h. Methanol was removed and cells were incubated with 5% milk blocking buffer at 37 °C for 1h. Followed by incubation of goat anti-RSV (Fisher, AB1128MI) and HRP-labeled donkey anti-goat (Fisher, AP180PMI) antibodies. Cells were incubated with 0.03% 4-Chloronapthol and 1% hydrogen peroxide at 25 °C. After 20 min incubation, the plate was dried upside down and plaques counted. The multiplicity of infection (MOI) was confirmed according to the virus titer from the plaque assay. RNA-seq data from infections with RSV Y125A mutant NS1 was generated for Chatterjee et al.[30] and re-analyzed here (GSE99298).

For infection, A549 cells were grown to approximately 80% confluence in cell culture plates and were infected with RSV at a MOI of 1. Mock infection was performed with phosphate buffered

saline (PBS). After 1 hpi, the inoculum was removed by aspiration. Cells were washed twice with PBS and incubated in complete medium at 37°C for different time points until harvesting. For experiments with KPT-335 (verdinexor), A549 cells were infected with RSV at a MOI of 1 or mock-infected with PBS. The inoculum was removed 1 hpi and cells were washed twice with PBS, followed by replacement of complete medium supplemented with KPT-335 (1μM, RayBiotech 331-21369-1) or dimethyl sulfoxide (DMSO) at 37 ℃ for 24 h incubation.

Transfection

For ChIP: A549 cells were maintained in F-12K medium supplemented with 10% fetal bovine serum and 5 mM penicillin–streptomycin at 37°C with 5% carbon dioxide. 293T (human embryonic kidney cell line) were maintained in DMEM medium supplemented with 10% fetal bovine serum and 5 mM penicillin–streptomycin at 37°C with 5% carbon dioxide. A549 cells (10e^6) in 15cm plates were transfected with 20ug of pCAGGS vector containing a 3x HA-tagged NS1 (HA-NS1) or 3x HA-Y125A NS1 (HA-Y125A) or no insert (empty vector) using Lipofectamine transfection reagent (Invitrogen, Carlsbad, CA, USA).

Chromatin immunoprecipitation- (ChIP) seq and ChIP-qPCR. A549 cells were cross-linked 24 hours post-transfection at 70-90% confluence or 48 hours post-infection with hRSV. Cross-linking: Cells in 15cm plates were washed two times with PBS, incubated for 15 minutes with 1.25mM EGS (ethylene glycol-bis(succinic acid N-hydroxysuccinimide ester), Sigma, St. Louis, MO, USA) in PBS, washed with 3 times PBS, incubated for 10 minutes with 1% formaldehyde (Sigma, St. Louis, MO, USA) in PBS. Crosslinking was quenched with the addition of glycine (125 mM final concentration, Sigma, St. Louis, MO, USA) for 5 minutes. Then cells were washed

1 time with PBS. All cross-linking steps were performed at room temperature. Cells were scraped from the plate in ice-cold PBS, aliquoted to 10e^6 cells per tube, snap-frozen on dry ice and stored at -80C until sonication. Prior to sonication, cells were lysed in an SDS buffer (1% SDS, 10 mM EDTA, 50 mM Tris [pH 8.0], dH2O) with EDTA-free protease inhibitors (Roche, France) for 20 minutes on ice, then chromatin was fragmented using a Bioruptor (Diagenode, Denville, NJ, USA) to an average size of less than 300 bp. Ten percent of each sonicated sample was set aside for the input sample and the remaining sonicated lysates were subjected to immunoprecipitation overnight at 4°C with antibody conjugated to Protein-A Dynabeads. The following antibodies were used: normal rabbit IgG (12-370, Millipore), anti-HA (ab9110, Abcam, Cambridge, UK), and Mediator (TRAP1/CRSP220, A300-739A, Bethyl Laboratories, Montgomery, TX, USA); 2.5ug per $10e^6$ cells. Chromatin-IP-bead mixtures were then washed three times with buffers: low-salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris [pH 8.0], 150 mM NaCl, dH2O), high-salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris [pH 8.0], 500 mM NaCl, dH2O), LiCl buffer (0.25 M LiCl, 1% NP-40, 1% deoxycholate, 1 mM EDTA, 10 mM Tris [pH 8.0], dH2O), and 1X Tris-EDTA buffer, and eluted with an SDS buffer (1% SDS, 0.1 M NaHCO3, dH2O). Chromatin cross-links were reversed in eluted ChIP and input samples by overnight incubation with 5M NaCl at 65°C and DNA was isolated using PCR MinElute spin purification columns (Qiagen) as per manufacturer's instructions. At least 10 ng of ChIP or input DNA was submitted for indexed library preparation to the Genome Technology Access Center at Washington University in St. Louis. Samples were indexed and pooled (9-11 per lane) and subjected to 50 bp single-end sequencing according to the manufacturer's protocol (Illumina HiSeq3000, San Diego, CA, USA). ChIP-qPCR was performed with equal volumes of eluted DNA isolated as for ChIPseq with the following primers: IFITB: GGTATGCCGACCTTGAGAGAG and

TTCCCACTAAGGGTCCTGTTC; IFITC: TGATGCGTGCCCTACTCTC and CTGTGTCTCTGCTGTTCCGA; IFITD: GGCTGTTTCCTTATTGTTGCTCT and AGCAGTCCTGGTTCTGTGAG; GAPDH: CGCAGAGCCTCGAGGAGAAG and ACAGGAGGACTTTGGGAACGAC.

RNA-seq

Infected cells were harvested and RNA was extracted using the Qiagen RNA Easy kit (Cat. No. / ID: 74004) and 500ng submitted to the Genome Technology Access Center at Washington University in St. Louis. Samples were indexed and pooled (3 per lane) and subjected to 100 bp paired-end sequencing according to the manufacturer's protocol (Illumina HiSeq3000, San Diego, CA, USA).

Luciferase reporter assays. Luciferase reporter assays were performed as previously described[106] with the following modifications. 2.5x10^4 293T cells per well of a 96-well plate were transfected with pNL1.1.TK (control NanoLuc vector, Promega, Madison, WI, USA), plus wild-type or Y125A mutant NS1 or empty vector (pCAGGS), with or without 4ug/ml poly(I:C) (for IFIT locus wide luciferase assay) or 14.3 ng/ml (for all other luciferase assays); IFN-β 1000 IU/ml; or TNF-α 20 ng/ml 5 hours before harvest (Sigma, St. Louis, MO, USA), plus luciferase reporter plasmids: minimal promoter pGL4.23 (Promega) or ISRE reporter pGL4.45 (Promega) or minimal promoter pGL4.23 (Promega) with one of the NS1 binding regions cloned in, and with the control reporter NanoLuc pNL1.1 (Promega). 293T cells were subjected to luciferase assay 24 hours post-transfection (triplicate wells per condition, 96-well plate). The Promega Nano-Glo system (#N1110) was used to measure the firefly and NanoLuc (control) luciferase activity for each well

according to the manufacturer's instructions in a BioTek Cytation5 plate reader. All experiments were performed at least twice.

ChIP- and RNA-sequencing analysis. Reads were aligned to hg19 with bowtie2 (v2.2.5) with default settings[107]. Peaks were called with HOMER (version v4.10.1) with the settings tbp = 1 and findPeaks cmd = findPeaks ns1_ha_align_tbp1/ -style factor -o auto -i ns1_input_align_tbp1/[108]. RPKM normalized genome browser tracks were created with deepTools (v3.1.0) bamCoverage utility with settings--binSize10--extendReads150—normalize using RPKM and visualized on the UCSC genome browser 29106570. ChIPQC (v1.14.0) was used for quality control[109]. The ChIPSeeker R package (v1.16.1) and HOMER were used to annotate peaks[108,110,111].

For new RNA-seq data acquired for this paper (0 h.p.i., 24 h.p.i., 48 h.p.i.), basecalls and demultiplexing were performed with Illumina's RTA version 1.9 and bcl2fastq2 software with a maximum of one mismatch in the indexing read. RNA-seq reads were then aligned to the Ensembl release 76 primary assembly with STAR version 2.5.1a[112]. Gene counts were derived from the number of uniquely aligned unambiguous reads by Subread:featureCount version 1.4.6-p5 (Liao et al., 2014). Isoform expression of known Ensembl transcripts were estimated with Salmon version 0.8.2[113]. Sequencing performance was assessed for the total number of aligned reads, total number of uniquely aligned reads, and features detected. The ribosomal fraction, known junction saturation, and read distribution over known gene models were quantified with RSeQC version 2.6.2[114]. All gene counts were then imported into the R/Bioconductor package EdgeR[115], and TMM normalization size factors were calculated to adjust for samples for differences in library size. Ribosomal genes and genes not expressed in the smallest group size minus one sample greater than one count-per-million were excluded from further analysis. The TMM size factors and the matrix

of counts were then imported into the R/Bioconductor package Limma[116]. Weighted likelihoods based on the observed mean-variance relationship of every gene and sample were then calculated for all samples with the voomWithQualityWeights[117,118]. The performance of all genes was assessed with plots of the residual standard deviation of every gene to their average log-count with a robustly fitted trend line of the residuals. Differential expression analysis was then performed to analyze for differences between conditions and the results were filtered for only those genes with Benjamini-Hochberg false-discovery rate adjusted *p*-values less than or equal to 0.05. Heatmaps and volcano plots were generated with R packages heatmap2 and ggplot2. Genotify (v1.2.1) was used for manual gene curation[55]. Graphing and statistical analyses were performed using the R ggplot2 (3.2.1) and stats (3.4.1) packages and with Graphpad Prism version 8.2.1.

ChIP-qPCR analysis. Triplicate PCR wells were averaged and input wells were also scaled to 100% (10% of total input chromatin had been set aside from each sample prior to ChIP): input Ct value -3.322. Percent input for each ChIP sample was calculated as % input =100 * 2^([scaled input]-[sample]). Log2 fold change was plotted in figure by 2^(% input-ChIP sample/% input IgG).

Luciferase reporter assay analysis. The average ratios of the firefly to NanoLuc luciferase replicate values for each condition were compared to the average ratio for the minimal promoter reporter to determine relative luciferase activity. Statistical analyses were performed using Graphpad Prism 8. A p value < 0.05 was considered statistically significant. For all graphs, mean values ± standard deviation (SD) are shown.

Statistical analysis. Unless otherwise indicated, statistical analyses were performed using Graphpad Prism 8. A p value < 0.05 was considered statistically significant. For all graphs, mean values ± standard deviation (SD) are shown.

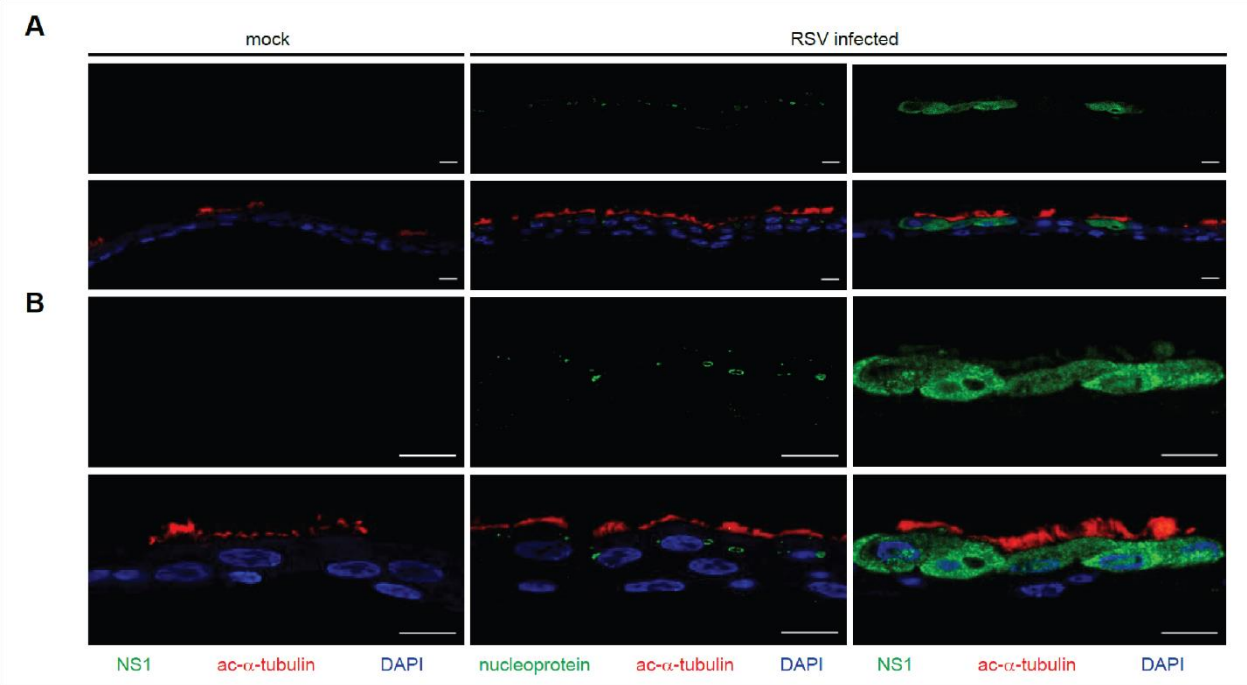
Table 2.1 Primers used to clone NS1 bound sites into pGL4.23

NB353	CGCGgagctcAGCCCAAGGTTGTAAACCACT	IFIT_D_For_1	chr10-176
NB354	TATAagatctCCATCTCAGGCTCAGGTCAG	IFIT_D_Rev_1	chr10-176
NB355	CGCGgagctcATTGCAGGTCTCAAGCCGTTA	IFIT_C_For_1	chr10-226
NB356	TATAagatctGGAAATAGCTGCACACAGGG	IFIT_C_Rev_1	chr10-226
NB357	CGCGgagctcCCTGGGAAGGAACACCACAC	IFIT_B_For_1	chr10-385
NB358	TATAagatctTTGAAGGCAGTTTTAGGGGCA	IFIT_B_Rev_1	chr10-385
NB359	CGCGgagctcACCCCTCATACAATCCTGCC	IFIT_A_For_1	chr10-214
NB360	TATAagatctGGTCTTCCAGGTCTGAAGCAA	IFIT_A_Rev_1	chr10-214
NB505	CGCGgctagcCTAGCCACTCCCACCACAAG	ISG20_ISRE_LUC_F1	chr15-12

NB506	TATActcgagCCCCATCCCCTGCCTTACC	ISG20_ISRE_LUC_R1	chr15-12
NB509	CGCGgctagcATTGCAGGTCTCAAGCCGTTA	IFIT3_ISRE_LUC_F1	chr10-226
NB510	TATActcgagGCCTGCACAGTAAGAAACTCAAC	IFIT3_ISRE_LUC_R1	chr10-226
NB511	CGCGgctagcGTAGCAGGCTCCAGAAGTTAGTTGTG	NFKB_incl_F	chr15-12
NB512	TATActcgagCAAGTGAAGTCAGGGGCGGA	NFKB_incl_R	chr15-12
NB513	CGCGgctagcAGTCCTGGGGATGTTTATTCTCTG	ISRE_incl_F	chr15-12
NB514	TATActcgagCATCGGCATCCCGACCCTG	ISRE_incl_R	chr15-12

Chapter 3: NS1 is distributed throughout the genome at several hundred discrete genomic regulatory elements and physically associates with the transcriptional regulatory complex Mediator

3.1 Work by several investigators in our groups indicates that NS1 partitions to the nucleus and associates with components of the Mediator complex[39]



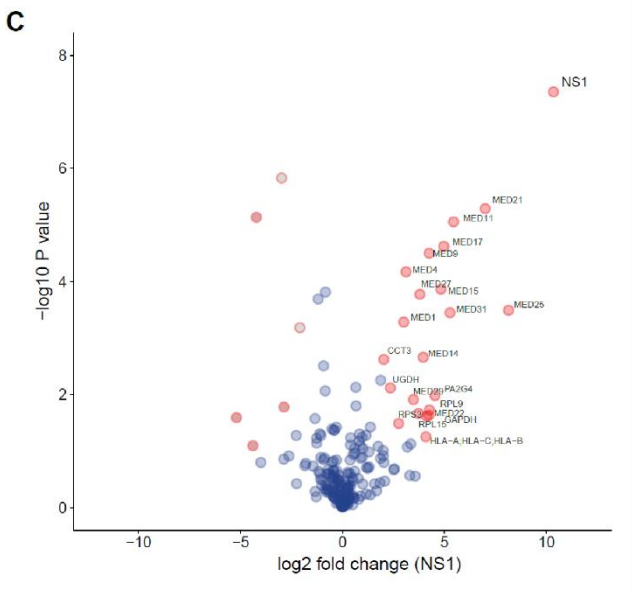


Figure 3.1. NS1 is found in the nucleus of RSV infected human airway epithelial cells and associates with components of the nuclear-localized Mediator complex. A and B, Low and high power magnified images of mock or RSV infected hTECs differentiated using air-liquid interface culture technique were stained for DAPI (blue, nuclear stain), ciliary marker anti-acetylated tubulin (red, ac-α-tubulin), and either hRSV nucleoprotein or NS1 as labeled (green). Scale bar is 10 μm. C, Affinity purification followed by mass spectrometry identified interacting proteins of NS1. Volcano plot shows fold change over unrelated control (n=3) and *p*-value. Significant interactors in red, other in blue, and depleted in grey. *Adapted from Pei, Beri, et al. Cell Reports 2021*

NS1 has several demonstrated immune antagonist functions in the nucleus as described in the Introduction. Immunofluorescence of primary human tracheobronchial epithelial cells (hTECs) differentiated into ciliary cells using air-liquid interface (ALI)[119,120] was performed to detect RSV proteins nucleoprotein or NS1 after infection. Differentiated cells were fixed and probed with anti-acetylated tubulin as a molecular marker of cilia, an organelle specific to ciliated epithelial cells, and DAPI as a nuclear stain. Antibody to hRSV nucleoprotein detected nucleoprotein localized to the cytoplasm, in a pattern representative of localization in inclusion bodies, wherein RSV is thought to replicate and avoid immune detection[121–123]. An antibody to NS1, a unique reagent recently generated by the Leung group[39], detected NS1 throughout the cell, in the nucleus as well as cytoplasm, but excluded from nucleoli (**Figure 3.1A, B**). Work by the Leung group (**Figure 3.1C**) and others[38,39] using affinity purification followed by mass spectrometry (AP/MS) showed that several components of the nuclear-localized Mediator complex, a central transcriptional regulatory complex, associate with NS1. Given this unexpected result, the Payton and Leung groups continued to investigate a potentially novel nuclear role for NS1.

