

NORTHWESTERN UNIVERSITY

Regulators of the Cellular Response to Viruses

A DISSERTATION

**SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS**

for the degree

DOCTOR OF PHILOSOPHY

Field of Interdisciplinary Biological Sciences

By

Roli Mandhana Kasat

EVANSTON, ILLINOIS

December 2018

ABSTRACT

Regulators of the Cellular Response to Viruses

Roli Mandhana Kasat

The cellular innate immune response to viruses is a defense mechanism executed by most cells in the human body to form the initial barrier to virus replication. Detection of viral nucleic acids initiates widespread gene expression changes that combine to establish an antiviral state and stimulate professional immune cell activation. This system co-evolved with viruses that can mutate to develop new evasion and escape strategies, forcing the innate immune system to adapt and form novel antiviral mechanisms as well. A deeper understanding of the complex cellular response to viruses is needed to detect and treat established and emerging viral infections as well as immune and autoimmune disorders. This level of understanding requires not only studying the components of the cellular innate immune response but also studying the viral factors that enable evasion and antagonism of the system. Harnessing these cellular and viral features could improve the development of diagnostics, antiviral drugs, and therapeutics for immune and autoimmune diseases. The work presented in this thesis furthers the knowledge about both cellular and viral components related to the innate immune response. Virus-regulated gene expression changes assayed by RNA-sequencing revealed the presence of thousands of previously-unannotated RNAs as well as other previously unrecognized virus-regulated RNAs. These primarily noncoding RNAs are candidates for regulatory or direct antiviral function and could be potential targets in pharmaceutical development for

antivirals. The use of viral inhibitors of the innate immune response in drug development is an under-explored area. A proof of this concept is demonstrated by experiments using a family of Paramyxovirus IFN evasion proteins to suppress excessive immune signaling due to laboratory-generated and patient-derived mutations that hyperactivate the RNA sensor, MDA5. The results presented in this thesis not only advance the current understanding of cellular innate immunity and elucidate the expanding roles of noncoding RNAs in antiviral responses, but also establish new directions for the development of diagnostics and therapeutics.

ACKNOWLEDGEMENTS

First and foremost, I must thank my advisor, Dr. Curt Horvath, for his guidance throughout graduate school. His door has always been open to me, providing constant support as well as constructive criticism when needed. He not only invested time in the development of my critical thinking and experimental skills but also helped me become a more well-rounded scientist. His passion and dedication to scientific inquiry are inspiring, and I will miss being around that excitement and insight.

My graduate school experience would not have been the same without the members of the Horvath lab. The birthday celebrations and holiday parties were fun milestones to share with them, but I'll most fondly look back on the lively sports, politics and pop culture discussions that we had. I was very fortunate to have incredible labmates who not only made working a fun experience, but were great resources when I needed a helping hand or scientific advice. I'm especially indebted to Dr. Jonathan Freaney whose ChIP-sequencing data motivated my project.

I'm also very grateful to have had Patrick Parisien and Nancy Au Yeung in the lab with me for the entirety of my graduate career. It has been incredible to know that Patrick is sitting right behind me when I need someone to share in my excitement or to comfort me. I will miss sharing our lives with each other every day. From my very first day in the lab, Nancy has been someone who I have turned to for help with troubleshooting experiments

or finding new directions for my projects. More recently she has been an amazing confidante as we navigate the latter stages of graduate school together.

When I joined IBiS I assumed I would be friendly with my classmates, but I did not realize at the time how close I would become to this group. I was very lucky to be surrounded by amazing scientists who supported one another as we went through graduate school together. I'll always cherish our Tuesday trivia nights at Nevin's! I'm especially grateful to have found some of my best friends in my classmates Adam Hockenberry and Kari Tanaka. I have grown so much thanks to their friendship, both scientifically and personally.

I could go on for many pages thanking my friends outside of IBiS and Northwestern for everything they have done for me over the years. They have been instrumental in keeping me grounded and motivated, reminding me that no person is one-dimensional.

Finally, I must thank my entire family for their unending love, support and encouragement throughout graduate school.

There are no words to express how grateful I am to my parents, Om & Manju Mandhana, and my brother and sister-in-law, Abhishek & Aditi Mandhana. Since I was young, they have always encouraged and supported me in doing my best. Because of their sacrifices, hard work and dedication, I could pursue every opportunity and experience that I have

ever wanted. Without their patient and loving support as I completed my education, I would not be here. I also want to thank my newest family, the Kasats, who have been amazingly warm and welcoming from the first day I met them. They have been right alongside my family cheering me on through these final stages.

And last but not least, I would like to thank my husband Sarish Kasat. He has been the ultimate partner as I navigated graduate school. He has been there with me celebrating every small victory and consoling me when I was down. I would not have made it this far without his unconditional support and love. With him in my corner, I know I can achieve whatever I desire next in my career.

LIST OF ABBREVIATIONS

AGS	Aicardi-Goutieres Syndrome	MDA5	Melanoma Differentiation-Associated protein 5
ATP	Adenosine triphosphate	MeV	Measles virus
bp	Base pair	MOI	Multiplicity of infection
CA-MDA5	Constitutively active MDA5	mRNA	Messenger RNA
CARD	Caspase activation and recruitment domain	MuV	Mumps virus
cDNA	Complementary DNA	NF-κB	Nuclear factor κ B
cGAS	Cyclic-GMP-AMP synthase	NiV	Nipah virus
ChIP	Chromatin immunoprecipitation	PAMP	Pathogen-associated molecular pattern
DNA	Deoxyribonucleic acid	PCR	Polymerase chain reaction
dsRNA	Double-stranded RNA	pfu	Plaque forming unit
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase	PIV5	Parainfluenza virus 5
GO	Gene ontology	poly(I:C)	Polyinosinic-polycytidylic acid
HeV	Hendra virus	PRR	Pattern recognition receptor
hg19	Human genome build 19	RIG-I	Retinoic acid-inducible gene 1
HSV-1	Herpes simplex virus 1	RLR	RIG-I-like receptor
HTS	High-throughput screen	RNA	Ribonucleic acid
IFN	Interferon	rRNA	Ribosomal RNA
IRF	Interferon regulatory factor	SeV	Sendai virus
ISG	Interferon stimulated gene	shRNA	Short hairpin RNA
ISGF3	Interferon stimulated gene factor 3	siRNA	Small interfering RNA
KEGG	Kyoto Encyclopedia of Genes and Genomes	SLE	Systemic lupus erythematosus
LGP2	Laboratory of genetics and physiology 2	TLR	Toll-like receptor
lncRNA	Long noncoding RNA	VSV	Vesicular stomatitis virus
MAVS	Mitochondrial antiviral signaling protein	WT	Wild type

TABLE OF CONTENTS

ABSTRACT	2
ACKNOWLEDGEMENTS	4
LIST OF ABBREVIATIONS	7
LIST OF FIGURES	11
LIST OF TABLES	13
CHAPTER I: INTRODUCTION	14
OVERVIEW OF CELLULAR INNATE IMMUNE RESPONSE	15
BACKGROUND	18
Induction of the cellular innate immune response	18
DETECTION OF NON-SELF NUCLEIC ACIDS	19
INNATE IMMUNE TRANSCRIPTION FACTORS: IRF3 AND NF- κ B.....	22
<i>Nuclear Factor κB (NF-κB)</i>	22
<i>Interferon Regulatory Factor (IRF)</i>	23
TYPE I INTERFERON	24
<i>Type I IFN production</i>	25
<i>Type I IFN Signaling</i>	26
Regulation of the cellular innate immune response	26
CELL-ENCODED REGULATORS.....	28
VIRUS-ENCODED REGULATORS.....	31
CONTRIBUTIONS OF THESIS RESEARCH	32
CHAPTER II: VIRUS INFECTION REGULATES EXPRESSION OF NOVEL RNAS ...	33
INTRODUCTION	34
Cellular innate immune response to viruses	34
Genome-wide sequencing increases detection of RNAs	35
Contributions of thesis research	36
RESULTS	37

RNA-sequencing and identification of differentially expressed RNAs	37
Gene enrichment analysis	43
VIRUS-REGULATED PATHWAYS THAT PROMOTE VIRUS INFECTION: CHANGES	
IN METABOLIC GENE EXPRESSION.....	49
Genomic distribution of annotated and unannotated RNAs	52
Protein coding analysis of annotated and unannotated RNAs.....	56
Vertebrate conservation analysis of virus-regulated annotated and unannotated	
RNAs	58
Transcription regulation of virus-induced RNAs	61
Inducibility of virus-induced unannotated RNAs by diverse stimuli.....	62
Identification of negative regulators of the cellular innate immune response.....	74
DISCUSSION	85
CHAPTER III: INHIBITION OF INTERFERONOPATHIC MDA5 BY PARAMYXOVIRUS	
V PROTEINS.....	93
INTRODUCTION	94
Interferonopathies caused by monogenic mutations	94
MDA5 and its role in chronic interferon production	97
Paramyxovirus evasion of antiviral signaling	98
RESULTS.....	100
Construction and Characterization of Constitutively Active (CA)-MDA5 Proteins ..	100
V Proteins Bind to CA-MDA5 Proteins	104
V Proteins Suppress CA-MDA5 Signal Transduction	105
PIV5 V Protein Suppresses CA-MDA5 Mediated IFNβ Gene Expression	105
PIV5 V Protein Suppresses Interferonopathic Gene Expression.....	108
PIV5 V Protein Suppresses Hyperactive Antiviral Response.....	108
DISCUSSION	111
CHAPTER IV: PERSPECTIVES.....	116
ENDOGENOUS REGULATORS.....	119
VIRUS-ENCODED REGULATORS	122
CONCLUSIONS	124
REFERENCES.....	126

APPENDIX A: MATERIALS AND METHODS.....	145
GENERAL METHODS	146
METHODS SPECIFIC TO CHAPTER II	148
METHODS SPECIFIC TO CHAPTER III	152
METHODS SPECIFIC TO APPENDIX D.....	154
APPENDIX B: PRIMERS USED FOR GENE EXPRESSION ANALYSIS.....	155
APPENDIX C: BIOLOGICAL REPLICATES OF AFF1-AS1 AND ZBED5-AS1 KNOCKDOWN EXPERIMENTS	158
APPENDIX D: ASSAY DEVELOPMENT FOR IDENTIFICATION OF SMALL MOLECULE REGULATORS OF CELLULAR INNATE IMMUNE SIGNALING.....	162
INTRODUCTION	163
DEVELOPMENT OF HIGH-THROUGHPUT SCREEN	164
Cell line and detection reagents	164
Determining assay conditions	164
Determining assay quality	166
HIGH-THROUGHPUT SCREEN TO IDENTIFY COMPOUNDS OF INTEREST	169
Titration analysis of compounds of interest.....	170
Future directions	170
APPENDIX E: ENRICHED GO BIOLOGICAL PROCESSES FOR VIRUS-INDUCED GENES	182

LIST OF FIGURES

Figure 1.1 – Co-evolution of host and virus.....	17
Figure 1.2 – Induction of the cellular innate immune response by viral nucleic acids	21
Figure 1.3 – Type I IFN signaling.....	27
Figure 1.4 – Various regulators of the cellular response to viruses	29
Figure 2.1 – Quantile distribution of average reads of differentially expressed RNAs	39
Figure 2.2 – Validation of differential gene expression induced by Sendai virus infection	40
Figure 2.3 – Graphical representation of functional enrichment analysis of Sendai virus- regulated RNAs.....	45
Figure 2.4 – Metabolic gene expression changes following Sendai virus infection.....	51
Figure 2.5 – Comparison of expression of Sendai virus-regulated previously-unannotated and previously-annotated RNAs	53
Figure 2.6 – Comparison of genomic distribution of Sendai virus-regulated previously- annotated and previously-unannotated RNAs	55
Figure 2.7 – Comparison of protein coding potential of Sendai virus-regulated previously- annotated and previously-unannotated RNAs.	57
Figure 2.8 – Comparison of vertebrate sequence conservation of Sendai virus-regulated annotated and unannotated RNAs.....	60
Figure 2.9 – Transcription factor occupancy at Sendai virus-induced loci.....	63
Figure 2.10 – Classification of virus-induced previously-unannotated RNA expression in Namalwa cells from virus infections and IFN treatment.....	68
Figure 2.11 – Expression of virus-induced previously-unannotated RNAs in 2fTGH cells from various stimuli	72
Figure 2.12 - Expression of virus-induced previously-unannotated RNAs in THP-1 cells from various stimuli	73

Figure 2.13 - Validation of AFF1-AS1 and ZBED5-AS1 expression induced by Sendai virus infection.....	81
Figure 2.14 – Expression of AFF1-AS1 and ZBED5-AS1 in Namalwa cells infected with diverse viruses	82
Figure 2.15 – Expression of AFF1-AS1 and ZBED5-AS1 in 2fTGH cells by diverse stimuli	84
Figure 2.16 – AFF1-AS1 knockdown by shRNA in 2fTGH cells	86
Figure 2.17 – ZBED5-AS1 knockdown by siRNA in 2fTGH cells.....	87
Figure 3.1 – Cellular machinery implicated in type I interferonopathies	96
Figure 3.2 – Construction and characterization of mutant MDA5 proteins	102
Figure 3.3 – CA-MDA5 mutants increase endogenous IFNβ mRNA expression.....	103
Figure 3.4 – V Proteins bind CA-MDA5 mutants.....	106
Figure 3.5 – V proteins suppress CA-MDA5 mutant signaling.....	107
Figure 3.6 – PIV5 V protein suppresses CA-MDA5-induced IFNβ expression	109
Figure 3.7 – PIV5 V protein suppresses CA-MDA5-induced ISG expression	110
Figure 3.8 – PIV5 V protein suppresses antiviral response of CA-MDA5 mutants	112
Figure C.1 – Biological replicates of AFF1-AS1 knockdown by shRNA in 2fTGH cells ..	159
Figure C.2 - Biological replicates of AFF1-AS1 knockdown by shRNA in 2fTGH cells ...	160
Figure C.3 – Biological replicates of ZBED5-AS1 knockdown by siRNA in 2fTGH cells.	161
Figure D.1 – Illustration of A549-Dual cell reporter induction.....	165
Figure D.2 – Preliminary experiment to determine assay conditions	167
Figure D.3 – Determining HTS assay quality	168
Figure D.4 – Screen of Spectrum Collection	172

LIST OF TABLES

Table 2.1 – The 10 most highly virus-induced previously-annotated RNAs.....	42
Table 2.2 – The 10 most highly virus-suppressed previously-annotated RNAs	42
Table 2.3 – Clusters of enriched GO biological processes for virus-suppressed genes ..	46
Table 2.4 – Factor occupancy in virus-induced previously-annotated RNAs	64
Table 2.5 – Factor occupancy in virus-induced previously-unannotated RNAs	66
Table 2.6 – Details of virus-induced previously-unannotated RNA expression in Namalwa cells from virus infections and IFN α treatment	69
Table 2.7 – Summary of virus-induced previously-unannotated RNA expression in multiple cell lines cells from virus infections, polyI:C and IFN α treatment.....	75
Table 2.8 – Details of virus-induced previously-unannotated RNA expression in multiple cell lines cells from virus infections, polyI:C and IFN α treatment.....	76
Table B.1 – Primers used for gene expression analysis	156
Table D.1 – Results of compounds of interest	173
Table D.2 – Dose-response analysis of compounds of interest	174
Table E.1 – Clusters of enriched GO biological processes terms for virus-induced genes	183

CHAPTER I: INTRODUCTION

OVERVIEW OF CELLULAR INNATE IMMUNE RESPONSE

In humans the initial response to viruses and other microbial pathogens is executed by the innate immune system, an evolutionarily conserved defense mechanism (Medzhitov and Janeway, 1997). The innate immune response occurs rapidly within a few hours of pathogen detection, in contrast to the more specialized adaptive immune response which takes several days to initiate (Barber, 2001). Though specialized immune cells such as dendritic cells, macrophages, and natural killer cells participate in the innate immune response, non-professional cells such as epithelial cells, endothelial cells and fibroblasts can initiate cell-intrinsic innate immune signaling (Akira et al., 2006). This cellular innate immune response to viruses is carried out by many different cells, and thus forms the first barrier to virus replication (de Weerd and Nguyen, 2012; Mogensen, 2009; Stark et al., 1998).