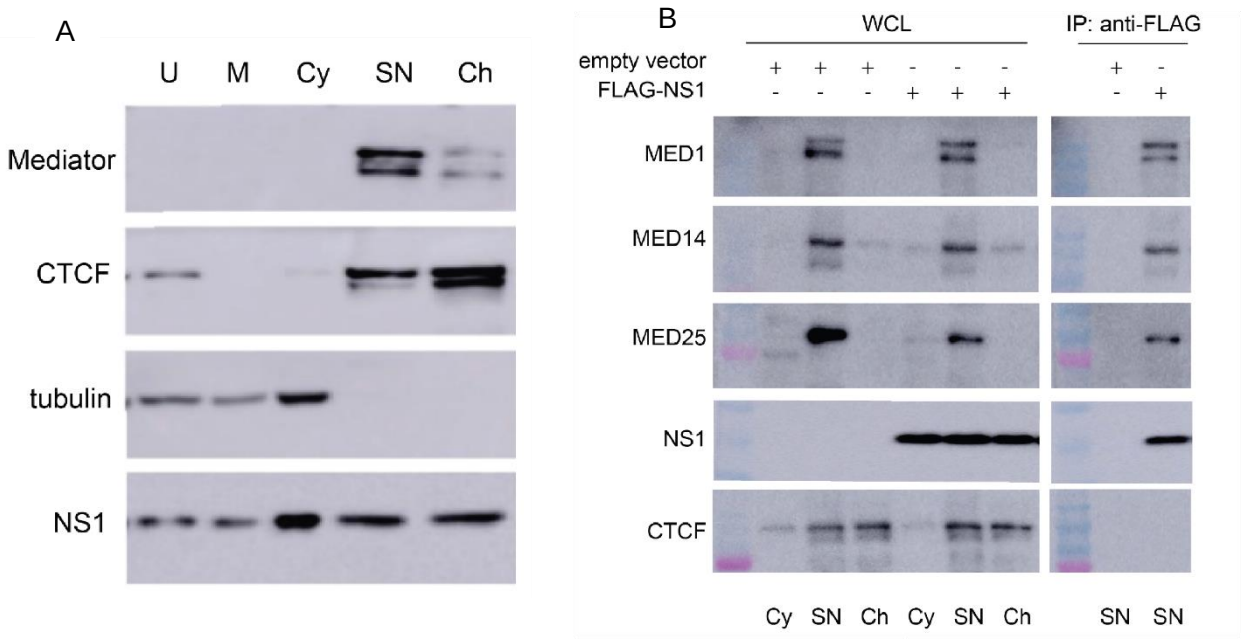


Figure 3.2. NS1 partitions to the soluble nuclear and chromatin fractions and co-immunoprecipitates with Mediator from the soluble nuclear fraction. A, Subcellular fractionation followed by Western blot for NS1 and host proteins (Jacqueline Payton). U, unfractionated; M, membrane associated; Cy, cytoplasmic; SN, soluble nuclear; Ch, chromatin. B, Western blots show co-immunoprecipitation of MED components MED1, MED14, and MED25 with an antibody to FLAG-NS1 from transfected A549 cells (Jingjing Pei). WCL, whole cell lysate. *Adapted from Pei, Beri et al Cell Reports 2021*

3.2 NS1 partitions to both soluble nuclear and chromatin associated fractions and co-immunoprecipitates with several subunits of the Mediator complex

To confirm that transfected NS1 partitions to the nucleus, my advisor Dr. Jacqueline Payton performed subcellular fractionation followed by Western blotting (**Figure 3.2A**) for NS1-transfected A549s. The blot demonstrates fractionation of Mediator and CTCF, which are nuclear and chromatin-associated proteins, and tubulin, a cytosolic protein, into the expected fractions. NS1 is detected throughout the cell including in the cytosolic, soluble nuclear, and chromatin fractions. The association between NS1 and Mediator was further confirmed by co-immunoprecipitation of transfected FLAG-NS1 with Mediator from the soluble nuclear fraction of A549s (**Figure 3.2B**).

3.3 NS1 enters the nucleus via active transport

NS1 is a protein of only 15.8 kDa, which is below the size limit of passive diffusion into the nucleus[39]. To test for active transport of NS1, Jingjing Pei of the Leung group performed immunofluorescence of GFP-tagged NS1 (**Figure 3.3**)[39]. Addition of the GFP tag increased molecular weight to ~ 42 kDa, the purpose being to increase the weight of NS1 to prevent passive transport. The cytoplasmic Cullin ligase adaptor protein Keap1 was used as a control. GFP-tagged Keap1 did not partition to the nucleus, while GFP-tagged NS1 did[39]. These data suggest that NS1 utilizes active transport to enter the nucleus. Further studies to elucidate the import mechanism are ongoing in the Leung lab. This section concludes the relevant experiments conducted by others within our groups.

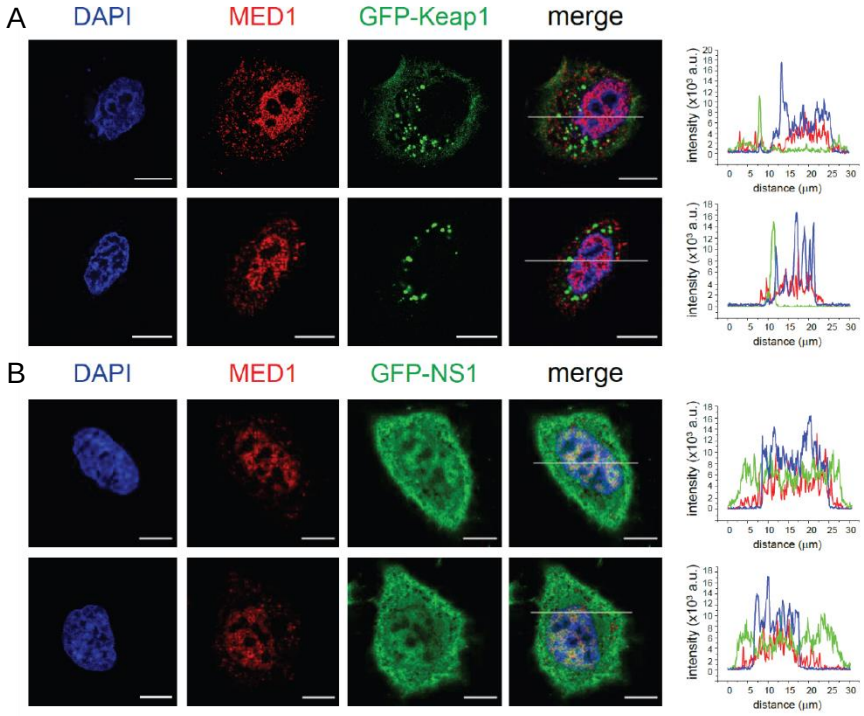


Figure 3.3. GFP-tagged NS1 partitions to the nucleus of hRSV infected cells. Confocal micrographs of A549s transfected for 24 hours with (A) GFP-tagged Keap1 or (B) NS1 and fixed. Cells were stained for DAPI (blue), anti-MED1 (red), or GFP (green). Fluorescence intensity across a representative slice of the cell is shown to the right. *Adapted from Pei, Beri, et al. Cell Reports 2021*

3.4 Chromatin immunoprecipitation for Mediator identified 21,020 binding sites genome wide

The studies above show that NS1 undergoes active transport into the nucleus and associates with chromatin and Mediator subunits. We next asked whether NS1 has a genome-wide chromatin binding profile. This protein has not previously been tested using ChIP-seq. In contrast, ChIP-seq to determine Mediator binding has been performed such that validated antibody and sequencing datasets are available[124]. Therefore, I carried out Mediator ChIP-seq as quality control to ensure my technique was working as well as to test for its genome-wide chromatin binding profile for hypothesis testing. I mapped Mediator binding with α-Med1 antibody in cells transfected with HA-NS1 or mock transfected and in cells infected with hRSV A2 or mock infected. Sequencing was performed at Washington University Genome Technology Access Center (GTAC) and reads were aligned to the human genome version hg19. Using the HOMER software suite, I used the findPeaks command line tool to call peaks.

I next validated the Mediator ChIP-seq dataset against previously published independent research. I searched the literature for Mediator ChIP datasets and downloaded a dataset from the Bilodeau group at the Université Laval in Canada[124]. They carried out ChIP-seq on three cell lines, including the A549 human lung epithelial cell line I used in my ChIP-sequencing studies, also aligned to hg19. Their group also used the same antibody against MED1, Bethyl A300-793A, that I used in my studies. However, I used EGS and formaldehyde to crosslink while they used

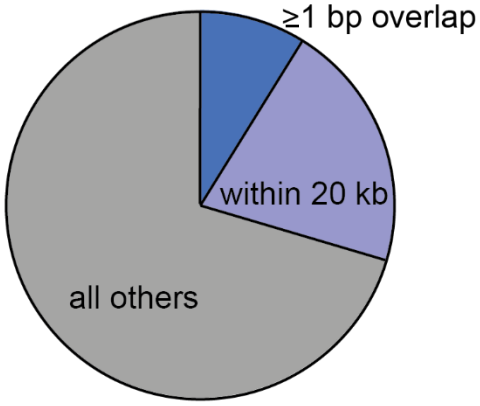


Figure 3.4 Overlap of 21,020 Mediator peaks in my dataset with previous published data (Bilodeau group). Dark blue, direct overlap of one or more nucleotide (9%); light blue, proximity of <20 kb; grey, all other Mediator peaks in my dataset.

formaldehyde only. From the Gene Expression Omnibus (GEO) repository, I downloaded the peak file generated by this group from their raw data. As they used a different peak calling algorithm, MACS2, with different specifications, the peaks from their dataset have different characteristics. In some cases, wider peaks in their dataset are called as two closely spaced peaks in my dataset. Some peaks called in their dataset were not identified in mine, and vice versa; however, 1,879 of the 21,020 peaks (8.9%) that I had identified directly overlapped peaks in the Fournier et al. dataset by at least one nucleotide, with 6,240 (30%) are within 20 kilobases of their peaks (**Figure 3.4**). With Mediator binding sites mapped genome-wide, I next evaluated NS1 binding genome-wide.

3.5 NS1 binds 1,756 sites throughout the genome at transcriptional regulatory elements

To map genome-wide chromatin binding of NS1, I used a 3HA-tagged NS1 construct to transfect A549 human lung epithelial cells. I used anti-HA to immunoprecipitate NS1 in the transfected cells and in empty vector-transfected cells as a negative control given the lack of an available α-NS1 antibody. I found that NS1 binds 1,756 discrete loci (**Figure 3.5A**). Forty-two percent coincide with Mediator binding sites in my dataset, including from NS1-transfected, empty-vector-transfected, RSV infected, and mock-infected cells (**Figure 3.5A**) with the majority overlapping at least one enhancer element[125]. Using bedtools, I found that 81% of NS1 sites lie within 10 kb of at least one gene. To calculate this, I compared the genomic coordinates of NS1 binding sites with the start –10kb and end +10kb of all genes (*Homo sapiens* (human) genome assembly GRCh37 (hg19) from Genome Reference Consortium[126]) to identify potential overlaps. This analysis mapped NS1 binding sites located within or near promoter and enhancer elements that may modulate transcription of these neighboring genes, although it was not a functional confirmation.

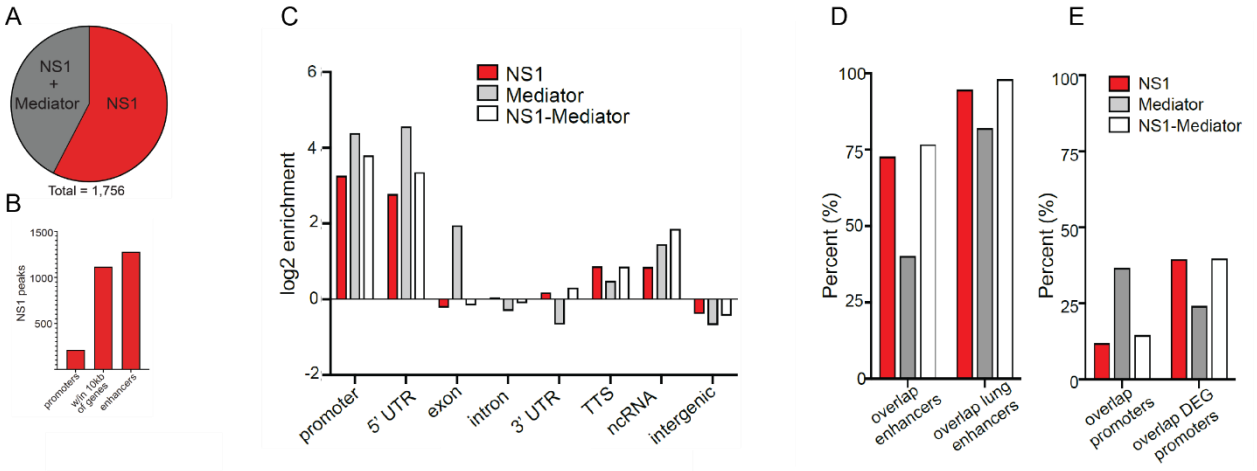


Figure 3.5. ChIP-sequencing of NS1 reveals binding sites associated with Mediator, transcriptional regulatory elements, and differentially expressed genes during hRSV infection. A, Pie chart shows total number of NS1 peaks called (red) and percent that overlap at least one Mediator peak (grey). B, Number ofNS1 sites (total of 1,756) that overlap the indicated genomic element. C, Bar graph shows genomic annotations of NS1 only (red), Mediator only (grey) or both (white) sites overlapping different genomic elements labeled on X axis. D, NS1 peaks, Mediator peaks, or NS1-Mediator overlapping peaks annotated as in (C) but overlapping all enhancers or lung specific enhancers; E, As in D but for all promoters or promoters of differentially expressed genes (DEGs). *Adapted from Pei, Beri, et al. Cell Reports 2021*

While forty percent of NS1 sites overlapped Mediator sites, a closer evaluation of the genomic annotations revealed important differences. We compared NS1 unique binding sites, Mediator only binding sites, and sites with NS1 and Mediator binding sites overlapping by at least one nucleotide. While both NS1 and Mediator are enriched at promoter and 5' UTR sites, Mediator alone is more enriched than NS1. At exons, Mediator but not NS1 is enriched (**Figure 3.5C**). We also quantified the percent of NS1 only, Mediator only, and combined binding sites that overlap enhancer elements. Seventy-three percent of NS1 binding sites overlapped enhancers, of which 95% were in lung tissue enhancers. In comparison, only 40% of Mediator sites overlapped enhancers, of which 82% were lung enhancers. For promoters, Mediator only peaks were found more frequently (40%) than NS1 only (17%) and NS1-Mediator (20%) binding sites. However, 34% of genes differentially expressed during RSV infection harbor an NS1 peak, a substantial increase over total promoters (10%). In comparison, there is a smaller proportion of differentially expressed genes

with Mediator peaks compared to total promoters (26% versus 32%). These data show that NS1 binding is enriched in transcriptional regulatory elements and that it preferentially binds near the promoters of genes that respond to RSV infection. Thus, I next evaluated the association of NS1 binding sites with specific genes differentially expressed during hRSV infection.

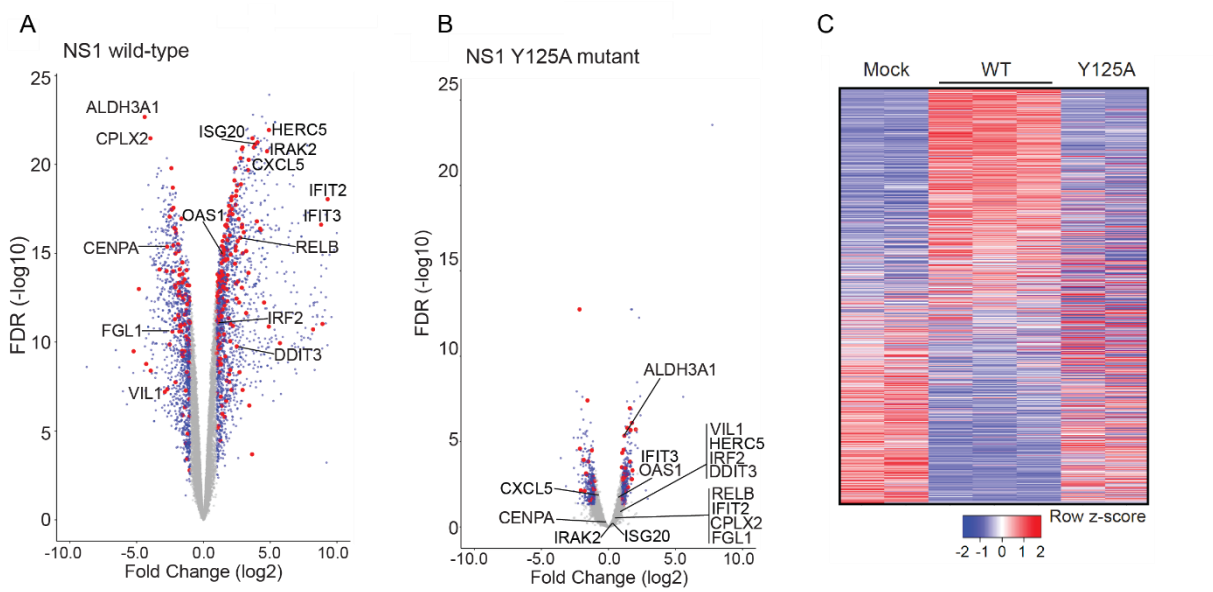


Figure 3.6. NS1 ChIP-seq reveals binding sites associated with Mediator, transcriptional regulatory elements, and differentially expressed genes during hRSV infection. A, volcano plot of gene expression in A549s at 96 h.p.i vs mock infection with hRSV bearing wild-type NS1. DEGs are in blue, and DEGs within 10 kb of an NS1 binding site are in red. B, same as (A) but comparing hRSV infection bearing NS1 Y125A to mock. C, heat map of differentially expressed genes within 10 kb of an NS1 peak in the indicated condition. *Adapted from Pei, Beri, et al. Cell Reports 2021*

3.6 NS1 peaks are enriched within 10 kilobases of genes differentially expressed during hRSV infection with WT but not Y125A NS1

To query for an association between NS1 binding sites and differentially expressed genes during hRSV infection required a relevant differential gene expression dataset. I used the DESeq2[127] package to generate differential expression data with RNA sequencing data[128] from A549 cells 96 hours after infection with WT hRSV or mock infection. Comparison to the NS1

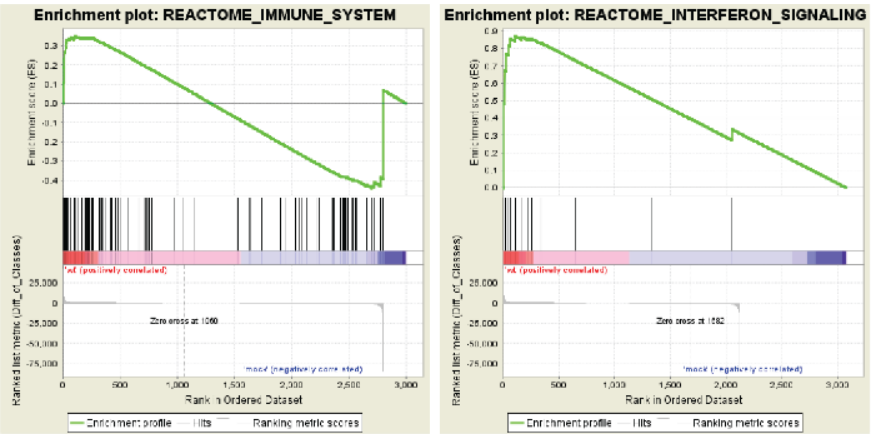


Figure 3.7 Immune response gene sets were enriched near NS1 peaks Reactome database gene sets related to the immune system (A) and interferon signaling (B) were among those enriched in genes within 10 kb of an NS1 peak that were differentially expressed between WT hRSV infection and mock infection using GSEA.

ChIP-seq data shows enrichment of differentially expressed genes (DEGs) within 10 kb of a NS1 peak. Specifically, 30% of genes within 10 kb of at least NS1 peak were DEGs (differential expression = *fold change > +/- 2; Benjamini-Hochberg adjusted false discovery rate < 0.05*, as compared to about 25% of genes overall. The volcano plot in **Figure 3.6A** shows genes that are differentially expressed between wild type and mock RSV infection at 96 hours post-infection. Genes within 10 kb of an NS1 peak are highlighted in red, demonstrating that a significant portion are among the most highly differentially expressed genes. In contrast, the gene expression profile of Y125A NS1 RSV (**Figure 3.6B, C**) infected cells is more like that of mock infected cells than that of WT NS1 RSV infected cells, with substantial reductions in fold change differences and significance values (**Figure 3.6A, C**).

Of the host genes with altered expression during hRSV infection, many such as *IRF2* and *IFIT2* are significantly up-regulated in WT NS1 hRSV infection (*IRF2*: *log2 fold change 0.996, adjusted p-value 1.3e-4; IFIT2: log2 fold change 3.7, adjusted p-value 6.9e-21*) but dramatically less so in Y125A NS1 hRSV infection (*IRF2: log2 fold change 0.49, adjusted p-value 0.25; IFIT2: log2 fold change 0.6, adjusted p-value 0.48*). I used gene set enrichment analysis (GSEA) to analyze all

genes differentially expressed in WT NS1 infection versus mock and within 10 kb of an NS1 peak. GSEA identified several enriched biological pathways, including innate immune system and interferon signaling (**Figure 3.7**) that were enriched in this subset of genes. Thus, the differential gene expression profile of hRSV infected A549s was associated with NS1 bound regulatory elements.

3.7 A cluster of NS1 binding sites at the IFIT *locus was validated via ChIP-qPCR for binding by both WT and Y125A NS1*

To validate the ChIP-seq data, I focused on NS1 binding sites near genes with altered expression during hRSV infection. The most promising of these was the *IFIT* locus, in which I identified four NS1 binding sites near four *IFIT* family genes. Site A is at a distal enhancer; site B is at the 3' UTR of *IFIT2*; site C is at the promoter of *IFIT3*; and site D is just upstream of the *IFIT1B* promoter (**Figure 3.8A**). The *IFIT* (*i*nter*f*eron *i*nduced proteins with *t*etratricopeptide repeats)[55] genes are a cluster of relatively recently diverged interferon stimulated genes in the same genomic neighborhood. HOMER called four NS1 binding sites, out of which two overlap Mediator binding sites in the same condition (NS1-transfected cells). Comparison of the NS1 and Mediator binding sites from my dataset with publicly available ENCODE datasets shows that general transcriptional coactivators such as P300, and tissue specific transcription factors may be detected in the same region (**Figure 3.8A**).

By ChIP-qPCR, I validated WT NS1 chromatin binding at three IFIT loci. I did not identify reliable and specific qPCR primers for peak A, so I assayed peaks B, C, and D (**Figure 3.8B**). DNA immunoprecipitated with HA tagged NS1 or with IgG was normalized to input for each condition. 3HA-tagged WT NS1 was enriched vs. IgG at regions C and D. I next asked whether

NS1 Y125A binds chromatin at these same regions. I found that Y125A NS1 bound DNA at least as well as WT NS1 at each of the three IFIT peaks B, C, and D.

My ChIP-seq and qPCR work described in this chapter builds on the work of my collaborators that established nuclear localization, chromatin associated subcellular partitioning, and association with Mediator complex subunits. Taken together, these findings suggest a role for the hRSV NS1 protein in modulating transcriptional regulation at chromatin associated regulatory elements.

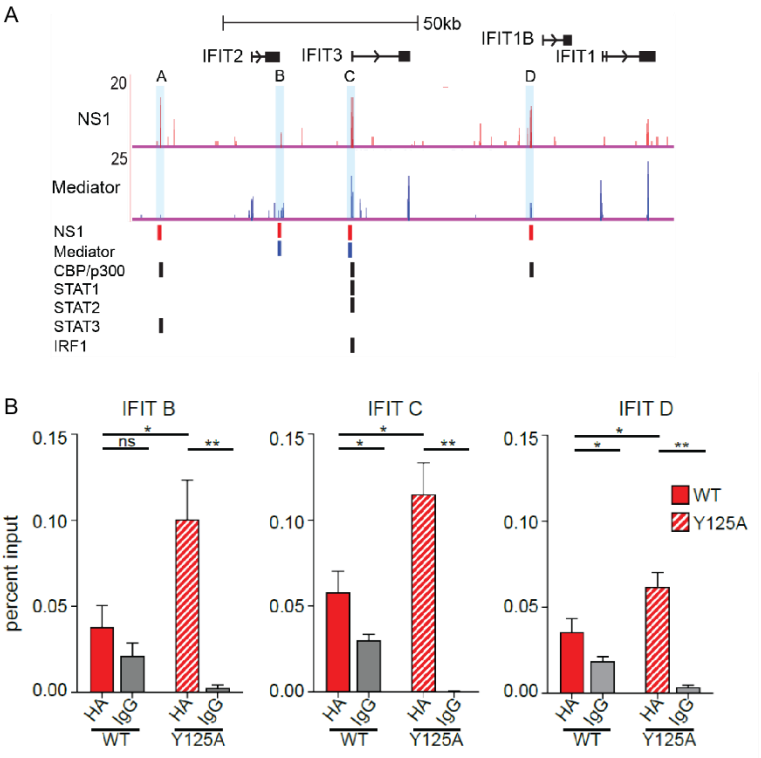


Figure 3.8. NS1 binding sites throughout the IFIT locus overlap host transcriptional regulatory factors, and both WT and Y125A NS1 bind at most of these loci. A, NS1 binding sites at the IFIT locus are labeled A through D above the raw read tracks in the UCSC Genome Browser and in red below them. Mediator tracks of raw data and called peaks are shown in blue. Host transcriptional regulators CBP/p300, STAT1, STAT2, STAT3, and IRF1 peaks from ENCODE (Sloan et al 2016) are shown below. B, ChIP-qPCR was performed on samples transfected with either 3HA-FLAG NS1 WT or 3HA-FLAG-NS1 Y125A. The results for HA-NS1 pulldown and IgG pulldown at IFIT loci B, C, and D are shown as a percent of input. *Adapted from Pei, Beri, et al. Cell Reports 2021*

3.8 Discussion

NS1 is a multifunctional interferon antagonist with several documented cytoplasmic functions[28,32,75,129]. The findings that NS1 partitions to the nucleus and associates with chromatin and several components of Mediator, a known transcriptional regulatory complex, support an additional nuclear role for NS1. The manifold roles of Mediator include its function as a

component of the pre-initiation complex[40,44], its role in bringing enhancer and promoter elements into closer proximity[45,124,130], and a role of the CDK8 domain of Mediator that is mutually exclusive with its association with RNA polymerase II[131–134].

Gene expression is regulated tightly by multiple interacting pathways[135–137]. Such tight regulation is a prominent feature of homeostasis and inappropriate deviations from the basal level of gene expression are associated with disease states (e.g., chronic inflammatory disease[138–140], cancer[141–143]). While the cytoplasmic activities of NS1 enable it to suppress immune response proteins, our data demonstrate that NS1 may also function to suppress the production of immune proteins at the level of transcription.

The presence of NS1 in the chromatin-associated fraction and in association with the Mediator suggests such a role in the modulation of gene expression. To participate in modulation of gene expression, chromatin-associated NS1 would be expected to be enriched at regulatory elements. These are non-coding regions which by merit of their sequence permit the binding of specific transcriptional regulatory proteins. Indeed, our chromatin immunoprecipitation of NS1 found it was bound at many transcriptional regulatory elements, both enhancers and promoters, and often overlapped Mediator binding sites. Together, these data demonstrated that NS1 may play a role in altered gene expression during hRSV infection via interaction with chromatin associated host transcriptional regulators at these regions.

Like the genomic annotations for the Mediator complex, NS1-only and NS1-Mediator binding sites are enriched at promoter and 5’ UTR elements. These regions are required for transcription of the gene under their control. The pre-initiation complex (PIC) and RNA polymerase II assemble at these regions. As NS1 is present at several promoter elements, it may interfere with PIC assembly, or prevent RNA Polymerase II from carrying out its elongation

function when present at the 5' UTR, as well as at genic elements such as exons and introns at the 5' end of the gene. NS1 promoter peaks are enriched near differentially expressed genes during hRSV infection, so it is possible that NS1 alters viral responsive changes in host gene expression under the control of these elements. NS1 is also enriched at lung specific enhancer elements. For enhancers that are dependent on innate immune signaling, the presence of NS1 may alter the kinetics of enhancer activation. Furthermore, in its capacity as a binding partner of the Mediator complex, NS1 may interfere in its function in bringing enhancers and promoters in proximity at actively transcribed genes. As GSEA indicates that enriched biological functions at NS1 peak proximal genes include interferon signaling and innate immune response, expression of such viral response related genes may be altered at regulatory elements at which NS1 is bound.

For instance, NS1 binding sites were identified at several regulatory elements in the *IFIT* locus. During hRSV infection the genes within this locus are differentially expressed, and the enrichment of NS1 binding sites at regulatory elements in this region suggests that NS1 plays a role in the modulation of their expression. As innate immune response genes and interferon stimulated genes, the gene products contribute to the antiviral state[144–146]. Disrupting the expression of such genes could prevent the cell from fully adopting an antiviral state.

Previous work showed that cells infected with recombinant hRSV containing NS1 Y125A mutant exhibited fewer differentially expressed genes and these had lower fold change throughout infection[128]. One potential explanation for this would be that the NS1 Y125A mutation negatively impacts NS1's ability to bind chromatin. However, ChIP-qPCR for the IFIT loci for both WT and Y125A 3HA-tagged NS1 showed that NS1 Y125A binds chromatin at least as well as WT NS1 at these regions. Another possibility is that Y125A NS1 has altered capacity to interact with key

transcriptional regulatory proteins compared to WT hRSV. Future experiments such as ChIP-seq and AP-MS for Y125A NS1 are needed to address these questions.

Signal-dependent gene expression relies on tight regulation to appropriately respond to stimulus[88,97,147]. The predominant value of the innate immune system, specifically the interferon signaling pathway that is induced in response to viral infection, lies in its capacity to rapidly detect and respond to infection[64]. Impeding this rapid response is a major function of NS1, documented to act in the cytoplasm as an interferon antagonist[26,28,30,33,34,49]. In the nucleus, targeted antagonism of host innate immune response genes would add another dimension to this function. Potentially, transcription of genes encoding products required for hRSV life cycle functions might be unaffected or even induced further while transcription of genes encoding innate immune response factors might be suppressed.

Chapter 4: In the presence of NS1, gene expression is altered in several conditions relevant to the innate immune response

4.1 Genomic regulatory elements overlapping NS1 binding sites drive altered transcription in the presence of NS1

The genome-wide chromatin binding studies showed that NS1 often overlaps regulatory elements, transcription factor binding sites, and Mediator binding sites, suggesting that this viral protein may impact host transcriptional regulation. Therefore, I next sought to define the functional effect of NS1 on gene expression. Given that Mediator does not coincide with all NS1 binding sites, I hypothesized that NS1 may interact with other transcription factors and/or transcriptional regulatory proteins. Indeed, I identified overlap of IRF, STAT, and p300 binding sites with NS1 peaks in the IFIT locus (**Figure 3.8**).

Therefore, I subjected all NS1 binding sites to transcription factor motif analysis and comparison to publicly available transcription factor ChIP-seq data. I used HOMER findMotifsGenome to identify transcription factor binding motifs enriched at the NS1 peaks. Although no single transcription factor dominated the enrichment results, motifs for the AP-1 (p-*value 1e-373, 13-fold increase over background*) and FOXA1 (*p-value 1e-433, 6-fold increase over background*) transcription factors were identified. While highly enriched, these transcription factors are ubiquitous and not specific to innate immune response. Enriched motifs for transcription factors involved in innate immune response included STAT3/IL-21 (*p-value 1e-3, 1.4 fold increase over background*), CHOP (*p-value 1e-12, 2.7 fold increase over background*), and STAT1 (*p-value 1e-2, 1.5 fold increase over background*).

As queries using HOMER gave in silico predictions, I next queried Factorbook[148], a repository of publicly available transcription factor binding datasets, for ChIP-seq datasets for relevant

transcription factors. Because NS1 is a multifunctional interferon antagonist[22,28–30,32,33,129], I searched for transcription factors within the STAT and IRF family. These are mediators of interferon induction and signaling pathways[36,37,69,71,149].

To identify the ChIP-seq datasets most relevant to my project, I sought ChIP-seq datasets performed in A549 cells that were used in my NS1 and Mediator ChIP-seq experiments. I did not find datasets for innate immune response transcription factors STAT1, STAT2, STAT3, STAT5A, and IRF3 and regulatory factors PGC1A and CEBPB in A549s, but rather used existing datasets from HeLa, IMR90, K562, and HepG2 cell lines. I searched for STAT and IRF family transcription factors as they are involved in the innate immune response (**Figure 1.2**). PGC1A is a component of the epigenetic silencing Polycomb complex[150] . Next, I used bedtools window to identify peaks in my NS1 ChIP-seq dataset associated via proximity with significantly upregulated genes in hRSV infection. I used window rather than intersect, the standard bedtools function used for identifying overlapping genomic features, because regulatory elements may be some distance away from the genes for which they control expression. Starting with a 1-kb window upstream and downstream of all genes, I tested a 5-kb window, 10-kb window, and 20-kb window. After increasing past 10 kb, I found that there were few additional NS1-gene pairs identified.

The four IFIT locus NS1 peaks were noteworthy, as in addition to sequencing data, I had validated binding at three of the four (B, C, and D) via ChIP-qPCR for both WT and Y125A NS1. Thus, I cloned the regulatory elements containing these peaks into luciferase reporter plasmids to test whether these regulatory elements drove altered gene expression in the presence of NS1.

4.2 Luciferase reporter assay was used to test changes in transcription driven by bound regulatory elements in the presence of NS1

Given that NS1 associates with Mediator and is bound at regulatory elements enriched near genes differentially expressed during hRSV infection, the evidence to this point suggests that NS1 may alter gene expression controlled by regulatory elements at which it binds. The luciferase reporter assay is commonly used to test the effect of a stimulus or other molecular factors on

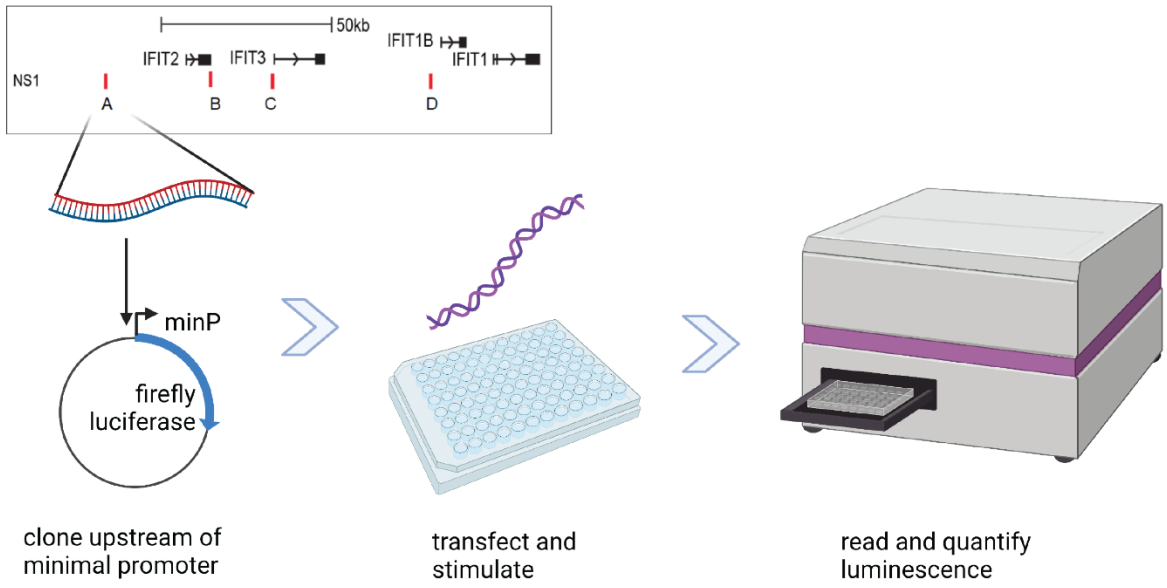


Figure 4.1. Luciferase assay workflow. A 400-1000 base pair region encompassing an NS1 binding site was cloned into the pGL4.45 firefly luciferase plasmid upstream of the minimal promoter (minP). A plasmid constitutively expressing NS1 (or empty vector), a constitutively active nanoLuc expressing plasmid (pNL1.1.TK), an immune stimulus (or none), and the experimental luciferase plasmid were transfected into 293Ts. 24 hours later, firefly luminescence and nanoLuc luminescence were read on a BioTek plate reader with Gen5 software (version 3.08). The ratio of firefly to nanoLuc was taken to normalize for transfection efficiency and the ratio for each well was compared to that for the minimal promoter in the same condition.

transcription under the control of a known regulatory element or to compare transcription driven by different regulatory elements. Thus, I used a luciferase reporter assay to test whether transcription driven by IFIT locus regulatory elements was altered in the presence of NS1 (**Figure 4.1**). As known NS1 functions antagonize the innate immune response, and infections with virus

bearing a3 helix structure-based mutants of NS1 were attenuated in IFN-competent but not IFN-null cells[30], the nuclear function of NS1 is likely to be associated with the type I IFN response. Therefore, I tested the effect of NS1 on reporter gene expression in stimulated cells, having activated the signaling pathways culminating in upregulation of innate immune response genes.