The cell-intrinsic innate immune response is activated by a small number of pattern-recognition receptors (PRRs) expressed in many different cell types that can identify a variety of pathogens by their conserved pathogen-associated molecular patterns (PAMPs) (Medzhitov and Janeway, 1997). Recognition of a PAMP initiates innate immune signaling in the host cell and leads to the production of cytokines and hundreds of other proinflammatory molecules, which not only form the first line of defense but also alert the adaptive immune system for further response to the pathogen (Fearon and Locksley, 1996).

While the cellular innate immune response is rapid and robust, in a healthy, normal cell the response is only transient (Coccia et al., 2006; Larner et al., 1986; Whittemore and Maniatis, 1990). Innate immune effectors are potent molecules that signal to limit viral infection by various mechanisms including inhibiting transcription and translation and initiating apoptosis (McNab et al., 2015; Stark et al., 1998). Excessive or inappropriate expression of these effectors is poisonous to the cell and can cause a variety of immune and autoimmune disorders (Baechler et al., 2003; Crow and Wohlgemuth, 2003; Crow et al., 2015; Gresser et al., 1980; Rice et al., 2013). As a result, the expression of these molecules attenuates rapidly through various regulatory mechanisms, including transcriptional control and proteasomal degradation, to maintain a healthy immune response without causing harmful diseases (Kobayashi and Flavell, 2004).

This complex cellular innate immune system has evolved under selective pressure from viruses (Daugherty and Malik, 2012; Kosiol et al., 2008). When a virus mutates to successfully escape the host immune system, it creates pressure on the host to adapt and evolve a new strategy to limit the virus. However, just as the host faces selective pressure from the virus, the virus is also constantly under selective pressure from the host, creating an arms race between the two (Fig. 1.1).

Achieving a deeper understanding of the cell-intrinsic innate immune response requires not only identifying and characterizing all the cell-encoded regulators but also evaluating how exogenous factors are able to evade the system. The research presented in this

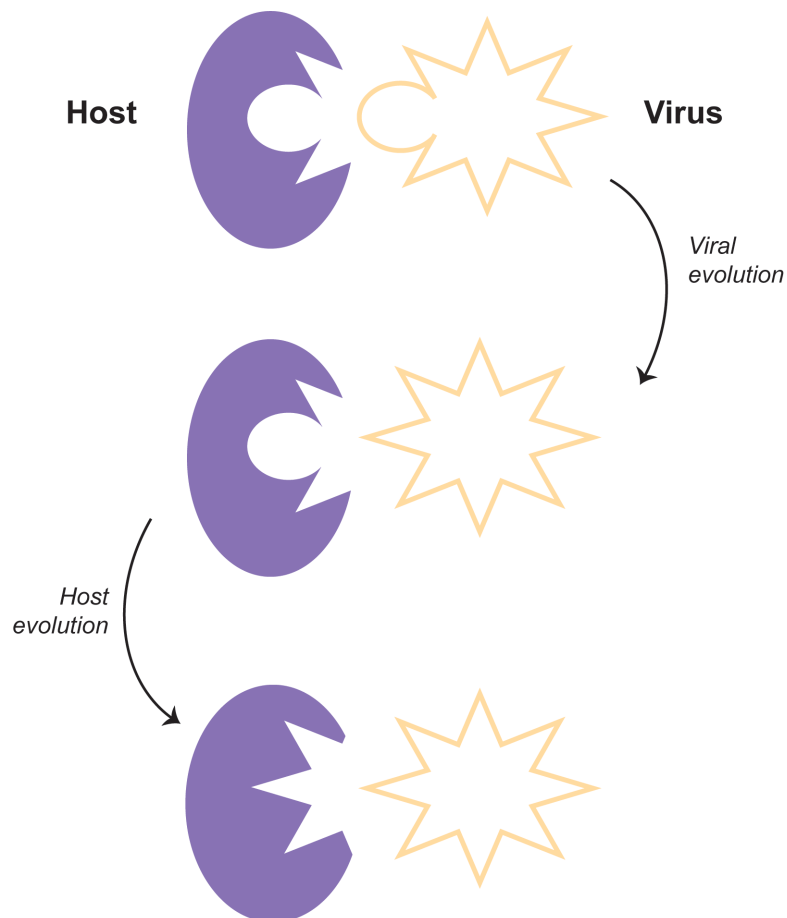


Figure 1.1 – Co-evolution of host and virus

An illustration of the selective pressure viruses and hosts exert on one another, driving their evolution over time. When a virus mutates to escape host recognition, this forces the host to adapt and evolve to recognize the new virus. The evolution of the host then pressures the virus to mutate once again to escape host recognition. This back and forth between the two creates an arms race that has shaped the human innate immune system and modern viruses.

thesis aimed to further the knowledge about the regulation of the cellular response to viruses from both of these angles. Due to the complexity of innate immune signaling and ChIP-sequencing analysis of virus-induced transcriptional regulation (Freaney et al., 2013), we hypothesized that there are many previously unrecognized cellular molecules with regulatory roles that could be identified through new deep-sequencing approaches. In addition to these unrecognized cellular features, there are also exogenous factors that could be novel targets for the development of antiviral and immune therapies as well as diagnostics. Chapter 2 analyzes the virus-regulated transcriptome to identify previously-unrecognized RNAs with potential functions in the cellular response to viruses. Chapter 3 provides a proof of concept for harnessing virus-encoded IFN evasion molecules to suppress excessive innate immune signaling. In a parallel experimental system detailed in Appendix D, a high-throughput screen was established to identify novel small molecules that could be used to control aberrant innate immune signaling. Together, the results in these chapters expand the current understanding about the regulators of the cellular response to viruses.

BACKGROUND

Induction of the cellular innate immune response

The cellular innate immune response to viruses is induced when host PRRs detect one of several viral components (Kawai and Akira, 2006a; Medzhitov and Janeway, 1997). A virion, the infectious virus particle, is comprised of enclosed viral genomic material, either single- or double-stranded DNA or RNA. Some viruses are enclosed in a protein shell

known as the capsid, while others, such as the Paramyxoviridae family of RNA genome viruses, have a phospholipid envelope containing viral glycoproteins. The genomic material as well as glycoproteins serve as PAMPs that trigger signaling downstream of host PRRs. However, the main PAMP for viruses is their nucleic acid.

DETECTION OF NON-SELF NUCLEIC ACIDS

Cellular PRRs are responsible for distinguishing self versus non-self nucleic acids to prevent improper triggering of the innate immune response (Janeway, 1992). The specific recognition of non-self nucleic acids relies on several characteristics that differ between self and non-self nucleic acids, including cellular localization, structure and chemical modifications (Schlee and Hartmann, 2016).

For non-self DNA, the hallmark is its localization in the cytoplasm. Since DNA is usually absent from the cytoplasm, its detection by cytoplasmic sensors indicates the presence of 'dangerous' DNA. For non-self RNAs, the major characteristic is the presence of long dsRNAs as these are a key component of viral replication and usually absent from uninfected cells (Weber et al., 2006). Additionally, chemical modifications of both RNAs and DNAs can also distinguish endogenous nucleic acids from foreign nucleic acids. While cellular processing removes 5' phosphates from self RNAs, many viral RNAs are marked by 5' di- or tri-phosphates (Bruns and Horvath, 2014). Endogenous RNAs are instead capped at the 5' end by an N7-methylated guanosine and further modified at the +1 ribonucleotide by 2'-O-methylation (Zust et al., 2011), providing another method of

distinguishing them from viral RNAs. Bacterial DNAs frequently exhibit unmethylated CpG dinucleotides that distinguish them from host DNAs. In addition to these specific motifs that physically differentiate the self and non-self nucleic acids, improper triggering of PRRs by accumulated self nucleic acids at steady state is prevented by endogenous nucleases that quickly degrade the cellular nucleic acids. All these features contribute to the selective activation of PRRs based on detection of non-self nucleic acids.

Distinct host PRRs detect non-self nucleic acids based on the identity and localization of the ligand. As shown in Figure 1.2, extracellular nucleic acids are detected by Toll-like receptors (TLRs), while intracellular nucleic acids are detected by the cyclic-GMP-AMP synthase (cGAS) and RIG-I-like receptors (RLRs).