To stimulate cells, I first used polyI:C, a double stranded RNA mimic that strongly induces signaling through both RIG-I and TLR3[76,93]. Both the ISRE and NF-kB regulatory elements are bound by transcription factors involved in type I IFN production[71,96,97,147], as IRF family members and NF-kB are activated downstream of double-stranded RNA detection, but ISRE motifs alone are bound by ISGF3 in the induction of interferon stimulated genes (**Figure 1.2**).

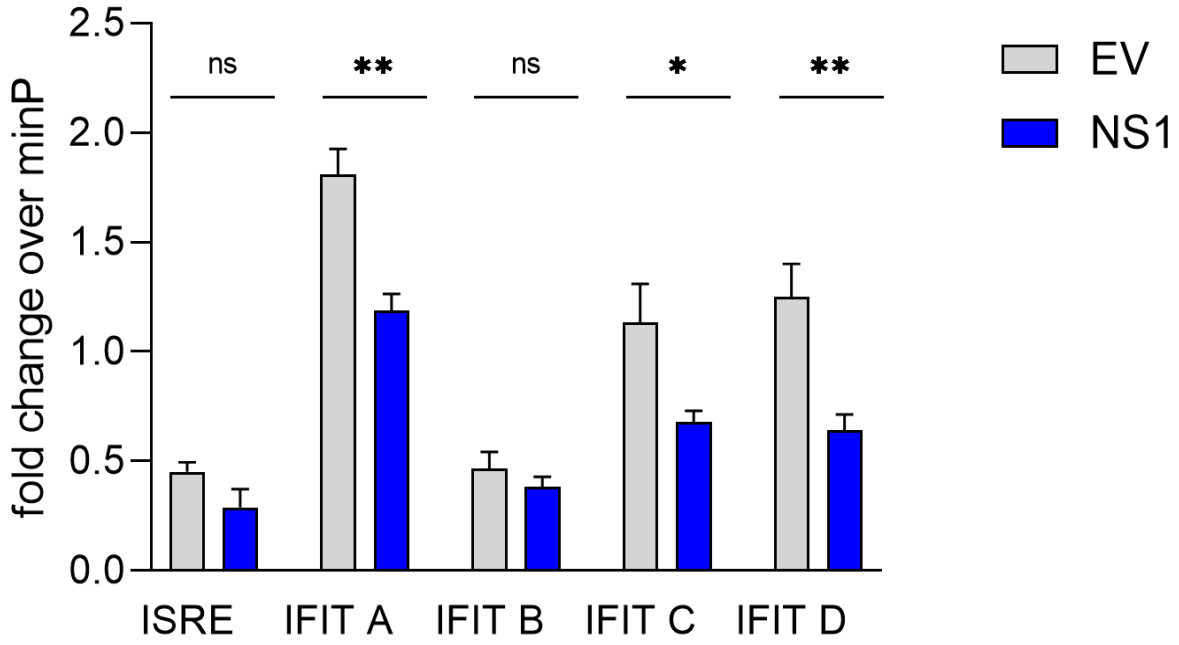


Figure 4.2. Transcription driven by NS1 binding sites in IFIT locus regulatory elements is decreased in the presence of NS1. Bar graph shows luciferase activity for the ISRE-5x plasmid (positive control) and regulatory elements cloned from NS1 binding sites in the *IFIT* locus (see locus map). Cells were transfected with the indicated reporter plasmid, control luciferase plasmid (nanoLuc), empty vector (EV) or NS1 plasmid, and stimulated with polyI:C. Twenty-hour hours later, luciferase luminescence was measured. Fold change = (ratio of firefly ISRE-5x or IFIT plasmid to nanoLuc) / (ratio of minimal promoter reporter plasmid to nanoLuc). EV, pCAGGS empty vector. NS1, pCAGGS-3HA-FLAG-NS1 WT. Representative of two independent experiments. Unpaired t-test, * $p < 0.05$, ** $p < 0.005$, ns, not significant. Bars show mean with standard deviation. *Adapted from Pei, Beri, et al, Cell Reports 2021*

4.2 Reporter expression driven by IFIT locus regulatory elements is decreased in the presence of NS1

Using luciferase reporter assays, I tested the regulatory impact of NS1 on four NS1 bound regulatory elements throughout the *IFIT* locus. I cloned a several hundred base pair region encompassing each NS1 binding site upstream of a minimal promoter in a luciferase reporter plasmid. In triplicate wells of a 96-well plate, I transfected 293Ts with 4 ug/ml polyI:C, a pGL4-based firefly luciferase reporter plasmid, the transfection control plasmid pNL1.1.TK, and a plasmid expressing NS1 or an empty vector. Twenty-four hours after plating, I read the luciferase luminescence on a plate reader. I took the ratio of firefly luciferase activity to the nano luciferase, the latter of which was driven by a constitutive TK promoter and functioned to normalize for transfection efficiency. I next normalized this ratio to the baseline luciferase activity level of the minimal promoter, which was the same backbone plasmid into which each NS1 binding site region was cloned. For each condition, I calculated these ratios for cells co-transfected with empty vector or NS1.

As a positive control, I used the interferon stimulated regulatory element repeat (ISRE-5x) which contains an optimized commercially available promoter with the ISRE motif repeated five times. It is highly responsive to polyI:C stimulus[151–153]. For the ISRE-5x promoter under these conditions, there was not a significant change in expression upon the addition of NS1 (**Figure 4.2**). However, later experiments using 14.3 ug/ml polyI:C did yield upregulation in ISRE-5x mediated reporter expression suggesting that more stimulus was required to observe a strong activation.

For *IFIT* peaks A, C, and D, there was a significant decrease in expression in the presence of NS1. A is at a distal enhancer nearest to *IFIT2* and D is at an enhancer just upstream of the *IFIT1B*

promoter, while C overlaps the *IFIT3* promoter. This data shows that gene transcription is significantly lower in the presence of NS1 when driven by *IFIT* genomic regulatory elements at which NS1 binds, including both promoter and enhancer elements.

4.3 Both WT and Y125A NS1 decrease transcription driven by optimized ISRE-5X and NF-kB RE-4X regulatory elements

Having shown that transcription driven by NS1-binding genomic regulatory elements is decreased in the presence of WT NS1, I next tested the effect of the Y125A point mutant of NS1. Briefly, infection with recombinant RSV containing this variant of NS1 substantially reduces the number of DEGs at 96 h.p.i compared to WT NS1 (**Figure 3.6**). In a luciferase reporter assay, IFNβ promoter-driven transcription was inhibited by NS1 Y125A, but less than half as much as by WT NS1[128]. Thus, I asked whether the Y125A mutant would impact the capacity of NS1 to suppress transcription driven by these promoters.

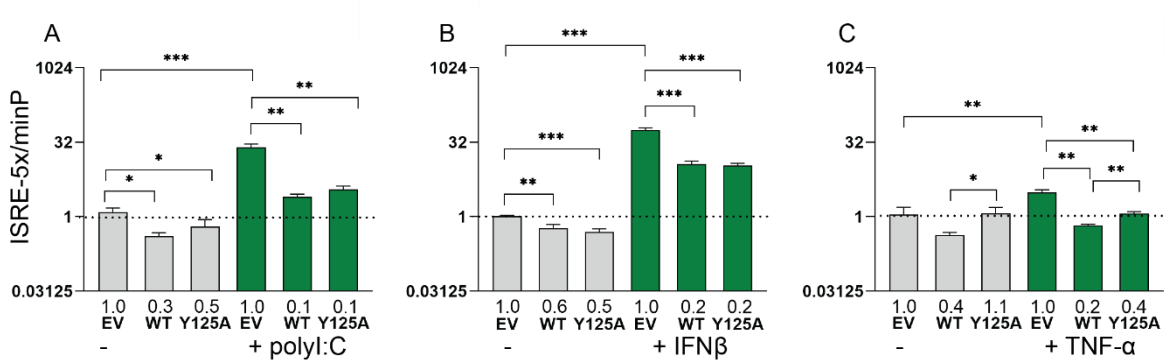


Figure 4.3. Transcription driven by optimized ISRE-5x reporter construct is lower in the presence of either WT or Y125A NS1. Bar graphs show the fold change of the firefly/nanoLuc ratio for the ISRE-5x promoter normalized to the minimal promoter/nanoLuc ratio for the same condition. Grey bars, unstimulated and green bars, stimulated in A, with polyI:C stimulus; B, with IFNβ stimulus; C, with TNF-α. EV, pCAGGS empty vector; WT, pCAGGS-3HA-FLAG-NS1 WT; Y125A, pCAGGS-3HA-FLAG-NS1 Y125A. Numbers below bars indicate the level of expression compared to empty vector transfected cells with the same stimulus. Unpaired *t*-test, *, $p < 0.05$; **, $p < 0.005$; ***, $p < 0.0005$. Representative of three or more independent experiments.

Having the ISRE-5x construct at my disposal, I tested the effect of WT and Y125A NS1 on ISRE-driven transcription in stimulated 293Ts. A plasmid encoding NS1 WT or Y125A point mutant (or an empty vector plasmid) was transfected with the ISRE-5x reporter construct and cells were stimulated with the indicated reagent. To demonstrate the varying impact of different immune stimulants on different reporter constructs, **Figures 4.3 – 4.6** show unstimulated and stimulated data for each stimulus/reporter combination. As expected, polyI:C stimulation of 293Ts transfected with the ISRE-5x plus an empty vector construct showed 21-fold increased luciferase activity over the ISRE-5x in unstimulated cells (**Figure 4.3A**). For cells transfected with WT NS1, polyI:C stimulus led to only a six-fold increase in reporter expression, three-fold lower than the change in stimulated cells transfected with the empty vector. For cells transfected with Y125A NS1 there was a five- and a half-fold increase in luciferase activity in the polyI:C stimulated condition vs unstimulated, or similar to the six-fold increase observed for polyI:C stimulated cells transfected with WT NS1. Notably, even without polyI:C stimulation, the presence of NS1 decreases luciferase activity driven by the ISRE compared to cells transfected with EV: three-fold lower than EV for WT NS1 and two fold lower for Y125A NS1.

The same experiment was repeated in interferon-β stimulated cells (**Figure 4.3B**), and a similar pattern was observed. A 55-fold increase in ISRE-driven reporter expression was observed upon IFN-β stimulus in EV-transfected cells, but only a 20-fold increase in the presence of either WT or Y125A NS1. In the uninduced condition, expression was halved in the presence of either WT or Y125A NS1. The fifty five-fold increase in luciferase activity of the ISRE-5x induced by IFNβ stimulus was almost three times that induced by polyI:C. In the presence of WT or Y125A NS1 luciferase activity decreased almost three-fold compared to IFNβ plus empty vector, which

was similar to the reduction of expression observed in polyI:C stimulated cells in the presence of NS1.

Finally, I tested the impact of tumor necrosis factor α (TNF-α) stimulus on ISRE-5x (**Figure 4.3C**). TNF-α stimulus led to only a three-fold increase in ISRE-5x driven reporter expression. This is expected because TNF-α does not strongly activate transcription factors that bind ISRE-5x. The increase in expression in TNF-α stimulated cells was twenty fold less than that of IFN-β and eight fold less than that of polyI:C. Expression of WT or Y125A NS1 in TNF-α stimulated cells reduced luciferase activity about five-fold (WT) or three-fold (Y125A) compared to TNF-α plus empty vector. Activity in the presence of WT or Y125A NS1 was similar to that with IFN-β but less reduction than with polyI:C, but for Y125A was less reduction than with polyI:C or IFN-β.

Next, I asked whether the suppressive effect on transcription observed in the presence of NS1 was specific to pathways that activate transcription factors specific to the ISRE-5x construct or whether it was applicable to other transcriptional pathways involved in the innate immune response. NF-KB is another major pathway of immune response[154–157]. While the NF-kB pathway (**Figure 1.2**) is induced by TLR3 downstream of dsRNA detection, it is not known to be strongly induced downstream of IFNβ stimulus, but rather is a target of signaling downstream of TNF-

α[68,154,155]. Therefore, I decided to test an optimized NF-kB reporter construct (NF-kB-4x) in these luciferase assays.

The NF-kB-4x luciferase construct is induced with polyI:C two and a half-fold over unstimulated, which, as expected, is lower (eight-fold) than the ISRE-5x construct (**Figure 4.4A**).

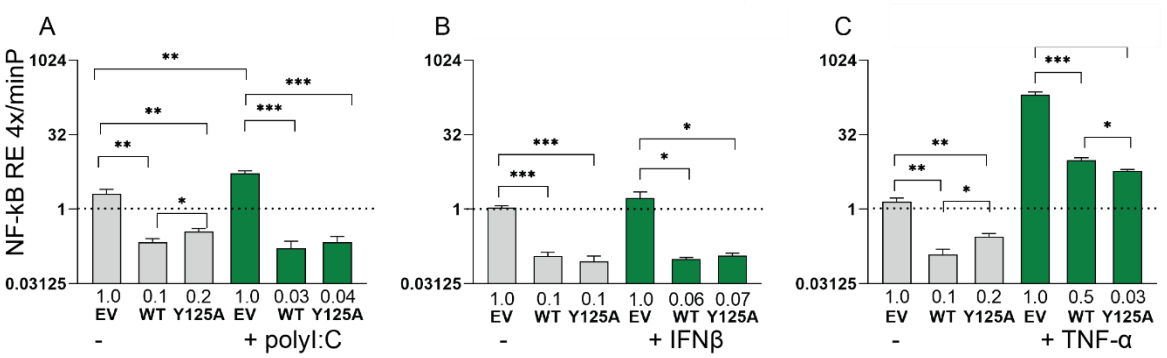


Figure 4.4. Transcription driven by optimized NF-kB-4x reporter construct is lower in the presence of either WT or Y125A NS1. Bar graphs show the fold change of the firefly/nanoLuc ratio for the ISRE-5x promoter normalized to the minimal promoter/nanoLuc ratio for the same condition. Grey bars, unstimulated and green bars, stimulated in A, with polyI:C stimulus; B, with IFNβ stimulus; C, with TNF-α. EV, pCAGGS empty vector; WT, pCAGGS-3HA-FLAG-NS1 WT; Y125A, pCAGGS-3HA-FLAG-NS1 Y125A. Numbers below bars indicate the level of expression compared to empty vector transfected cells with the same stimulus. Unpaired *t*-test, *, $p < 0.05$; **, $p < 0.005$; ***, $p < 0.0005$. Representative of three or more independent experiments.

Even without stimulation, luciferase activity was decreased ten-fold in the presence of NS1 WT and five to ten-fold in the presence of NS1 Y125A. The magnitude of the decrease in reporter expression driven by the NF-kB-4x in the presence of NS1 (WT or Y125A) was about three-fold greater than that driven by the ISRE-5x promoter in polyI:C or IFN-β stimulated cells. In TNF-α stimulated cells the difference in magnitude between ISRE-5x and NF-kB RE-4x was even greater, four- (WT) to twelve-fold (Y125A). PolyI:C-induced NFkB-RE-4x driven reporter expression was decreased three-fold more in the presence of NS1 than was ISRE-5x driven reporter expression. Unlike ISRE-5x and as expected, NF-kB-4x driven transcription is not induced by IFNβ treatment (**Figure 4.4B**). NF-kB-4x driven reporter expression was reduced around ten-fold in the presence

of NS1 WT or Y125A for the unstimulated cells, 30-fold in polyI:C stimulated cells, about 20-fold in IFN-β stimulated cells, and 20- (WT) to 40-fold (Y125A) in TNF-α stimulated cells. In contrast, and as expected, the NF-kB-driven luciferase activity is strongly induced >140-fold in response to TNF-α stimulus (**Figure 4.4C**). TNF-α induced NF-kB-driven luciferase expression was reduced 20-fold in the presence of either WT or Y125A NS1. This reduction is four (WT) to eight (Y125A) times greater than the five-fold decrease in IFNβ stimulated ISRE-5x activity in the presence of WT NS1.