The TLRs are transmembrane proteins expressed predominantly in immune cells that detect various extracellular PAMPs (Kawai and Akira, 2006b). DNA is detected by TLR9 and RNA is detected by TLR3, TLR7, and TLR8 (Kawasaki and Kawai, 2014). The extracellular ligand binding domain of the TLRs detects PAMPs and the TIR domain initiates intracellular signaling by interacting with one of various cytoplasmic adapters.

The main cytoplasmic PRRs are cGAS and the RLRs. cGAS is a cytoplasmic DNA sensor that produces cGAMP in a DNA-dependent manner (Ablasser et al., 2013; Diner et al., 2013; Gao et al., 2013; Sun et al., 2013; Wu et al., 2013b). Binding of dsDNA by cGAS forms a dimeric complex that synthesizes 2'3'-cGAMP, a ligand for ER-membrane

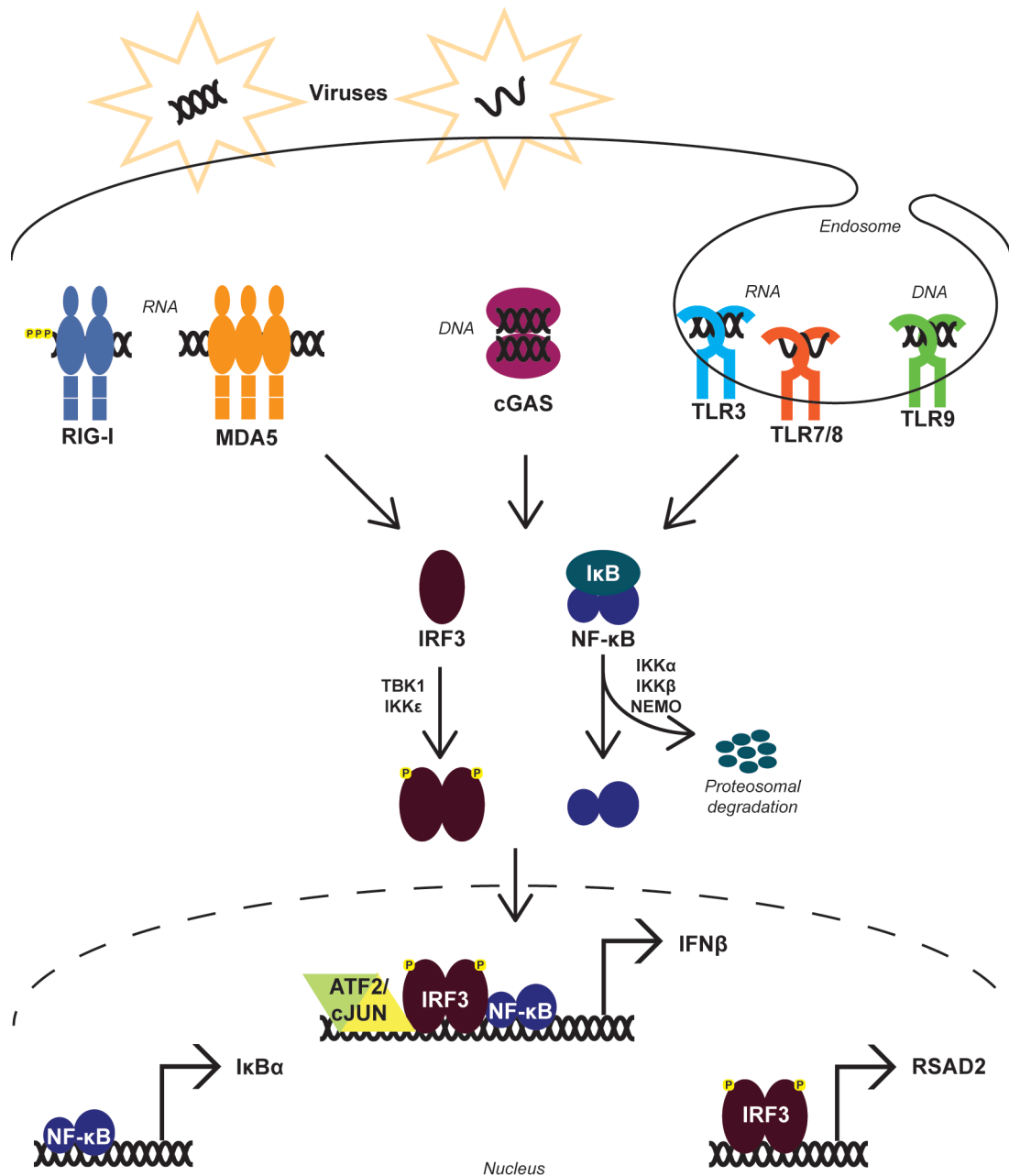


Figure 1.2 – Induction of the cellular innate immune response by viral nucleic acids

An illustration of the various host pattern recognition receptor (PRR) pathways that detect viral nucleic acids. The RLR family members RIG-I and MDA5 detect cytoplasmic RNA and cGAS detects cytoplasmic DNA. The TLR family members detect extracellular RNA and DNA. All pathways converge at the activation of the transcription factors IRF3 and NF-κB, which translocate into the nucleus following activation. These transcription factors drive the transcription of known immune genes such as IκBα, IFNβ, and RSAD2.

signaling adapter STING (Ishikawa et al., 2009; Zhang et al., 2013). The RLR family members RIG-I, MDA5, and LGP2 are widely expressed cytoplasmic RNA sensors that share a conserved central DECH-Box helicase domain that binds RNA and hydrolyzes ATP (Bruns and Horvath, 2014; Cordin et al., 2006). While LGP2 is able to bind dsRNA, it lacks the caspase activation and recruitment domain (CARD) required for downstream signaling. MDA5 and RIG-I bind distinct RNAs and oligomerize to form filaments and signal through their CARDS. MDA5 binds longer dsRNA, while RIG-I binds short RNA with 5'-triphosphate ends (Baum and Garcia-Sastre, 2010; Bruns and Horvath, 2014). Upon oligomerization, MDA5 and RIG-I filaments signal through the MAVS receptor on the mitochondrion.

INNATE IMMUNE TRANSCRIPTION FACTORS: IRF3 AND NF- κ B

All of the distinct nucleic acid-sensing PRR pathways converge at the activation of the transcription factors IRF3 and NF- κ B. These transcription factors are the master regulators of the innate immune response, coordinately binding to hundreds of loci following virus infection (Freaney et al., 2013). They are essential for the production of many antiviral effectors, including the type I interferons (IFNs).

Nuclear Factor κ B (NF- κ B)

The nuclear factor κ B (NF- κ B) transcription factors are master regulators of inflammation, immunity and development (Baeuerle and Henkel, 1994; Oeckinghaus and Ghosh, 2009). There are 5 family members: RelA (p65), RelB, c-Rel, and p105 and p100, which are

processed into the mature p50 and p52, respectively. All five proteins share the Rel homology domain which is needed for dimerization, nuclear translocation and DNA binding. These proteins form several hetero- and homodimeric complexes that function as distinct transcription factors (Oeckinghaus and Ghosh, 2009). In unstimulated cells, the transcription factors are inactive and sequestered in the cytoplasm by I κ B proteins. Virus infection predominantly activates the p50/RelA (p65) complex. As shown in Figure 1.2, the PRR signaling pathways activate the IKK complex consisting of IKK α , IKK β , and NEMO (IKK γ), to induce the ubiquitination and degradation of I κ B α , which releases NF- κ B into the nucleus (Oeckinghaus and Ghosh, 2009). Upon nuclear translocation, NF- κ B binds response elements in target gene promoters to induce their transcriptional activation.