To summarize, I found that gene expression in the presence of NS1 decreases for three of four NS1-binding *IFIT* locus regulatory element reporters. I also found that NS1 represses transcription induced by polyI:C- and IFN-β-mediated signaling. Indeed, the ISRE-5x construct is repressed in

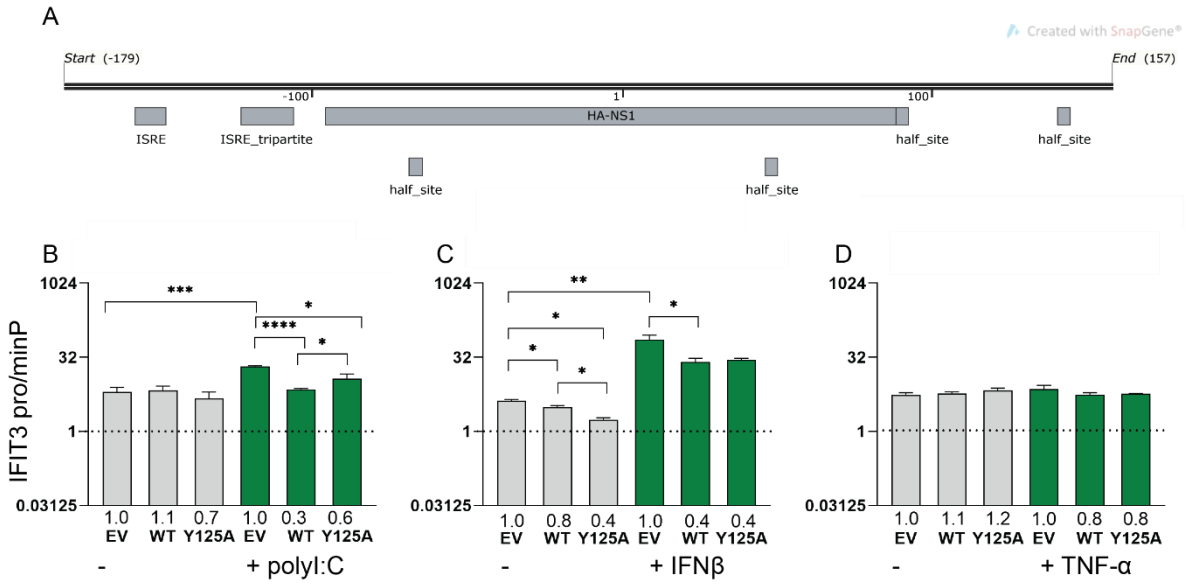


Figure 4.5. Transcription driven by the *IFIT3* promoter is lower in the presence of WT or Y125A NS1. A, map of the *IFIT3* promoter region cloned upstream of the minimal luciferase promoter in pGL4.23. TSS is at position 1. B – D, Grey bars, unstimulated and green bars, stimulated. Stimuli: B, polyI:C; C, IFNβ; D, TNF-α. EV, pCAGGS empty vector; WT, pCAGGS-3HA-FLAG-NS1 WT; Y125A, pCAGGS-3HA-FLAG-NS1 Y125A. The y axis shows the fold change the firefly/nanoLuc ratio for the ISRE-5x promoter over that for the minimal promoter in the same condition. Numbers below bars indicate the level of expression compared to empty vector transfected cells with the same stimulus. Unpaired *t*-test, *, $p < 0.05$; **, $p < 0.005$; ***, $p < 0.0005$. Representative of three or more independent experiments.

the presence of WT NS1 in both polyI:C (ten-fold) and IFN-β (five-fold) stimulated cells. Furthermore, in the presence of Y125A NS1, reporter expression driven by this construct also decreases in polyI:C (seven-fold) and IFN-β (five-fold) stimulated cells. Finally, in an uninduced condition, reporter expression is suppressed up to ten-fold in the presence of WT or Y125A NS1.

4.4 In the presence of NS1, IFIT3 promoter drives lower transcription while ISG20 *promoter full length and truncation variants drive similar or higher levels of transcription in a motif-specific manner*

The ISRE-5x and NF-kB-4x reporter constructs used in studies above were optimized for activation by their respective stimulation pathways. Genomic regulatory elements are multifaceted and multifunctional, with some having complex networks of transcription factors that enable tight regulation of transcription. For stress responsive genes such as those involved in the innate immune response, tight regulation is especially important[158–160]. Runaway or uninhibited gene expression carries the risk of adverse reactions. Therefore, I wanted to ask whether, like the optimized luciferase promoter constructs, NS1-bound regulatory elements would drive reduced transcription in the presence of NS1.

I first selected the *IFIT3* promoter, which encompasses an NS1 binding site (**Figure 4.5**)**.** *IFIT3* is an interferon stimulated gene whose protein product binds and inhibits cellular protein activity during infection[145,146]. *IFIT3* is upregulated during hRSV infection (**Figure 3.6A**). The *IFIT* C reporter contains the promoter for *IFIT3* (**Figure 4.2**) but was cloned using different primers. This cloned fragment is a 336 bp fragment overlapping the transcription start site of an *IFIT3* isoform. It contains one NS1 binding sites and two variations of the ISRE motif as well as several GAAA half-sites.

I tested the *IFIT3* promoter using luciferase assays in cells stimulated with polyI:C, IFN-β, or TNF-α (**Figure 4.5**). Unlike the ISRE-5x, *IFIT3* promoter-driven activity was not consistently reduced by NS1 when uninduced. Both polyI:C (three-fold) and IFN-β (18-fold) induced luciferase expression. In the polyI:C induced condition (**Figure 4.5B**), addition of WT (three-fold) and Y125A NS1 (two-fold) was associated with decreased expression driven by the *IFIT3* promoter. With IFN-β induction (**Figure 4.5C**), expression in the presence of WT or Y125A NS1 was decreased three-fold relative to EV. However, TNF-α (**Figure 4.5D**) did not alter expression driven by the *IFIT3* promoter, and there was no significant change in expression in the presence of either WT or Y125A NS1 in TNF-α treated cells. The transcriptional changes observed with these stimuli agree with their roles in the innate immune response. *IFIT3* is an interferon stimulated gene, and when I searched the full sequence for transcription factor binding sites (**Figure 4.4A**) there were no kappa B sites. Thus, upregulation of gene expression under the control of the *IFIT3* promoter is driven by ISRE binding transcription factors, but not by NF-kB family members. *IFIT3* promoter-driven luciferase expression decreased in the presence of NS1 WT in polyI:C or IFN-β stimulated cells. Y125A NS1 also decreased IFIT3-driven luciferase, but less than WT, and the difference was significant only in the presence of in polyI:C, not IFN-β.

While the ISRE-5x construct was induced 21-fold by polyI:C relative to unstimulated, the *IFIT3* promoter was induced only three-fold. Similarly, IFN-β induced 55-fold higher expression from the ISRE-5x construct and 18-fold increased expression from the *IFIT3* promoter. Expression decreased by 10-fold in the presence of WT NS1 in polyI:C-stimulated cells under control of the ISRE-5x, but only three-fold for the *IFIT3* promoter. ISRE-5x driven expression decreased seven-fold and *IFIT3* promoter driven expression decreased two-fold in the presence of NS1 in polyI:C stimulated cells.

For IFN-β stimulated cells, ISRE-5x drove five-fold lower expression in the presence of WT NS1 or Y125A NS1 than in the presence of empty vector. For the *IFIT3* promoter, expression in IFN-β stimulated cells decreased by about three-fold with the addition of WT NS1 or Y125A NS1.

4.5 Reporter expression driven by the ISG20 *promoter is increased in the presence of NS1*

Next, I returned to my list of RSV-infection DEGs with NS1 binding sites at nearby regulatory elements. *ISG20* was notable as this gene encodes an endoribonuclease that has activity against multiple RNA viruses[161–164]. There is an NS1 binding site just upstream of the transcription start site for two of three isoforms shown on GENCODE V36.

I tested the *ISG20* promoter (**Figure 4.6**) in a luciferase promoter assay as before. In this case, in unstimulated cells expression increased 1.5-fold in the presence of WT NS1 but not NS1

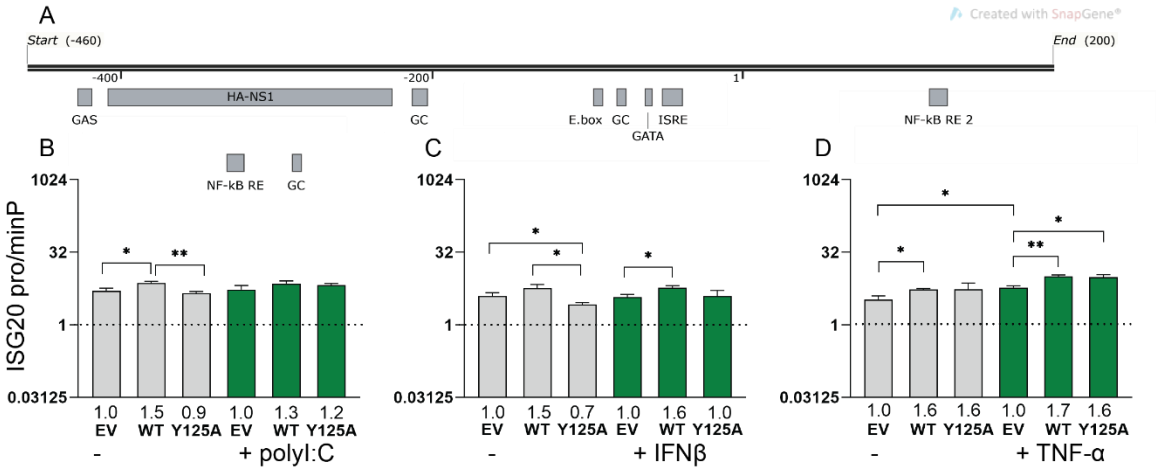


Figure 4.6. Transcription driven by the *ISG20* promoter reporter is increased in the presence of either WT or Y125A NS1. A, cartoon of the cloned region of the *ISG20* promoter, with transcription factor binding sites identified by Gongora et al and/or myself indicated. Gene start is at position 1 and proceeds to 200. Created in SnapGene Viewer. Grey bars, unstimulated and green bars, stimulated in B, with polyI:C stimulus; C, with IFNβ stimulus; D, with TNF-α. EV, pCAGGS empty vector; WT, pCAGGS-3HA-FLAG-NS1 WT; Y125A, pCAGGS-3HA-FLAG-NS1 Y125A. The y axis shows the fold change the firefly/nanoLuc ratio for the ISRE-5x promoter over that for the minimal promoter in the same condition. Numbers below bars indicate the level of expression compared to empty vector transfected cells with the same stimulus. Unpaired *t*-test, *, $p < 0.05$; **, $p < 0.005$; ***, $p < 0.0005$. Representative of three or more independent experiments.

Y125A. There was no significant change in expression in polyI:C stimulated cells in the presence of either WT or Y125A NS1. In IFN-β stimulated cells, expression again increased 1.5-fold with the addition of WT NS1 but did not significantly change with the addition of Y125A NS1. In TNF-α stimulated cells, both WT and Y125A NS1 were associated with a 1.67-fold increase in luciferase expression vs. empty vector.

To conclude this section, I found that in the presence of NS1, the *IFIT3* promoter drives about three-fold less transcription in polyI:C stimulated cells than without. Similarly, it drives about three-fold as much transcription in IFN-β stimulated cells in the presence of NS1 compared to without. In unstimulated cells, or those stimulated with polyI:C or IFN-β, decreased reporter expression is also observed. In contrast, the *ISG20* promoter transcription is essentially unchanged with NS1 and polyI:C or IFNβ, but drives higher levels of transcription in the presence of and TNF-α.

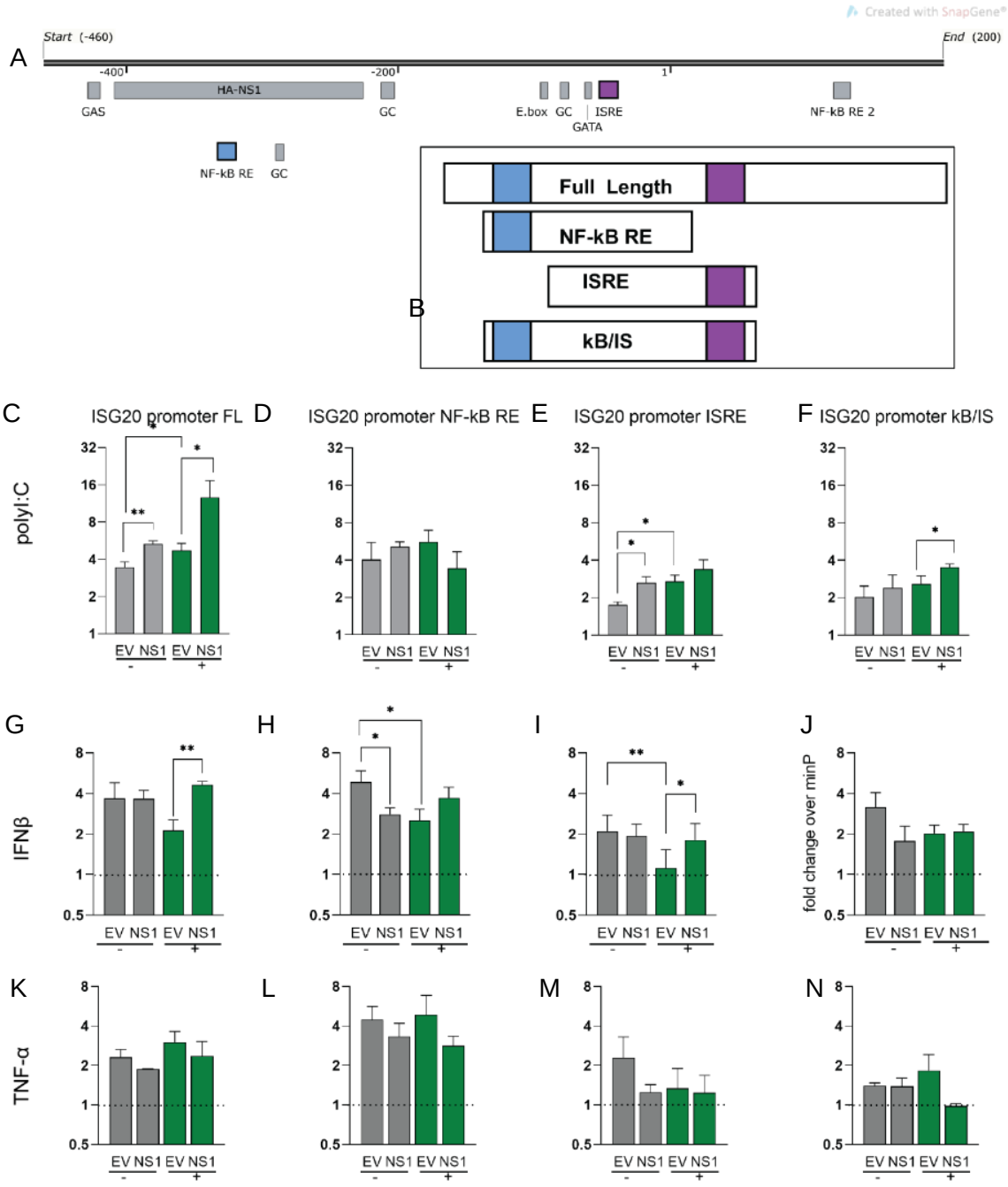


Figure 4.7. Truncated variants of the *ISG20* promoter may drive either a decrease or increase in expression in response to stimulus and in the presence of NS1. A, the *ISG20* promoter as shown in **Figure 4.6** but with the motifs of interest highlighted: NF-kB RE in blue, and ISRE in purple. B, a simplified cartoon of the full length *ISG20* promoter and truncation mutants indicating the relative locations of the indicated transcription factor binding sites. C-N, luciferase reporter assay was performed as before using the indicated variant of the *ISG20* promoter to drive luciferase expression. Bars in grey are unstimulated and bars in green are stimulated. Unpaired t-test, *, $p < 0.05$; **, $p < 0.005$. All graphs are representative of at least two independent experiments.

4.6 Truncation variants of the ISG20 *promoter respond differently to stimulus and in the presence of NS1*

About two decades ago, the Mechti group[165] identified several transcription factor binding sites (TFBS) throughout the *ISG20* promoter: GAS, NF-kB RE, all GC boxes, E.box, GATA, and ISRE, some of which agree with elements I manually identified: HA-NS1, NF-kB RE, ISRE, and NF-kB RE 2. Therefore, I decided to break down the *ISG20* promoter construct, as originally cloned, into segments containing either the NF-kB motif (I20 NF-kB RE), the ISRE motif (I20 ISRE), or both (I20 kB/IS), but excluded the GAS identified by the Mechti group and the second NF-kB RE that I manually identified (**Figure 4.7B**).

I then tested the full length *ISG20* promoter, or I20 FL, along with I20 NF-kB RE; I20 ISRE; and I20 kB/IS (**Figure 4.7C-N**). Expression driven by I20 FL in cells stimulated with polyI:C increased two and a half-fold in the presence of NS1. I20 NF-kB RE was induced by one and a half fold. I20 ISRE was induced by two-fold but had basal expression of about half that of I20 FL. I20 kB/IS was also upregulated two-fold by NS1. In the experiment shown, the *ISG20* promoter was not induced by IFN-β or TNF-α.

With IFN-β stimulus, reporter expression decreased under the control of all *ISG20* promoter variants. When NS1 was added to IFN-β stimulated cells, expression increased two-fold for the ISG20 FL construct and one and a half fold for either the ISG20 NF-kB RE or ISG20 ISRE construct but remained about the same for the ISG20 kB/IS construct.

With TNF-α stimulus, reporter expression was not upregulated in the experiment shown for any of the tested *ISG20* promoter variants. Neither was there a significant change in expression in the presence of NS1.

However, the Mechti group observed an induction of 7.2-fold[165] by polyI:C of their longest *ISG20* promoter construct[165]. This construct is 170 nt longer than my ISG20 FL construct. This indicates to me that I could use a longer promoter construct to observe higher induction. However, the Mechti group observed consistent upregulation of *ISG20* promoter driven reporter expression even for truncated variants of the promoter[165]. Thus, quantity and choice of IFN stimulus could be further optimized for induction of truncated *ISG20* promoter variants. The experiment could be conducted in different cells to better recapitulate the physiological conditions of induction.

4.7 Discussion

In Chapter 4, I asked whether NS1 acted as a transcriptional modulator of gene expression driven by genomic regulatory elements at which it binds. I used optimized reporter plasmids to test the effect of NS1 on promoters regulated by specific TFs. Then I cloned NS1-bound genomic regulatory elements to test the effect of NS1 on a more physiological DNA sequence. To provide immune signaling context, I used a dsRNA mimic, IFN-β, or TNF-α to activate signaling pathways that alter transcription downstream of detected pathogens.

Because I found that fewer than half of NS1 sites overlapped Mediator binding sites, yet nearly all were in transcriptional regulatory elements, I hypothesized that NS1 interacts with other transcription factors. RIG-I signaling activates the IRF3 and IRF7 transcription factors, members of the IRF family of transcription factors[36,76,82]. IRF family transcription factors, which are involved in the innate immune response, bind the ISRE motif, for which the consensus is GAAANNGAAA[87,149,149,166–168]. TLR3 is the other major pattern recognition receptor that recognizes dsRNA[64,93,94,94,95]; the NF-kB transcription factor is activated downstream of the TLR3 signaling pathway. The consensus motif for NF-kB is GGGNNTTCCC[154,155,169,170].

I used optimized, commercially available reporter plasmids to test gene expression in the presence of NS1 under the control of promoters carrying each motif. An optimized reporter plasmid ISRE-5x, which contains five repeats of the GAAANNGAAA motif, drives increased reporter expression in response to innate immune stimuli, including polyI:C, IFN-β. It also drives increased expression TNF-α stimulated cells. I showed that ISRE-5x also drives lower gene expression in the presence of NS1. Another optimized reporter plasmid, NF-kB-4x, which contains four repeats of GGGNNTTNCC and is stimulated by polyI:C and TNF-α, also drove lower reporter gene expression in the presence of NS1. However, the NF-kB-4x drove relatively lower expression upon the addition of NS1 in polyI:C stimulated cells than did the ISRE-5x. For both ISRE-5x and NF-kB-4x optimized reporter plasmids, expression in unstimulated cells as well as polyI:C and IFN-β stimulated cells was lower in the presence of either WT or Y125A NS1. ISRE-5x and NF-kB-4x mediated reporter expression in the presence of either WT and Y125A NS1 was lower in TNF-α stimulated cells as well. PolyI:C and IFN-β are both innate immune signals directly involved in the response to acute viral infection[64,73,74,92,171–175], whereas TNF-α is directly relevant for NF-kB activation and inflammatory signaling[154,155]. IRF family member transcription factors are directly activated downstream of polyI:C and IFN-β stimulus[176]. However, NF-kB family member transcription factors are directly activated downstream of polyI:C or TNF-α, but not IFN-β, stimulus[154,155,173,177].

Genomic regulatory elements, which carry their current sequence and are conserved for their functionality rather than engineering, are those with which NS1 would interact in the course of viral infection. Testing NS1 bound regulatory elements within the *IFIT* locus with the luciferase reporter assay provided evidence that such regulatory elements drive altered gene expression in the presence of NS1. The genes regulated by these elements are interferon-stimulated genes

strongly induced during hRSV infection. As these elements drive lower reporter gene expression in the presence of NS1, they may similarly drive lower expression of the *IFIT* genes in hRSV infected cells. While the RNA-sequencing for NS1 Y125A hRSV infected A549s shows that the IFIT transcriptome is more like that of mock than of WT NS1 hRSV infection, this is in context of an infection in which the WT but not Y125A NS1 bearing virus continues to replicate at 96 hpi[128].

I next tested the effect of NS1 Y125A on transcription driven by the *IFIT3* promoter. I found that reporter expression driven by the *IFIT3* promoter in polyI:C stimulated cells was markedly less decreased in the presence of Y125A NS1 (about twice as much total expression vs WT, $p < 0.05$), whereas gene expression driven by the ISRE-5x promoter in polyI:C stimulated cells underwent a ten-fold decrease in the presence of both WT and Y125A NS1. While optimized promoter constructs drive a consistent decrease in reporter expression in the presence of NS1, they lack the complexity of genomic regulatory elements. This difference may explain why transcription driven by the optimized promoters was more profoundly reduced in the presence of NS1 than by the genomic regulatory elements.

Within the genome, variants of consensus motifs bind transcription factors with different affinities[103,166–168], which in turn modulates the level of regulation that any one transcription factor has on a specific gene. Discrete combinations of transcription factors such as IRF3, IRF7, and NF-kB have also been shown to modulate expression of individual genes[103,178–180]. At a gene such as *IFIT3*, the promoter element may bind IRF3 homodimers or IRF3/7 heterodimers activated downstream of RIG-I signaling. The *IFIT3* promoter also has a tripartite ISRE where an IRF3 trimer may bind. While the ISRE-5x promoter contains five consecutive GAAANNGAAA motifs and measures 75 bp, the IFIT3 promoter contains one bipartite and one tripartite ISRE motif and

measures 336 bp. Therefore, motifs not shared between the ISRE-5x and the IFIT3 promoters may account for the demonstrated difference in gene expression in the presence of NS1 WT vs. Y125A.

Changes in gene expression may also be ascribed to the signaling pathway that is activated. ISRE consensus-like motifs are bound by IRF3/7 downstream of RIG-I signaling, but by ISGF3 downstream of IFN-β signaling[71,97]. The ISRE-5x is more strongly induced by IFN-β than by polyI:C, and there is a stronger decrease in gene expression in the presence of NS1 downstream of polyI:C than downstream of IFN-β. IRF9 and IRF3 only share 26% sequence identity by protein BLAST[181]. NS1 is known to interact strongly with IRF3, but it is not known whether it strongly interacts with other family members such as IRF9.

I next asked whether promoter regions near DEGs outside of the IFIT locus continued the pattern of lower expression in the presence of NS1. I focused on the *ISG20* promoter for this next set of experiments, as in addition to an ISRE motif, the *ISG20* promoter harbors at least one instance of the NF-kB RE. In contrast to the *IFIT3* promoter, in the presence of NS1 the *ISG20* promoter drove increased gene expression. I hypothesized that additional regulatory elements within the *ISG20* promoter or a combination not present in the *IFIT3* promoter were responsible for the opposite effect on gene expression. As well as ISRE and NF-kB regulatory elements, the *ISG20* promoter contains several additional regulatory sequences that might explain the distinct transcriptional response to NS1. These include GAS, E.box, GATA, and GC elements[165]. Previous work by the Mechti group indicates that the *ISG20* expression may be induced by NF-kB and IRF1 in Daudi cells[164]. Additionally, GC boxes may bind SP1 family transcription factors. These GC boxes were discussed in Gongora et al[165] as potential motifs driving basal *ISG20* transcription. However, the same paper demonstrated that a truncated promoter lacking the GC box elements but including the ISRE and a 5' UTR overlapping region of the promoter still drove increased

expression vs that of a control vector[165]. Independent work by the Pentecost group identified *ISG20* as an estrogen modulated gene[182]. Serial truncations indicated that several motifs throughout the promoter region increase the level of transcription[165] from the basal rate. Therefore NF-kB and IRF family transcription factors alone fail to fully account for *ISG20* promoter driven transcription. The SP-1 transcription factor may also bind the GC box motifs[183] and Myc/Max heterodimers may bind the E. box motif[184]. SP-1 has both activating and suppressive functions dependent on context. In some tissues SP-1/NF-kB complexes lead to upregulation of gene expression[185,186]. Determining which transcription factors bind throughout the *ISG20* promoter downstream of innate immune signaling will be important to determine which are involved in the upregulation of gene expression in the presence of NS1.

Alternatively, the presence of both NF-kB RE and ISRE motifs within the same promoter could affect gene transcription differently than the presence of only one or the other. As the second NF-kB RE element that I annotated manually is within the 5' UTR of the *ISG20* promoter fragment, I suggest that it is involved in *ISG20* promoter-driven transcription. While both IRF3 and NF-kB may be present within the nucleus, NS1 may disrupt the interaction of IRF3 with regulatory elements due to its interaction with Mediator and chromatin. Where IRF3 and NF-kB might bind the same regulatory element in the absence of NS1, then in the presence of NS1 one or the other may be sterically excluded. This could disrupt the timing and duration of transcriptional upregulation. Furthermore, if a transcription factor such as SP-1, which can act as either inducer or repressor in a context dependent manner, were to remain bound to the promoter, then absence of an innate immune specific transcription factor required for the antiviral action of SP-1 (either inducing or repressing) might permit SP-1 to act as the opposite: repressing where the antiviral action would be inducing, or vice versa. Such a mechanism might involve the combined activity

of SP-1 with NF-kB[183,185,186], but the interaction of NS1 with NF-kB might abolish its association with SP-1.

The non-enzymatic IFIT proteins interact with several binding partners including eIF3[187] in complex with which they inhibit translation. IFIT3 specifically enhances the antiproliferative activity of cell cycle mediators p21 and p27[188,189], and enhances RIG-I signaling[190], among several functions[144–146]. Furthermore, the protein encoded by *ISG20* is an endonuclease preferentially active against single-stranded RNAs[162–164] and has known antiviral activity against influenza virus, a negative sense RNA virus[191] like RSV, as well as multiple positive sense RNA viruses[192]. Transcription is highly induced during hRSV infection. Taken together, these results indicate that *ISG20* is upregulated during acute viral infection as an antiviral defense measure. However, sustained expression of *ISG20* may increase the amount of cellular RNAs processed by ISG20. This could increase cell stress and raise the risk of long-term inflammation, a known effect of hRSV infection.

Chapter 5: Future directions and concluding remarks

5.1 DNA affinity precipitation assay and electrophoretic mobility shift assay may be used to confirm association of NS1 with host transcription factors at genomic TF binding motifs

A biochemical assay can be used to demonstrate an association of NS1 with host transcription factors on chromatin. Two methods commonly used in the literature are the DNA affinity precipitation assay (DAPA), previously utilized by our group to identify cancer associated single nucleotide polymorphisms that alter transcription factor binding affinity[193], and the electrophoretic mobility shift assay (EMSA)[194].

In the DAPA assay, I transfected 293Ts with an HA-tagged IRF3-5D expressing plasmid, then incubated with oligonucleotides containing an interferon stimulated regulatory element (ISRE) or scrambled oligo (**Figure 5.1**). By overexpressing IRF3 and/or NS1 and incubating with oligonucleotides encoding the ISRE monomer or a scrambled sequence, this assay can be used to test whether addition of NS1 reduces the binding capacity of IRF3 to the ISRE.

While the IRF3-5D construct was used in the experiments illustrated below, IRF3 is not the only transcription factor that should be considered. The well characterized phosphomimetic, IRF3-5D[36,37,86,87] can be used in experiments requiring activated IRF3 without viral infection or the addition of stimulus. However, multiple phosphomimetic forms of IRF3 are available including 6D, which has an additional aspartic acid change[166]. In Daudi cells, the IRF1 transcription factor has been shown to associate via EMSA with the ISRE motif within the *ISG20* promoter[165], and so this combination of transcription factor and oligonucleotide should also be tested in A549s with NS1. As NS1 has been shown to suppress or otherwise alter reporter transcription downstream of IFN-βstimulus, IRF9, the IRF component of the ISGF3 complex, is an additional viable candidate.

NF-kB and SP-1 should also be tested as they are candidates for interaction with NS1 as discussed in Chapter 4.

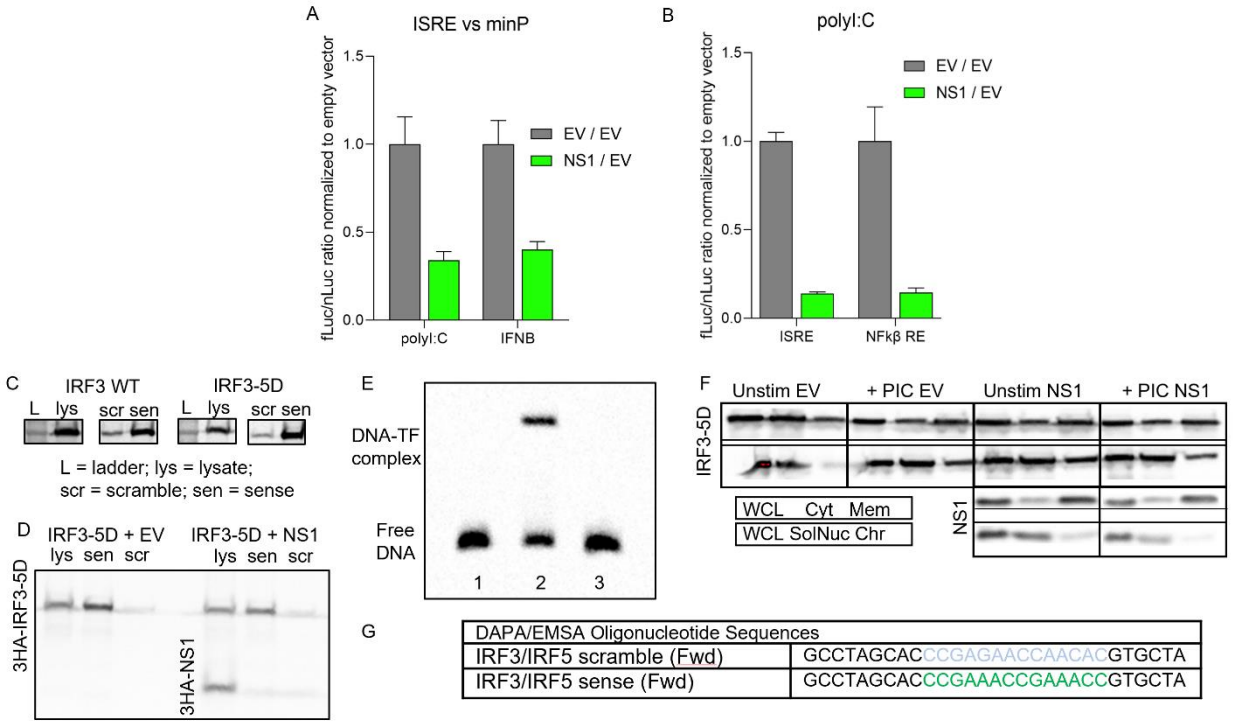


Figure 5.1. ISRE motif oligonucleotides bind IRF3-5D. IRF3 WT, IRF3-D, and NS1 were all 3HA tagged. A, B, (Luciferase) 293Ts were stimulated with polyI:C or IFNB (A), incubated for 24 hours with a firefly luciferase reporter plasmid with either a minimal promoter, ISRE promoter, or NF-kB (B) regulatory element promoter and either empty vector (EV) or an NS1 expressing plasmid. The firefly luciferase (fLuc) ratio over nano luciferase (nLuc, transfection control) was calculated for both EV and NS1 conditions and normalized to EV for that promoter. C, (DAPA) Cell lysate transfected with plasmid expressing HA-tagged IRF3 WT (left) or IRF3-5D (right) was incubated with biotinylated oligonucleotides (sequences shown panel G) and detected with anti-HA antibody. D, (DAPA) Cell lysate transfected with IRF3-5D + either empty vector or an NS1 expressing plasmid was incubated and detected as above. E, (EMSA) Control DNA and oligonucleotides included with the LightShift Chemiluminescent EMSA kit were subjected to EMSA according to the manufacturer's directions. F, (Cell fractionation) Cells transfected with IRF3-5D and either empty vector or NS1, with or without polyI:C stimulus, were fractionated according to the Thermo Scientific Subcellular Protein Fractionation Kit for Cultured Cells. G, (Oligonucleotide sequences) Forward oligonucleotide sequences used in DAPA and EMSA experiments. Portion unique to the scrambled oligo is shown in grey and the corresponding portion unique to the IRF3-IRF5 binding sequence is shown in green. F, (EMSA) Cell lysate transfected with IRF3 + either empty vector or NS1 was subjected to EMSA with the oligonucleotides shown in panel G. *Figure adapted from June 2021 thesis progress report.*

5.2 NS1 a3 helix variants and other hRSV proteins should be subjected to chromatin profiling

Several residues within the α3 helix of NS1 may modulate of the transcriptional regulatory effect of the protein. NS1 variants should be immunoprecipitated with chromatin through chromatin immunoprecipitation, or else by a similar technique such as CUT&RUN[195]. CUT&RUN offers advantages over ChIP in that there is lower inherent background and that lower cell counts may yield sufficient DNA for identification of peaks.

5.3 Studies may be expanded to additional cell types

hRSV is a respiratory virus, thus physiologically relevant tissues throughout the airway should be tested for NS1 chromatin occupancy. Primary human tracheobronchial epithelial cells (hTECs) can be infected with hRSV. With the novel anti-NS1 antibody as a tool, infected hTECs could be subjected to CUT&RUN-sequencing to profile chromatin bound NS1.

Concluding remarks

The studies detailed here demonstrate a novel, chromatin associated role for NS1, an interferon antagonist protein previously shown to act only in the cytoplasm. First, I demonstrated that NS1 associates with several hundred regulatory elements throughout the genome. Furthermore, I showed these regulatory elements are enriched at genes that are differentially expressed during hRSV infection and involved in the innate immune response. At least a subset of these regulatory elements drive altered reporter gene expression in the presence of NS1. This transcriptional regulatory role of NS1 provides a more complete picture of the extent to which hRSV disrupts of host immune response. Together, these studies provide an exciting new avenue through which to explore the biological basis for severe or recurring hRSV infection.

Works cited

1. Nair, H. *et al.* Global burden of acute lower respiratory infections due to respiratory syncytial virus in young children: a systematic review and meta-analysis. *Lancet Lond. Engl.* **375**, 1545–1555 (2010).
2. Scheltema, N. M. *et al.* Global respiratory syncytial virus-associated mortality in young children (RSV GOLD): a retrospective case series. *Lancet Glob. Health* **5**, e984–e991 (2017).
3. Stockman, L. J., Curns, A. T., Anderson, L. J. & Fischer-langley, G. Respiratory Syncytial Virus-associated Hospitalizations Among Infants and Young Children in the United States, 1997–2006. *Pediatr. Infect. Dis. J.* **31**, 5–9 (2012).
4. Kim, L. *et al.* Identifying Gaps in Respiratory Syncytial Virus Disease Epidemiology in the United States Prior to the Introduction of Vaccines. *Clin. Infect. Dis.* **65**, 1020–1025 (2017).
5. Scott, P. D. *et al.* Molecular Analysis of Respiratory Syncytial Virus Reinfections in Infants from Coastal Kenya. *J. Infect. Dis.* **193**, 59–67 (2006).
6. Glezen, W. P., Taber, L. H., Frank, A. L. & Kasel, J. A. Risk of Primary Infection and Reinfection With Respiratory Syncytial Virus. *Am. J. Dis. Child.* **140**, 543–546 (1986).
7. Farrag, M. A. & Almajhdi, F. N. Human Respiratory Syncytial Virus: Role of Innate Immunity in Clearance and Disease Progression. *Viral Immunol.* **29**, 11–26 (2016).
8. Kikkert, M. Innate Immune Evasion by Human Respiratory RNA Viruses. *J. Innate Immun.* **12**, 4–20 (2020).
9. Hall, C. B. *et al.* The burden of respiratory syncytial virus infection in young children. *N. Engl. J. Med.* **360**, 588–598 (2009).
10. Griffin, M. P. *et al.* Single-Dose Nirsevimab for Prevention of RSV in Preterm Infants. *N. Engl. J. Med.* **383**, 415–425 (2020).