Interferon Regulatory Factor (IRF)

The interferon regulatory factor (IRF) family of transcription factors, which modulates antiviral defense, immune responses, cell growth and apoptosis, consists of 9 different proteins, IRF1 through IRF9 (Paun and Pitha, 2007; Tamura et al., 2008). All family members have a conserved N-terminal DNA-binding domain with a unique helix-turn-helix structure that binds a consensus DNA sequence (Tamura et al., 2008). The C-terminal region has a non-conserved IRF association domain (IAD) that enables hetero- and homodimer formations with other transcription factors and increases the specificity and diversity in DNA binding. The primary IRF transcription factor downstream of PRR signaling is IRF3, which is constitutively expressed in the cytoplasm of unstimulated cells

(Tamura et al., 2008). Phosphorylation of IRF3 by TBK1 and IKK ϵ induces dimerization and nuclear translocation, where it binds to IRF3 response elements in target gene promoters to induce their transcriptional activation (Fig. 1.2).

TYPE I INTERFERON

Though activation of the innate immune transcription factors by virus infection induces the expression of many genes, the most highly induced and well-studied genes are interferons (IFNs), key cytokines in the antiviral response. The IFNs were named for their ability to interfere with virus infection by establishing cellular protection and influencing downstream immune responses (Stark et al., 1998).

There are several types of IFN (Stark et al., 1998). Type I IFN includes the single IFN β , as well as a number IFN α subtypes that all bind to a common IFN α/β receptor (Stark et al., 1998); Type II IFN includes the single IFN γ that binds to the distinct IFN γ receptor (Stark et al., 1998); Type III IFN includes IFN λ members, also known as IL-28 and IL-29, that bind to a distinct IFN λ receptor (Ank et al., 2006).

Type I IFNs are especially important as their receptors are found on most cells where they can induce the transcription of hundreds of downstream effectors to establish an antiviral state (de Weerd and Nguyen, 2012; Schoggins et al., 2011). The antiviral effectors degrade host and viral RNAs, inhibit transcription and translation, induce growth arrest, activate apoptosis, and trigger autophagy pathways that together lead to virus

suppression and infection clearance (McNab et al., 2015; Stark et al., 1998). IFN is also an immune modulator, regulating the activity of dendritic cells, macrophages, B cells, T cells, and natural killer cells to counteract infections and stimulate lasting immunity (Kiefer et al., 2012; McNab et al., 2015; Tough, 2012; Tough et al., 1996; Trinchieri, 2010).

Type I IFN production

The type I IFNs are very closely related virus-inducible proteins, though there are differences in their transcriptional regulation. The IFN β promoter is one of the most well-studied inducible gene promoters and has been previously shown to be transcriptionally silent due to the presence of a TSS-obscuring nucleosome that is displaced following virus infection (Freaney et al., 2014; Goodbourn et al., 1985; Zinn et al., 1983). Virus-induced binding of the transcription factors IRF3, NF- κ B, and ATF-2/c-Jun to the positive regulatory domains (PRDs) (Honda and Taniguchi, 2006; Kim and Maniatis, 1997) of the IFN β promoter forms an enhanceosome complex that recruits additional transcriptional co-activators to mediate re-organization of this +1 nucleosome and enable transcription of the gene (Agalioti et al., 2000; Kim and Maniatis, 1997; Lomvardas and Thanos, 2001). Though work from the Horvath lab has shown that all the IFN α promoters also have a well-positioned +1 nucleosome that blocks access to the TSS (Freaney et al., 2013), their transcriptional mechanism is less well understood. The IFN α promoters differ from the IFN β promoter and only consist of PRD-like elements that bind IRFs (Honda and Taniguchi, 2006; Ryals et al., 1985), indicating that even though IFN α and IFN β are expressed following virus infection, they are differentially regulated.

Type I IFN Signaling

Once expressed, the IFNs are secreted from the cell and bind a common transmembrane IFN α/β receptor comprised of the IFNAR1 and IFNAR2 chains to activate their own signaling pathway (Fig. 1.3) (Novick et al., 1994; Uze et al., 1990). Ligand binding induces a series of janus kinase (JAK)-mediated tyrosine phosphorylation events that activate the latent cytoplasmic proteins STAT1 and STAT2 (Darnell et al., 1994; Imbrota et al., 1994; Shuai et al., 1993; Yan et al., 1996). Phosphorylated STAT1 and STAT2, along with the DNA binding subunit IRF9, form the heterotrimeric transcription factor interferon stimulated gene factor 3 (ISGF3) (Fu et al., 1990; Kessler et al., 1990). ISGF3 translocates to the nucleus and binds to interferon stimulated response element (ISRE) sequences present in interferon stimulated gene (ISG) promoters to induce their transcription (Levy et al., 1988; Reich et al., 1987). Through this pathway the type I IFNs rapidly induce the expression, in some cases more than 100-fold (Blomstrom et al., 1986; Friedman et al., 1984), of hundreds of antiviral effectors within a few hours (Colonno and Pang, 1982; Lerner et al., 1984). The products of ISGs such as PKR (Pindel and Sadler, 2011), which inhibits protein translation, RNase L (Chakrabarti et al., 2011), which cleaves single-stranded RNA, and MxA (Haller et al., 2015), which inhibits viral transcription, combine to establish the cellular antiviral state that prevents virus replication and stimulates professional immune cell activation (Fearon and Locksley, 1996).

Regulation of the cellular innate immune response

The cellular innate immune response rapidly induces many signaling pathways to enable

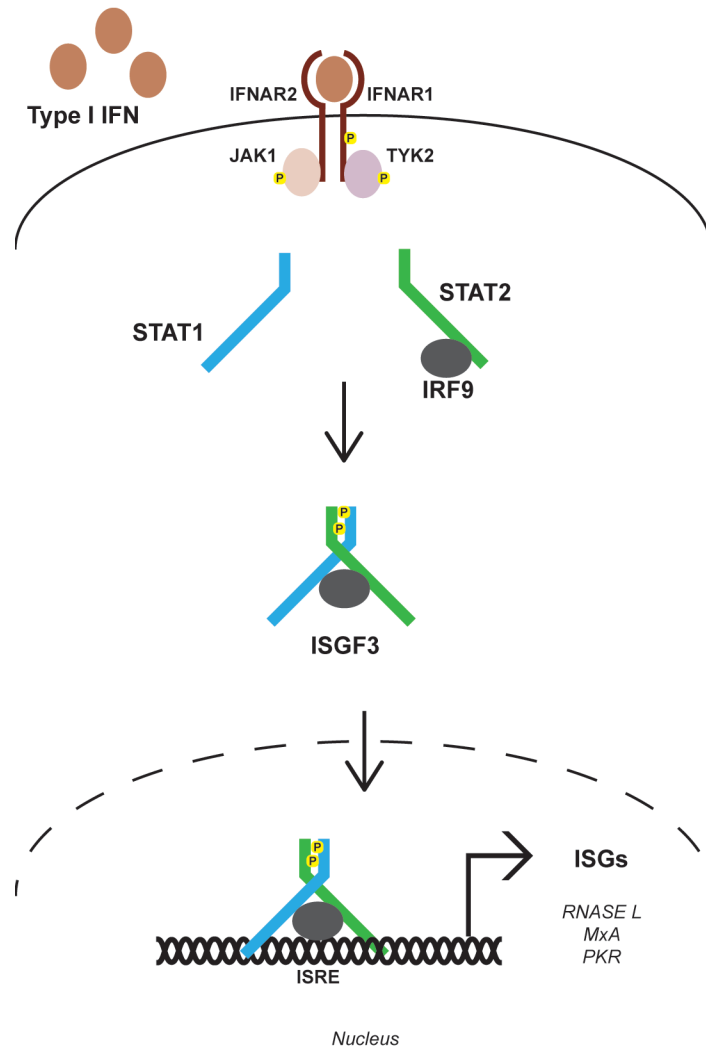


Figure 1.3 – Type I IFN signaling

Type I IFNs bind a common IFN α/β receptor consisting of the IFNAR1 and IFNAR2 receptor chains. STAT1 and STAT2 are phosphorylated by receptor associated kinases and along with IRF9 form a heterotrimeric transcription factor, ISGF3. Binding of this transcription factor to the IFN stimulated response element (ISRE) in gene promoters induces the expression of IFN stimulated genes (ISGs) such as RNASE L, MxA and PKR.

the host to counteract a virus infection (McNab et al., 2015; Stark et al., 1998). However, effective negative regulation of this response is critical for survival for both the host and the virus (Fig. 1.4). For the host, tight control of the innate immune response ensures that the signaling is only transient, as excessive signaling could cause immune and autoimmune disorders (Gresser et al., 1980; Mogensen, 2009; Trinchieri, 2010). For the virus, antagonizing the innate immune response allows it to successfully establish an infection and replicate in the host.

CELL-ENCODED REGULATORS

There are many different layers of cellular negative regulation in the response to viruses. Post-translational modifications of proteins are a common method of regulation. In some instances phosphorylation and SUMOylation are activating marks that must later be removed by phosphatases as in the STAT1 dephosphorylation by TCPTP (ten Hoeve et al., 2002) or the deSUMOylation of RIG-I and MDA5 to enable their ubiquitination and subsequent degradation (Hu et al., 2017). However, these marks in other contexts can be inhibitory as in the SUMOylation of STAT1 by PIAS1 following IFN stimulation (Ungureanu et al., 2005) or the phosphorylation of RIG-I and MDA5 at steady state to keep them inactive (Wies et al., 2013).

In addition to post-translational modifications, several physical mechanisms are also employed to negatively regulate the cellular innate immune response. Degradation of molecules, both nucleic acids and proteins, is used limit signaling following virus infection.

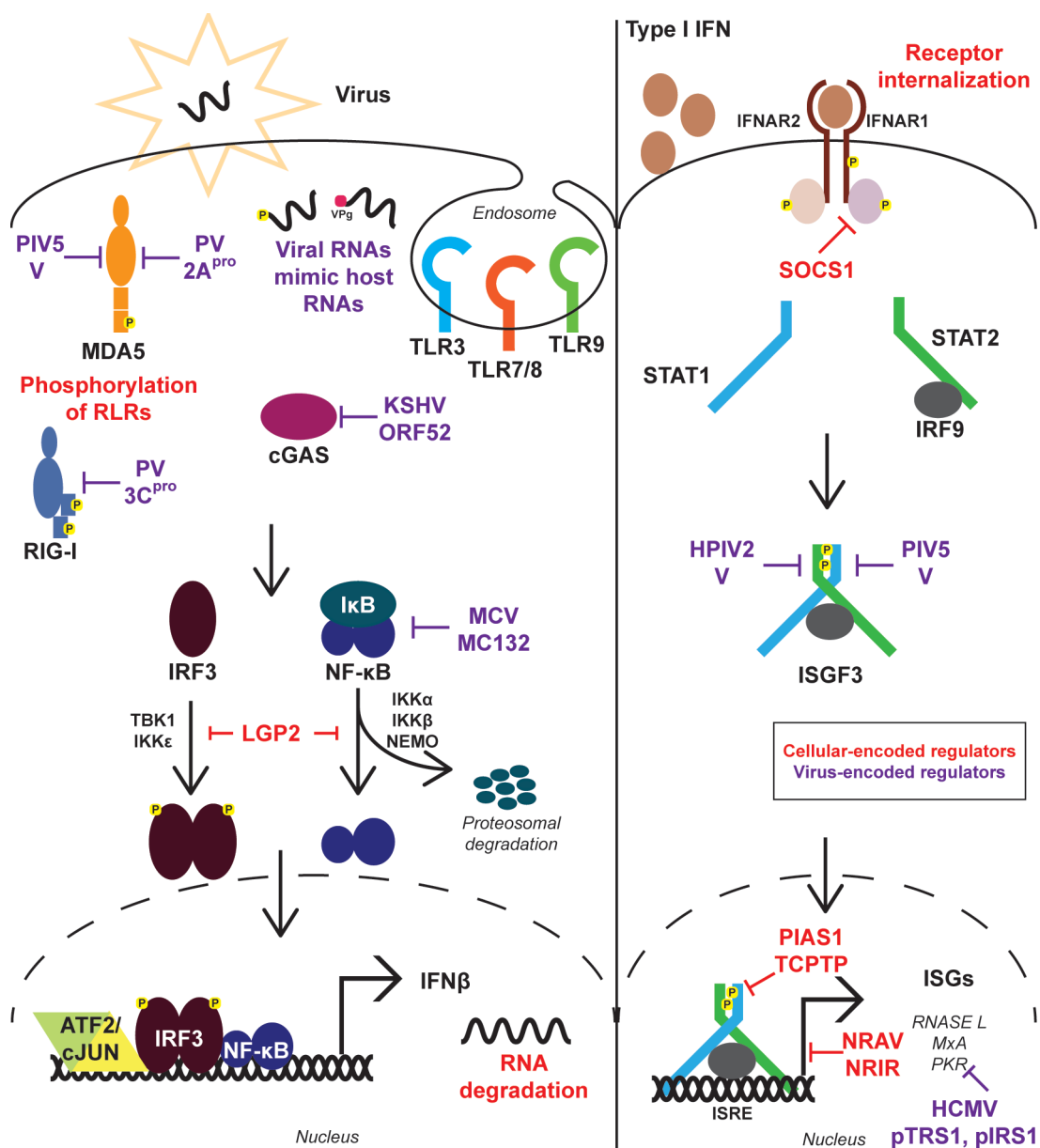


Figure 1.4 – Various regulators of the cellular response to viruses

This diagram depicts examples of the many layers of cellular- (red) and virus-encoded regulators (purple) of the cellular innate immune response. The host encodes various negative regulators to prevent inappropriate activation of the innate immune response such as the phosphorylation of RLRs. There are many different factors that attenuate the system once it has been activated such as the feedback inhibitors LGP2 and SOCS1. Viruses also inhibit the innate immune response by targeting the host PRRs, transcription factors and antiviral effectors.

TREX1 prevents excessive accumulation of DNA (Yang et al., 2007), while RNA exonucleases limit accumulation and translation of immune genes (Decker and Parker, 1994; Whittemore and Maniatis, 1990). Signaling adapter proteins such as STING are ubiquitinated and degraded to limit signal transduction following PAMP detection (Wang et al., 2015; Zhong et al., 2009). IFNAR is internalized (Zheng et al., 2011) to curtail signaling downstream of IFN. Transcriptional control, such as LGP2 inhibition of IRF3 and NF- κ B activation (Parisien et al., 2018) or SOCS1 inhibition of phosphorylation by janus kinases (Piganis et al., 2011), is also used to suppress the transcription of innate immune genes. Some of the inhibitors participating in these processes are present in the cell at constitutive levels, while many of them are induced by a feedback loop during the activation of the innate immune response.

A majority of the well-characterized negative regulators of the innate immune response are protein-coding genes. However, as noncoding RNAs emerge as key regulatory factors in other cellular processes (Djebali et al., 2012; Fitzgerald and Caffrey, 2014; Guttman and Rinn, 2012; Rinn and Chang, 2012; Yoon et al., 2013), they are now also being linked to the regulation of the cellular response to viruses. The antiviral long noncoding RNAs (lncRNAs) identified thus far mostly regulate the transcription of a small set of antiviral genes. Two lncRNAs, NRAV and NRIR, both negatively regulate the transcription of a few ISGs and suppress viral replication (Kambara et al., 2014b; Ouyang et al., 2014; Valadkhan and Gunawardane, 2016). Though the exact mechanism of NRIR inhibition is not known yet, NRAV affects the histone modification of target gene

promoters. The lncRNA NKILA modulates the expression of a larger set of genes by attenuating NF- κ B activation (Liu et al., 2015).

This is just a small subset of key cellular negative regulators of the innate immune response. The coordinated effort of many different cellular proteins and RNAs is essential to properly control this response.