11. Mohammed, M. H. A., Agouba, R., Obaidy, I. E., Alhabshan, F. & Abu-Sulaiman, R. Palivizumab prophylaxis against respiratory syncytial virus infection in patients younger than 2 years of age with congenital heart disease. *Ann. Saudi Med.* **41**, 31–35 (2021).
12. Englund, J. A. & Chu, H. Y. Vaccines Against Respiratory Syncytial Virus: The Time Has Come. *J. Infect. Dis.* **215**, 4–7 (2017).
13. Chu, H. Y. & Englund, J. A. Maternal immunization: Maternal Immunization. *Birth Defects Res.* **109**, 379–386 (2017).
14. Madhi, S. A. *et al.* Respiratory Syncytial Virus Vaccination during Pregnancy and Effects in Infants. *N. Engl. J. Med.* **383**, 426–439 (2020).
15. Mascola, J. R. & Fauci, A. S. Novel vaccine technologies for the 21st century. *Nat. Rev. Immunol.* **20**, 87–88 (2020).
16. GlaxoSmithKline. *A Phase III, Randomized, Double-blind, Placebo-controlled Multi-country Study to Demonstrate Efficacy of a Single Dose of Unadjuvanted RSV Maternal Vaccine, Administered IM to Pregnant Women 18 to 49 Years of Age, for Prevention of RSV Associated LRTIs in Their Infants up to 6 Months of Age*. https://clinicaltrials.gov/ct2/show/NCT04605159 (2021).
17. National Institute of Allergy and Infectious Diseases (NIAID). *VRC 317: A Phase I Randomized, Open-Label Clinical Trial to Evaluate Dose, Safety, Tolerability and Immunogenicity of a Stabilized Prefusion RSV F Subunit Protein Vaccine, VRC-RSVRGP084-00-VP (DS-Cav1), Alone or With Alum Adjuvant, in Healthy Adults*. https://clinicaltrials.gov/ct2/show/NCT03049488 (2020).
18. Ruckwardt, T. J. *et al.* Safety, tolerability, and immunogenicity of the respiratory syncytial virus prefusion F subunit vaccine DS-Cav1: a phase 1, randomised, open-label, dose-escalation clinical trial. *Lancet Respir. Med.* (2021) doi:10.1016/S2213-2600(21)00098-9.
19. Amarasinghe, G. K. *et al.* Taxonomy of the order Mononegavirales: update 2017. *Arch. Virol.* **162**, 2493–2504 (2017).

20. Canedo-Marroquín, G. *et al.* Modulation of Host Immunity by Human Respiratory Syncytial Virus Virulence Factors: A Synergic Inhibition of Both Innate and Adaptive Immunity. *Front. Cell. Infect. Microbiol.* **7**, (2017).
21. Cao, D., Gao, Y. & Liang, B. Structural Insights into the Respiratory Syncytial Virus RNA Synthesis Complexes. *Viruses* **13**, 834 (2021).
22. Kipper, S. *et al.* New Host Factors Important for Respiratory Syncytial Virus (RSV) Replication Revealed by a Novel Microfluidics Screen for Interactors of Matrix (M) Protein. *Mol. Cell. Proteomics* **14**, 532–543 (2015).
23. Collins, P. L. & Melero, J. A. Progress in understanding and controlling respiratory syncytial virus: Still crazy after all these years. *Virus Res.* **162**, 80–99 (2011).
24. Bermingham, A. & Collins, P. L. The M2-2 protein of human respiratory syncytial virus is a regulatory factor involved in the balance between RNA replication and transcription. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 11259–11264 (1999).
25. Jin, H. *et al.* Recombinant respiratory syncytial viruses with deletions in the NS1, NS2, SH, and M2-2 genes are attenuated in vitro and in vivo. *Virology* **273**, 210–218 (2000).
26. Atreya, P. L., Peeples, M. E. & Collins, P. L. The NS1 Protein of Human Respiratory Syncytial Virus Is a Potent Inhibitor of Minigenome Transcription and RNA Replication. *J. Virol.* **72**, 1452–1461 (1998).
27. Evans, J. E., Cane, P. A. & Pringle, C. R. Expression and characterisation of the NS1 and NS2 proteins of respiratory syncytial virus. *Virus Res.* **43**, 155–161 (1996).
28. Spann, K. M., Tran, K. C. & Collins, P. L. Effects of Nonstructural Proteins NS1 and NS2 of Human Respiratory Syncytial Virus on Interferon Regulatory Factor 3, NF-□B, and Proinflammatory Cytokines. *J. Virol.* **79**, 5353–5362 (2005).
29. Swedan, S., Musiyenko, A. & Barik, S. Respiratory Syncytial Virus Nonstructural Proteins Decrease Levels of Multiple Members of the Cellular Interferon Pathways. *J. Virol.* **83**, 9682–9693 (2009).

30. Chatterjee, S. *et al.* Structural basis for human respiratory syncytial virus NS1-mediated modulation of host responses. *Nat. Microbiol.* **2**, 17101 (2017).

31. Amarasinghe, G. K. *et al.* Taxonomy of the order Mononegavirales: update 2019. *Arch. Virol.* **164**, 1967–1980 (2019).

32. Spann, K. M., Tran, K.-C., Chi, B., Rabin, R. L. & Collins, P. L. Suppression of the Induction of Alpha, Beta, and Gamma Interferons by the NS1 and NS2 Proteins of Human Respiratory Syncytial Virus in Human Epithelial Cells and Macrophages. *J. Virol.* **78**, 4363–4369 (2004).

33. Bitko, V. *et al.* Nonstructural Proteins of Respiratory Syncytial Virus Suppress Premature Apoptosis by an NF-□B-Dependent, Interferon-Independent Mechanism and Facilitate Virus Growth. *J. Virol.* **81**, 1786–1795 (2007).

34. Elliott, J. *et al.* Respiratory Syncytial Virus NS1 Protein Degrades STAT2 by Using the Elongin-Cullin E3 Ligase. *J. Virol.* **81**, 3428–3436 (2007).

35. Ren, J. *et al.* A novel mechanism for the inhibition of interferon regulatory factor-3-dependent gene expression by human respiratory syncytial virus NS1 protein. *J. Gen. Virol.* **92**, 2153–2159 (2011).

36. Hiscott, J. *et al.* Triggering the Interferon Response: The Role of IRF-3 Transcription Factor. *J. Interferon Cytokine Res.* **19**, 1–13 (1999).

37. Lin, R., Heylbroeck, C., Pitha, P. M. & Hiscott, J. Virus-Dependent Phosphorylation of the IRF-3 Transcription Factor Regulates Nuclear Translocation, Transactivation Potential, and Proteasome-Mediated Degradation. *Mol. Cell. Biol.* **18**, 2986–2996 (1998).

38. Wu, W. *et al.* The Interactome of the Human Respiratory Syncytial Virus NS1 Protein Highlights Multiple Effects on Host Cell Biology. *J. Virol.* **86**, 7777–7789 (2012).

39. Pei, J. *et al.* Nuclear-localized human respiratory syncytial virus NS1 protein modulates host gene transcription. *Cell Rep.* **37**, 109803 (2021).

40. Carlsten, J. O. P., Zhu, X. & Gustafsson, C. M. The multitalented Mediator complex. *Trends Biochem. Sci.* **38**, 531–537 (2013).

41. Soutourina, J. Mammalian Mediator as a Functional Link between Enhancers and Promoters. *Cell* **178**, 1036–1038 (2019).

42. André, K. M., Sipos, E. H. & Soutourina, J. Mediator Roles Going Beyond Transcription. *Trends Genet. TIG* (2020) doi:10.1016/j.tig.2020.08.015.

43. Soutourina, J. Transcription regulation by the Mediator complex. *Nat. Rev. Mol. Cell Biol.* **19**, 262–274 (2018).

44. Abdella, R. *et al.* Structure of the human Mediator-bound transcription preinitiation complex. *Science* (2021) doi:10.1126/science.abg3074.

45. Kagey, M. H. *et al.* Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**, 430–5 (2010).

46. Whyte, W. A. *et al.* Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013).

47. Pei, J. *et al.* Structural basis for IFN antagonism by human respiratory syncytial virus nonstructural protein 2. *Proc. Natl. Acad. Sci. U. S. A.* **118**, (2021).

48. Ling, Z., Tran, K. C. & Teng, M. N. Human Respiratory Syncytial Virus Nonstructural Protein NS2 Antagonizes the Activation of Beta Interferon Transcription by Interacting with RIG-I. *J. Virol.* **83**, 3734–3742 (2009).

49. Lo, M. S., Brazas, R. M. & Holtzman, M. J. Respiratory Syncytial Virus Nonstructural Proteins NS1 and NS2 Mediate Inhibition of Stat2 Expression and Alpha/Beta Interferon Responsiveness. *J. Virol.* **79**, 9315–9319 (2005).

50. Förster, A., Maertens, G. N., Farrell, P. J. & Bajorek, M. Dimerization of Matrix Protein Is Required for Budding of Respiratory Syncytial Virus. *J. Virol.* **89**, 4624–4635 (2015).

51. Ghildyal, R., Ho, A. & Jans, D. A. Central role of the respiratory syncytial virus matrix protein in infection. *FEMS Microbiol. Rev.* **30**, 692–705 (2006).

52. Ghildyal, R., Baulch-Brown, C., Mills, J. & Meanger, J. The matrix protein of Human respiratory syncytial virus localises to the nucleus of infected cells and inhibits transcription. *Arch. Virol.* **148**, 1419–1429 (2003).

53. Ghildyal, R. *et al.* Nuclear Import of the Respiratory Syncytial Virus Matrix Protein Is Mediated By Importin β1 Independent of Importin α. *Biochemistry* **44**, 12887–12895 (2005).

54. Jorquera, P. A. *et al.* Verdinexor (KPT-335), a Selective Inhibitor of Nuclear Export, Reduces Respiratory Syncytial Virus Replication In Vitro. *J. Virol.* **93**, e01684-18 (2019).

55. Andrews, J., El-Alawi, M. & Payton, J. Genotify: Fast, lightweight gene lookup and summarization. *Journal of Open Source Software* https://joss.theoj.org (2018) doi:10.21105/joss.00885.

56. The Respiratory Syncytial Virus Matrix Protein Possesses a Crm1-Mediated Nuclear Export Mechanism | Journal of Virology. https://jvi-asm-org.beckerproxy.wustl.edu/content/83/11/5353.long.

57. Bharaj, P. *et al.* The Matrix Protein of Nipah Virus Targets the E3-Ubiquitin Ligase TRIM6 to Inhibit the IKKε Kinase-Mediated Type-I IFN Antiviral Response. *PLoS Pathog.* **12**, e1005880 (2016).

58. Evidence for Ubiquitin-Regulated Nuclear and Subnuclear Trafficking among Paramyxovirinae Matrix Proteins. https://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1004739.

59. Wang, Y. E. *et al.* Ubiquitin-Regulated Nuclear-Cytoplasmic Trafficking of the Nipah Virus Matrix Protein Is Important for Viral Budding. *PLOS Pathog.* **6**, e1001186 (2010).

60. Coleman, N. A. & Peeples, M. E. The matrix protein of Newcastle disease virus localizes to the nucleus via a bipartite nuclear localization signal. *Virology* **195**, 596–607 (1993).

61. Duan, Z. *et al.* Characterization of signal sequences determining the nuclear export of Newcastle disease virus matrix protein. *Arch. Virol.* **158**, 2589–2595 (2013).

62. McLinton, E. C. *et al.* Nuclear localization and secretion competence are conserved among henipavirus matrix proteins. *J. Gen. Virol.* **98**, 563–576 (2017).

63. Yu, X., Shahriari, S., Li, H.-M. & Ghildyal, R. Measles Virus Matrix Protein Inhibits Host Cell Transcription. *PLOS ONE* **11**, e0161360 (2016).

64. Owen, J. A., Punt, J., Stranford, S. A., Jones, P. P. & Kuby, J. *Kuby immunology*. (W.H. Freeman, 2013).

65. Castranova, V., Rabovsky, J., Tucker, J. H. & Miles, P. R. The alveolar type II epithelial cell: a multifunctional pneumocyte. *Toxicol. Appl. Pharmacol.* **93**, 472–483 (1988).

66. Tristram, D. A., Hicks, W. & Hard, R. Respiratory Syncytial Virus and Human Bronchial Epithelium. *Arch. Otolaryngol. Neck Surg.* **124**, 777–783 (1998).

67. Medzhitov, R. & Janeway, C. A. Decoding the Patterns of Self and Nonself by the Innate Immune System. *Science* **296**, 298–300 (2002).

68. Alexopoulou, L., Holt, A. C., Medzhitov, R. & Flavell, R. A. Recognition of double-stranded RNA and activation of NF-kappaB by Toll-like receptor 3. *Nature* **413**, 732–738 (2001).

69. Levy, D. E. & García-Sastre, A. The virus battles: IFN induction of the antiviral state and mechanisms of viral evasion. *Cytokine Growth Factor Rev.* **12**, 143–156 (2001).

70. Schmid, S., Mordstein, M., Kochs, G., García-Sastre, A. & tenOever, B. R. Transcription Factor Redundancy Ensures Induction of the Antiviral State. *J. Biol. Chem.* **285**, 42013–42022 (2010).

71. Basler, C. F. & García-Sastre, A. Viruses and the type I interferon antiviral system: induction and evasion. *Int. Rev. Immunol.* **21**, 305–337 (2002).

72. Schoggins, J. W. *et al.* A diverse range of gene products are effectors of the type I interferon antiviral response. *Nature* **472**, 481–485 (2011).

73. Schoggins, J. W. Interferon-Stimulated Genes: What Do They All Do? *Annu. Rev. Virol.* **6**, 567–584 (2019).

74. Hijano, D. R. *et al.* Role of Type I Interferon (IFN) in the Respiratory Syncytial Virus (RSV) Immune Response and Disease Severity. *Front. Immunol.* **10**, 566 (2019).

75. Ban, J. *et al.* Human Respiratory Syncytial Virus NS 1 Targets TRIM25 to Suppress RIG-I Ubiquitination and Subsequent RIG-I-Mediated Antiviral Signaling. *Viruses* **10**, (2018).

76. Hornung, V. *et al.* 5'-Triphosphate RNA is the ligand for RIG-I. *Science* **314**, 994–7 (2006).

77. Yu, M. & Levine, S. J. Toll-like receptor 3, RIG-I-like receptors and the NLRP3 inflammasome: Key modulators of innate immune responses to double-stranded RNA viruses. *Cytokine Growth Factor Rev.* **22**, 63–72 (2011).

78. Janeway, C. A. & Medzhitov, R. Innate immune recognition. *Annu. Rev. Immunol.* **20**, 197–216 (2002).

79. Schleimer, R. P., Kato, A., Kern, R., Kuperman, D. & Avila, P. C. Epithelium: At the interface of innate and adaptive immune responses. *J. Allergy Clin. Immunol.* **120**, 1279–1284 (2007).

80. Lau, C. M. *et al.* Epigenetic control of innate and adaptive immune memory. *Nat. Immunol.* **19**, 963–972 (2018).

81. Hosakote, Y. M., Brasier, A. R., Casola, A., Garofalo, R. P. & Kurosky, A. Respiratory Syncytial Virus Infection Triggers Epithelial HMGB1 Release as a Damage-Associated Molecular Pattern Promoting a Monocytic Inflammatory Response. (2016) doi:10.1128/JVI.01279-16.

82. Schlee, M. *et al.* Recognition of 5' triphosphate by RIG-I helicase requires short blunt double-stranded RNA as contained in panhandle of negative-strand virus. *Immunity* **31**, 25–34 (2009).

83. Reich, S. *et al.* Structural insight into cap-snatching and RNA synthesis by influenza polymerase. *Nature* **516**, 361–366 (2014).

84. Engel, D. A. The influenza virus NS1 protein as a therapeutic target. *Antiviral Res.* **99**, 409–416 (2013).

85. Leung, D. W., Prins, K. C., Basler, C. F. & Amarasinghe, G. K. Ebolavirus VP35 is a multifunctional virulence factor. *Virulence* **1**, 526–531 (2010).

86. Hiscott, J. Triggering the Innate Antiviral Response through IRF-3 Activation. *J. Biol. Chem.* **282**, 15325–15329 (2007).

87. Yoneyama, M., Suhara, W. & Fujita, T. Control of IRF-3 activation by phosphorylation. *J. Interferon Cytokine Res. Off. J. Int. Soc. Interferon Cytokine Res.* **22**, 73–76 (2002).

88. Onomoto, K., Onoguchi, K. & Yoneyama, M. Regulation of RIG-I-like receptor-mediated signaling: interaction between host and viral factors. *Cell. Mol. Immunol.* (2021) doi:10.1038/s41423-020-00602-7.

89. Andersen, J., VanScoy, S., Cheng, T.-F., Gomez, D. & Reich, N. C. IRF-3-dependent and augmented target genes during viral infection. *Genes Immun.* **9**, 168–175 (2008).

90. Jing, T. *et al.* The Structural Basis of IRF-3 Activation upon Phosphorylation. *J. Immunol. Baltim. Md 1950* **205**, 1886–1896 (2020).

91. Grandvaux, N. *et al.* Transcriptional Profiling of Interferon Regulatory Factor 3 Target Genes: Direct Involvement in the Regulation of Interferon-Stimulated Genes. *J. Virol.* **76**, 5532–5539 (2002).

92. Zhou, Y. *et al.* TLR3 activation efficiency by high or low molecular mass poly I:C. *Innate Immun.* **19**, 184–192 (2013).

93. Groskreutz, D. J. *et al.* Respiratory Syncytial Virus Induces TLR3 Protein and Protein Kinase R, Leading to Increased Double-Stranded RNA Responsiveness in Airway Epithelial Cells. *J. Immunol.* **176**, 1733–1740 (2006).

94. Plociennikowska, A. *et al.* TLR3 activation by Zika virus stimulates inflammatory cytokine production which dampens the antiviral response induced by RIG-I-like receptors. *J. Virol.* (2021) doi:10.1128/JVI.01050-20.

95. Karpala, A. J., Lowenthal, J. W. & Bean, A. G. Activation of the TLR3 pathway regulates IFNβ production in chickens. *Dev. Comp. Immunol.* **32**, 435–444 (2008).

96. Barik, S. Respiratory Syncytial Virus Mechanisms to Interfere with Type 1 Interferons. in 173–191 (Springer, Berlin, Heidelberg, 2013). doi:10.1007/978-3-642-38919-1_9.

97. Chen, K., Liu, J. & Cao, X. Regulation of type I interferon signaling in immunity and inflammation: A comprehensive review. *J. Autoimmun.* **83**, 1–11 (2017).

98. Gibbert, K., Schlaak, J., Yang, D. & Dittmer, U. IFN-α subtypes: distinct biological activities in anti-viral therapy. *Br. J. Pharmacol.* **168**, 1048–1058 (2013).

99. Borden, E. C. *et al.* Interferons at age 50: past, current and future impact on biomedicine. *Nat. Rev. Drug Discov.* **6**, 975–990 (2007).

100. Mostafavi, S. *et al.* Parsing the Interferon Transcriptional Network and Its Disease Associations. *Cell* **164**, 564–578 (2016).

101. Carnero, E. *et al.* Type I interferon regulates the expression of long non-coding RNAs. *Front. Immunol.* **5**, 1–14 (2014).

102. Petermann, F. *et al.* The Magnitude of IFN-γ Responses Is Fine-Tuned by DNA Architecture and the Non-coding Transcript of Ifng-as1. *Mol. Cell* **75**, 1229-1242.e5 (2019).

103. Freaney, J. E., Kim, R., Mandhana, R. & Horvath, C. M. Extensive cooperation of immune master regulators IRF3 and NFκB in RNA Pol II recruitment and pause release in human innate antiviral transcription. *Cell Rep.* **4**, 959–973 (2013).

104. Czerkies, M. *et al.* Cell fate in antiviral response arises in the crosstalk of IRF, NF-κB and JAK/STAT pathways. *Nat. Commun.* **9**, (2018).

105. Johnson, B. *et al.* Dimerization Controls Marburg Virus VP24-dependent Modulation of Host Antioxidative Stress Responses. *J. Mol. Biol.* **428**, 3483–3494 (2016).

106. Andrews, J. M. *et al.* Novel cell adhesion/migration pathways are predictive markers of HDAC inhibitor resistance in cutaneous T cell lymphoma. *EBioMedicine* (2019) doi:10.1016/j.ebiom.2019.07.053.

107. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).

108. Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).

109. Carroll, T. S., Liang, Z., Salama, R., Stark, R. & de Santiago, I. Impact of artifact removal on ChIP quality metrics in ChIP-seq and ChIP-exo data. *Front. Genet.* **5**, 75 (2014).

110. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS J. Integr. Biol.* **16**, 284–287 (2012).

111. Yu, G., Wang, L.-G. & He, Q.-Y. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382–2383 (2015).

112. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinforma. Oxf. Engl.* **29**, 15–21 (2013).

113. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).

114. Wang, L., Wang, S. & Li, W. RSeQC: quality control of RNA-seq experiments. *Bioinforma. Oxf. Engl.* **28**, 2184–2185 (2012).

115. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinforma. Oxf. Engl.* **26**, 139–140 (2010).

116. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).

117. Luo, W., Friedman, M. S., Shedden, K., Hankenson, K. D. & Woolf, P. J. GAGE: generally applicable gene set enrichment for pathway analysis. *BMC Bioinformatics* **10**, 161 (2009).

118. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).

119. Horani, A., Nath, A., Wasserman, M. G., Huang, T. & Brody, S. L. Rho-associated protein kinase inhibition enhances airway epithelial Basal-cell proliferation and lentivirus transduction. *Am. J. Respir. Cell Mol. Biol.* **49**, 341–347 (2013).

120. You, Y., Richer, E. J., Huang, T. & Brody, S. L. Growth and differentiation of mouse tracheal epithelial cells: selection of a proliferative population. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **283**, L1315-1321 (2002).

121. Blanchard, E. L. *et al.* Polymerase-tagged respiratory syncytial virus reveals a dynamic rearrangement of the ribonucleocapsid complex during infection. *PLoS Pathog.* **16**, e1008987 (2020).

122. Lifland, A. W. *et al.* Human respiratory syncytial virus nucleoprotein and inclusion bodies antagonize the innate immune response mediated by MDA5 and MAVS. *J. Virol.* **86**, 8245–8258 (2012).

123. García, J., García-Barreno, B., Vivo, A. & Melero, J. A. Cytoplasmic inclusions of respiratory syncytial virus-infected cells: formation of inclusion bodies in transfected cells that coexpress the nucleoprotein, the phosphoprotein, and the 22K protein. *Virology* **195**, 243–247 (1993).

124. Fournier, M. *et al.* FOXA and master transcription factors recruit Mediator and Cohesin to the core transcriptional regulatory circuitry of cancer cells. (2016) doi:10.1038/srep34962.

125. Gao, T. *et al.* EnhancerAtlas: a resource for enhancer annotation and analysis in 105 human cell/tissue types. *Bioinformatics* btw495 (2016) doi:10.1093/bioinformatics/btw495.

126. GRCh37 - hg19 - Genome - Assembly - NCBI. https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.13/.

127. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, (2014).

128. Chatterjee, S. *et al.* Structural basis for human respiratory syncytial virus NS1-mediated modulation of host responses. (2017) doi:10.1038/nmicrobiol.2017.101.

129. Bakre, A. *et al.* Human respiratory syncytial virus non-structural protein NS1 modifies miR-24 expression via transforming growth factor-β. *J. Gen. Virol.* **96**, 3179–3191 (2015).
130. Dorsett, D. & Merkenschlager, M. Cohesin at active genes: a unifying theme for cohesin and gene expression from model organisms to humans. *Curr. Opin. Cell Biol.* **25**, 327–333 (2013).
131. Davis, M. A. *et al.* The SCF-Fbw7 ubiquitin ligase degrades MED13 and MED13L and regulates CDK8 module association with Mediator. *Genes Dev.* **27**, 151–156 (2013).
132. Fant, C. B. & Taatjes, D. J. Regulatory functions of the Mediator kinases CDK8 and CDK19. *Transcription* **10**, 76–90 (2019).
133. Guo, Z., Wang, G., Lv, Y., Wan, Y. Y. & Zheng, J. Inhibition of Cdk8/Cdk19 Activity Promotes Treg Cell Differentiation and Suppresses Autoimmune Diseases. *Front. Immunol.* **10**, 1988 (2019).
134. Knuesel, M. T., Meyer, K. D., Bernecky, C. & Taatjes, D. J. The human CDK8 subcomplex is a molecular switch that controls Mediator coactivator function. *Genes Dev.* **23**, 439–451 (2009).
135. Charest, J. *et al.* Combinatorial Action of Temporally Segregated Transcription Factors. *Dev. Cell* **55**, 483-499.e7 (2020).
136. Dubois-Chevalier, J. *et al.* Organizing combinatorial transcription factor recruitment at cis-regulatory modules. *Transcription* **9**, 233–239 (2018).
137. Smale, S. T. Core promoters: active contributors to combinatorial gene regulation. *Genes Dev* **15**, 2503–8. (2001).
138. Smale, S. T. *Selective Transcription in Response to an Inflammatory Stimulus*. vol. 140 (2010).
139. Alashkar Alhamwe, B., Miethe, S., Pogge von Strandmann, E., Potaczek, D. P. & Garn, H. Epigenetic Regulation of Airway Epithelium Immune Functions in Asthma. *Front. Immunol.* **11**, 1747 (2020).

140. Moheimani, F. *et al.* The genetic and epigenetic landscapes of the epithelium in asthma. *Respir. Res.* **17**, 119 (2016).

141. Zhao, Z. & Shilatifard, A. Epigenetic modifications of histones in cancer. *Genome Biol.* **20**, 245 (2019).

142. Bradner, J. E., Hnisz, D. & Young, R. A. Transcriptional Addiction in Cancer. *Cell* **168**, 629–643 (2017).

143. Chen, Y., Xu, L., Lin, R. Y.-T., Müschen, M. & Koeffler, H. P. Core transcriptional regulatory circuitries in cancer. *Oncogene* **39**, 6633–6646 (2020).

144. Diamond, M. S. & Farzan, M. The broad-spectrum antiviral functions of IFIT and IFITM proteins. *Nat. Rev. Immunol.* **13**, 46–57 (2013).

145. Fensterl, V. & Sen, G. C. The ISG56/IFIT1 Gene Family. *J. Interferon Cytokine Res.* **31**, 71–78 (2011).

146. Fensterl, V. & Sen, G. C. Interferon-Induced Ifit Proteins: Their Role in Viral Pathogenesis. *J. Virol.* **89**, 2462–2468 (2015).

147. Wang, W., Xu, L., Su, J., Peppelenbosch, M. P. & Pan, Q. Transcriptional Regulation of Antiviral Interferon-Stimulated Genes. (2017) doi:10.1016/j.tim.2017.01.001.

148. Wang, J. *et al.* Factorbook.org: a Wiki-based database for transcription factor-binding data generated by the ENCODE consortium. *Nucleic Acids Res.* **41**, D171-176 (2013).

149. Ourthiague, D. R. *et al.* Limited specificity of IRF3 and ISGF3 in the transcriptional innate-immune response to double-stranded RNA. *J. Leukoc. Biol.* **98**, 119–128 (2015).

150. Margueron, R. & Reinberg, D. The Polycomb complex PRC2 and its mark in life. *Nat. Rev.* 3–9 (2011) doi:10.1038/nature09784.

151. pGL4.45[luc2P/ISRE/Hygro] Vector Protocol. https://www.promega.com/resources/protocols/product-information-sheets/a/pgl4-45-vector-protocol/.

152. Uccellini, M. B. & García-Sastre, A. ISRE-Reporter Mouse Reveals High Basal and Induced Type I IFN Responses in Inflammatory Monocytes. *Cell Rep.* **25**, 2784-2796.e3 (2018).

153. Reporter vector PGL4.45[luc2P/ISRE/Hygro], complete sequence. (2012).

154. Lawrence, T. The Nuclear Factor NF-kB Pathway in Inflammation.

155. Pahl, H. L. Activators and target genes of Rel/NF-kB transcription factors. 14.

156. Ji, Z., He, L., Regev, A. & Struhl, K. Inflammatory regulatory network mediated by the joint action of NF-kB, STAT3, and AP-1 factors is involved in many human cancers. *Proc. Natl. Acad. Sci.* **116**, 9453–9462 (2019).

157. NF-kB Target Genes » NF-kB Transcription Factors | Boston University. http://www.bu.edu/nf-kb/gene-resources/target-genes/.

158. Bahrami, S. & Drabløs, F. Gene regulation in the immediate-early response process. *Adv. Biol. Regul.* **62**, 37–49 (2016).

159. Honda, K. *et al.* Spatiotemporal regulation of MyD88-IRF-7 signalling for robust type-I interferon induction. *Nature* **434**, 1035–1040 (2005).

160. Li, M. *et al.* Dynamic regulation of transcription factors by nucleosome remodeling. *eLife* **4**, e06249 (2015).

161. Degols, G., Eldin, P. & Mechti, N. ISG20, an actor of the innate immune response. *Biochimie* **89**, 831–835 (2007).

162. Espert, L. Interferon-induced exonuclease ISG20 exhibits an antiviral activity against human immunodeficiency virus type 1. *J. Gen. Virol.* **86**, 2221–2229 (2005).

163. Espert, L. *et al.* ISG20, a New Interferon-induced RNase Specific for Single-stranded RNA, Defines an Alternative Antiviral Pathway against RNA Genomic Viruses. *J. Biol. Chem.* **278**, 16151–16158 (2003).

164. Espert, L. *et al.* The exonuclease ISG20 is directly induced by synthetic dsRNA via NF-kappaB and IRF1 activation. *Oncogene* **23**, 4636–4640 (2004).

165. Gongora, C., Degols, G., Espert, L., Hua, T. D. & Mechti, N. A unique ISRE, in the TATA-less human Isg20 promoter, confers IRF-1-mediated responsiveness to both interferon type I and type II. *Nucleic Acids Res.* **28**, 2333–2341 (2000).

166. Andrilenas, K. K. *et al.* DNA-binding landscape of IRF3, IRF5 and IRF7 dimers: implications for dimer-specific gene regulation. *Nucleic Acids Res.* **46**, 2509–2520 (2018).

167. Csumita, M. *et al.* Specific enhancer selection by IRF3, IRF5 and IRF9 is determined by ISRE half-sites, 5' and 3' flanking bases, collaborating transcription factors and the chromatin environment in a combinatorial fashion. *Nucleic Acids Res.* **48**, 589–604 (2020).

168. Dalskov, L. *et al.* Characterization of distinct molecular interactions responsible for IRF3 and IRF7 phosphorylation and subsequent dimerization. *Nucleic Acids Res.* **48**, 11421–11433 (2020).

169. Canino, C. *et al.* A STAT3-NFkB/DDIT3/CEBPβ axis modulates ALDH1A3 expression in chemoresistant cell subpopulations. *Oncotarget* **6**, 12637–12653 (2015).

170. Verkoczy, L. *et al.* A role for nuclear factor kappa B/rel transcription factors in the regulation of the recombinase activator genes. *Immunity* **22**, 519–531 (2005).

171. Li, L.-C. *et al.* Small dsRNAs induce transcriptional activation in human cells. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 17337–42 (2006).

172. Palchetti, S. *et al.* Transfected poly(I:C) activates different dsRNA receptors, leading to apoptosis or immunoadjuvant response in androgen-independent prostate cancer cells. *J. Biol. Chem.* **290**, 5470–5483 (2015).

173. Peisley, A. & Hur, S. Multi-level regulation of cellular recognition of viral dsRNA. doi:10.1007/s00018-012-1149-4.

174. Wu, B. *et al.* Structural Basis for dsRNA Recognition, Filament Formation, and Antiviral Signal Activation by MDA5. *Cell* **152**, 276–289 (2013).

175. Panne, D., Maniatis, T. & Harrison, S. C. An atomic model of the interferon-beta enhanceosome. *Cell* **129**, 1111–1123 (2007).

176. Ashley, C. L., Abendroth, A., McSharry, B. P. & Slobedman, B. Interferon-Independent Upregulation of Interferon-Stimulated Genes during Human Cytomegalovirus Infection is Dependent on IRF3 Expression. *Viruses* **11**, (2019).

177. Carty, M., Guy, C. & Bowie, A. G. Detection of viral infections by innate immunity. *Biochem. Pharmacol.* 114316 (2020) doi:10.1016/j.bcp.2020.114316.

178. Hernandez-Munain, C., Roberts, J. L. & Krangel, M. S. Cooperation among multiple transcription factors is required for access to minimal T-cell receptor alpha-enhancer chromatin in vivo. *Mol Cell Biol* **18**, 3223–3233 (1998).

179. Narlikar, G. J., Fan, H. Y. & Kingston, R. E. Cooperation between complexes that regulate chromatin structure and transcription. *Cell* **108**, 475–87. (2002).

180. Yang, H., Lin, C. H., Ma, G., Baffi, M. O. & Wathelet, M. G. Interferon Regulatory Factor-7 Synergizes with Other Transcription Factors through Multiple Interactions with p300/CBP Coactivators. *J. Biol. Chem.* **278**, 15495–15504 (2003).

181. Zhang, Z., Schwartz, S., Wagner, L. & Miller, W. A greedy algorithm for aligning DNA sequences. *J. Comput. Biol. J. Comput. Mol. Cell Biol.* **7**, 203–214 (2000).

182. Pentecost, B. T. Expression and estrogen regulation of the HEM45 MRNA in human tumor lines and in the rat uterus. *J. Steroid Biochem. Mol. Biol.* **64**, 25–33 (1998).

183. SP1 Sp1 transcription factor [Homo sapiens (human)] - Gene - NCBI. https://www.ncbi.nlm.nih.gov/gene/6667.

184. MYC MYC proto-oncogene, bHLH transcription factor [Homo sapiens (human)] - Gene - NCBI. https://www.ncbi.nlm.nih.gov/gene/4609.

185. Liu, S. *et al.* Sp1/NFkappaB/HDAC/miR-29b regulatory network in KIT-driven myeloid leukemia. *Cancer Cell* **17**, 333–347 (2010).

186. O'Connor, L., Gilmour, J. & Bonifer, C. The Role of the Ubiquitously Expressed Transcription Factor Sp1 in Tissue-specific Transcriptional Regulation and in Disease. *Yale J. Biol. Med.* **89**, 513–525 (2016).

187. Terenzi, F., Pal, S. & Sen, G. C. Induction and mode of action of the viral stress-inducible murine proteins, P56 and P54. *Virology* **340**, 116–124 (2005).

188. Li, D. *et al.* Rig-G is a growth inhibitory factor of lung cancer cells that suppresses STAT3 and NF-κB. *Oncotarget* **7**, 66032–66050 (2016).

189. Xiao, S. *et al.* RIG-G as a key mediator of the antiproliferative activity of interferon-related pathways through enhancing p21 and p27 proteins. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 16448–16453 (2006).

190. Liu, X.-Y., Chen, W., Wei, B., Shan, Y.-F. & Wang, C. IFN-induced TPR protein IFIT3 potentiates antiviral signaling by bridging MAVS and TBK1. *J. Immunol. Baltim. Md 1950* **187**, 2559–2568 (2011).

191. Dawson, W. K., Lazniewski, M. & Plewczynski, D. RNA structure interactions and ribonucleoprotein processes of the influenza A virus. *Brief. Funct. Genomics* **17**, 402–414 (2017).

192. Zhou, Z. *et al.* Antiviral activities of ISG20 in positive-strand RNA virus infections. *Virology* **409**, 175–188 (2011).

193. Koues, O. I. *et al.* Enhancer Sequence Variants and Transcription-Factor Deregulation Synergize to Construct Pathogenic Regulatory Circuits in B-Cell Lymphoma. *Immunity* **42**, 186–198 (2015).

194. Miller, D. E. *et al.* Screening for Functional Non-coding Genetic Variants Using Electrophoretic Mobility Shift Assay (EMSA) and DNA-affinity Precipitation Assay (DAPA). *J. Vis. Exp. JoVE* (2016) doi:10.3791/54093.

195. Hainer, S. J. & Fazzio, T. G. High-Resolution Chromatin Profiling Using CUT&RUN. *Curr. Protoc. Mol. Biol.* **126**, e85 (2019).