VIRUS-ENCODED REGULATORS

Viruses have evolved a number of strategies to target the host innate immune system at various points and mount a productive infection (Beachboard and Horner, 2016). A common strategy employed by many RNA genome viruses is to alter their RNA to mimic host RNAs and prevent recognition by PRRs, such as the capping of viral RNAs by influenza A virus, the converting of 5'-triphosphates to a monophosphate by Borna disease virus or the masking of the 5' end of their genome by picornaviruses (Bowie and Unterholzner, 2008). Viruses also interfere with PRR activation as in the cleaving of RLRs by poliovirus proteases (Feng et al., 2014), and binding and inactivation of cGAS by KSHV ORF52 protein (Wu et al., 2015a) and of MDA5 by paramyxovirus V proteins (Andrejeva et al., 2004; Childs et al., 2007). Suppression of adapter proteins, as in cleaving of MAVS by PRRSV (Dong et al., 2015) and HCV (Li et al., 2005), and inactivation of STING by KSHV (Ma et al., 2015), prevents signal transduction downstream of PRR activation. Targeting of TBK1 and IKK ϵ prevents activation of the transcription factors IRF3 and NF- κ B (Beachboard and Horner, 2016). Viruses also directly target these transcription factors

as in the degradation of NF- κ B by poxvirus (Brady et al., 2015) and STAT inhibition by various paramyxovirus V proteins (Ramachandran and Horvath, 2009) to prevent the expression of immune genes. The products of immune genes are also frequently targeted by viruses such as the inhibition of PKR by HCMV (Ziehr et al., 2016) to prevent establishment of an antiviral state in the host cell. These layers of viral targeting of the cellular innate immune response enable the virus to successfully infect the host cell and continue replicating.

CONTRIBUTIONS OF THESIS RESEARCH

The cellular innate immune response is a complex signaling system in humans. Appropriate induction of this response is essential to restrict virus infections, but aberrant signaling can be detrimental and cause diseases. In the following chapters, results demonstrate that there are previously unrecognized and underappreciated aspects of the cellular innate immune response to viruses. Results identify a large cadre of virus-regulated novel, noncoding RNAs that may have regulatory or direct antiviral effects. Additionally, proof of concept studies identify viral antagonists and small molecule inhibitors as potential regulators of hyperactive immune signaling. Together, these results not only provide greater insights about the cellular innate immune response to viruses but also provide novel targets for drug design for diagnostics and therapeutics.

CHAPTER II: VIRUS INFECTION REGULATES EXPRESSION OF NOVEL RNAs

INTRODUCTION

Note: This chapter contains work under review for publication at Scientific Reports

Virus infection of human cells activates the cellular innate immune response through host cell signaling cascades that result in widespread changes in gene expression. These changes not only produce robust antiviral responses that form the basis of innate and adaptive immunity (Mogensen, 2009; Takeuchi and Akira, 2009), but also negatively regulate the response to prevent immune disorders. Though many factors involved in the cellular innate immune response have already been well-characterized, recent advances in genome-wide studies are revealing a major gap in the current understanding of these pathways. The ENCODE project demonstrated that contrary to previous understanding, most of the genome, including intronic and intergenic regions, is in fact transcribed into RNAs (Djebali et al., 2012). Though that analysis was done at basal conditions, a ChIP-sequencing study of virus-induced transcriptional regulation suggested that transcription from intronic and intergenic regions was not just limited to steady-state transcription (Freaney et al., 2013). Since many key factors in the response to viruses are themselves virus-regulated, these findings led us to hypothesize that there are previously unrecognized, virus-regulated RNAs with antiviral roles in the cellular innate immune response.

Cellular innate immune response to viruses

Induction of the cellular innate immune response occurs upon recognition of viral components by host pattern recognition receptors (PRRs). Though RNA- and DNA-

genome viruses are detected by distinct PRRs (Bruns and Horvath, 2014; Kawasaki and Kawai, 2014), the cellular response to most virus infections culminates in the activation of master transcription factors IRF3 and NF- κ B. These transcription factors drive the expression of many antiviral genes that mediate virus interference and signal amplification, including the type I interferons (IFNs) (Honda and Taniguchi, 2006; Kim and Maniatis, 1997; Taniguchi et al., 2001). The IFNs bind their cognate receptors and induce the expression of hundreds of downstream antiviral effectors known as interferon stimulated genes (ISGs) (McNab et al., 2015). The ISGs establish an antiviral state in the host cells and limit virus replication.

Genome-wide sequencing increases detection of RNAs

ChIP-sequencing in human cells demonstrated that early responses to RNA-genome virus infections initiate genome-wide binding of IRF3 and NF- κ B at diverse loci to recruit and activate RNA polymerase II (Pol II), assemble transcriptional machinery, and induce transcription (Freaney et al., 2013). Many of these virus-inducible binding sites were in intergenic and intronic regions of the genome, suggesting the existence of novel virus-regulated RNAs (Freaney et al., 2013).

The rise in genome-wide sequencing studies has revealed many previously unrecognized RNA species including noncoding RNAs with diverse roles in cellular processes such as gene expression, post-transcriptional regulation, translation, cell cycle regulation, and immune responses (Djebali et al., 2012; Fitzgerald and Caffrey, 2014; Guttman and Rinn,

2012; Rinn and Chang, 2012; Yoon et al., 2013). Noncoding RNAs are also being linked to the cellular response to viruses. For example, BISPR is an IFN α -induced long noncoding RNA (lncRNA) that regulates the expression of tetherin, a cellular antiviral factor that blocks viral release (Barriocanal et al., 2014; Kambara et al., 2014a). NEAT1 is a virus-induced lncRNA that regulates the expression of antiviral genes such as IL-8 (Imamura et al., 2014). In the case of lncRNA-ACOD1, expression promotes rather than interferes with virus replication by regulating cellular metabolism (Wang et al., 2017). The identification of novel virus-induced transcription factor binding sites along with the recent increase in identification of individual novel RNAs with functions related to the regulation of virus infection highlight the need for more thorough analysis of the cellular response to virus infections. There is currently a lack of understanding about the extent of RNA transcription following virus infection and what roles these virus-regulated RNAs have in the cellular innate immune response. More complete characterization of this response could allow for more rational development of diagnostic tools as well as antiviral and other immune therapies.

Contributions of thesis research

The Sendai virus-regulated transcriptome was analyzed by RNA-sequencing to bridge the current gap in knowledge about the extent of virus-regulated RNAs in the host cellular innate immune response. Not only were previously-unannotated virus-regulated RNA-encoding loci identified throughout the genome, but previously RefSeq-annotated RNAs were also identified as virus-regulated and functional in the antiviral response. The

previously-unannotated virus-regulated RNAs were primarily noncoding RNAs expressed from intergenic regions, unlike most previously-annotated RNAs which were mRNAs transcribed from exonic regions. Analysis of a subset of newly-identified previously-unannotated virus-induced RNAs revealed that they can be induced by RNA and DNA-genome viruses as well as by direct stimulation with the type I IFN, IFN α , in both general and cell-specific fashion. These findings expand the extent of known virus-regulated transcription and suggest that both coding and noncoding RNA expression are hallmarks of the cellular innate immune response to virus infections.

RESULTS

RNA-sequencing and identification of differentially expressed RNAs

To more completely appreciate the extent of virus-regulated RNA transcription, human Namalwa B cells were infected with Sendai virus. Sendai virus, a negative sense, single-stranded RNA virus of the Paramyxoviridae family, is known to potently activate the transcription factors IRF3 and NF- κ B which then induce the expression of type I IFN and other antiviral genes (Freaney et al., 2013; Johnston, 1981). Mock-infected cells and cells infected with 5 plaque forming units (pfu) of Sendai virus per cell for 6 hours were subjected to paired-end RNA sequencing. Approximately 200 million reads of 100 bp read length were obtained for each sample. On an average, 90% of total reads mapped to the human reference genome (GRCh37/hg19), indicating an overall high quality and accuracy of sequencing.

Enrichment of RNAs in mock-infected or virus-infected samples was analyzed with the DESeq2 program at a false discovery rate (FDR) of 5%. The average number of reads in the differentially expressed RNAs varied dramatically from 3 to 3.9×10^5 reads, with 50% of the RNAs having more than 2.5×10^2 reads on average (Fig. 2.1). RNAs with expression above the 75th percentile ranged from 2×10^3 to 3.9×10^5 average reads, indicating very high expression in a subset of the RNAs. The lowest expressed quantile of RNAs was excluded from this study as it would be difficult to analyze in further assays due to the low abundance of those RNAs.

From the remaining RNAs, 6210 differentially expressed RNAs with a fold change of at least 1.5 were identified for further study. Among these RNAs, 3300 were induced by virus infection, while 2910 were suppressed. The expression changes of representative groups of these RNAs were confirmed in independent samples by RT-qPCR using primers specific for virus-induced previously-annotated genes (Fig. 2.2A) and previously-unannotated RNAs (Fig. 2.2B), as well as for RNAs suppressed by virus infection (Fig. 2.2C). In all categories, changes in RNA abundance levels measured by RT-qPCR closely matched the expression changes measured in RNA-sequencing analysis.

Of the 3300 virus-induced RNAs, 1755 had an existing RefSeq annotation (Karolchik et al., 2014). The most highly induced of these annotated RNAs (Table 2.1) code for proteins known to play key roles in the cellular response to viruses. IFN, including type I IFN (Stark et al., 1998) (IFN β , IFN α 8, and IFN α 13) and type III IFN (Ank et al., 2006) (IL-29, IL-28 α

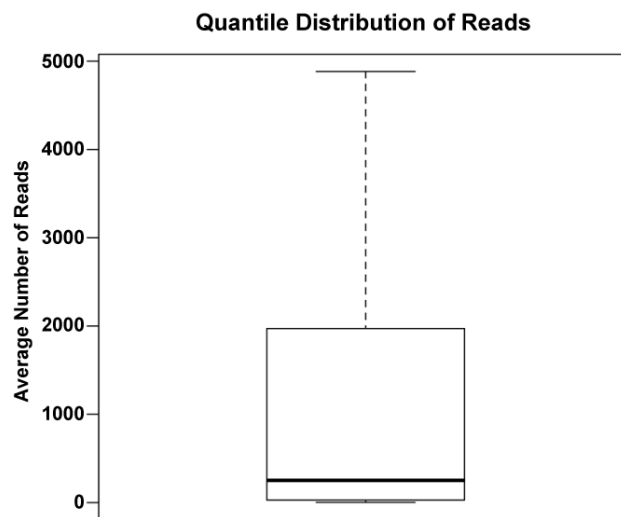


Figure 2.1 – Quantile distribution of average reads of differentially expressed RNAs

The quantile distribution of the average reads of all the differentially expressed RNAs in Namalwa cells infected with Sendai virus for 6 hours was analyzed. The average reads range from 3 to 3×10^5 reads. However, the outliers, RNAs with reads more than 1.5 times the interquartile range above the upper quantile, are not displayed for clarity. The box indicates the interquartile range or middle 50% of RNAs, with average number of reads ranging from 28 reads to 1971 reads. The median average number of reads is indicated by the solid line in the middle of the box at 252 reads.

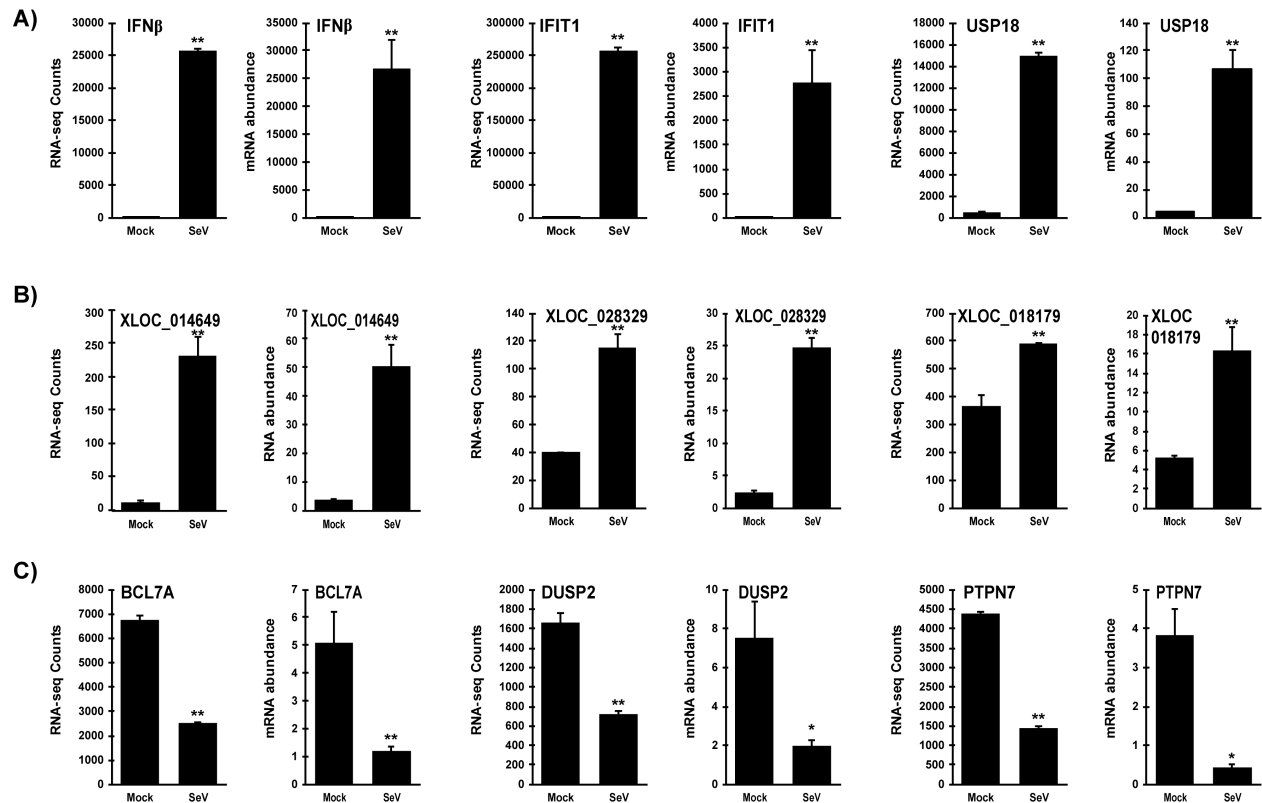


Figure 2.2 – Validation of differential gene expression induced by Sendai virus infection

Namalwa cells were infected with Sendai virus for 6 hours and total RNA was analyzed by RNA-sequencing or by RT-qPCR in independent samples. For each indicated RNA, RNA-sequencing counts are plotted on the graph on the left and RNA abundance from RT-qPCR on the right. Expression of virus-induced (A) previously-annotated genes and (B) previously-unannotated RNAs, and (C) virus-suppressed genes was validated. RNA abundance data are representative of ≥ 2 replicate experiments and are shown normalized to GAPDH expression. Bars indicate average values of technical replicates ($n=3$) with error bars representing standard deviation. Statistical analysis was done using a two-tailed Student's t-test for RT-qPCR measurements (* p-value < 0.05, ** p-value < 0.005).

and IL28 β), induces the expression of downstream antiviral effectors that interfere with viral infections by multiple methods including restricting viral replication, inhibiting transcription and translation, degrading RNAs and inducing apoptosis (McNab et al., 2015; Stark et al., 1998). The chemokines CXCL10 and CXCL11 direct T cell trafficking and differentiation by signaling through their shared receptor CXCR4 (Groom and Luster, 2011). The antiviral mediator 2'-5' oligoadenylate synthetase-like (OASL) binds the cytoplasmic RNA sensor RIG-I and enhances its activation (Zhu et al., 2015). HERC5, an E3 ligase, mediates the conjugation of ISG15 to target proteins activating various antiviral target-dependent effects such as enhancing of IRF3 activity, blocking nuclear localization of viral proteins, and blocking viral budding from the host cell (Dastur et al., 2006; Wong et al., 2006; Zhang and Zhang, 2011). In addition to the expression of these antiviral genes, 115 of the 389 previously-identified type I ISGs (Schoggins et al., 2011) were found to be induced by at least 1.5 fold in Namalwa cells infected with Sendai virus. The expression of these well-established antiviral genes verifies robust activation of the cellular innate immune response in the experimental system.

There were 1850 previously RefSeq-annotated RNAs among the 2910 virus-suppressed RNAs (Karolchik et al., 2014). Unlike the highly induced RNAs, the most highly suppressed annotated RNAs do not encode proteins that are well-known factors with critical roles in the cellular response to viruses (Table 2.2). Nevertheless, a few of them have been linked to the response to viruses. S1PR1, Sphingosine-1-phosphate receptor 1, negatively regulates IFNAR1 expression in plasmacytoid dendritic cells and its

Table 2.1 – The 10 most highly virus-induced previously-annotated RNAs

Gene	Log₂ fold change	FDR	PhastCons mean score
IL29	12.21 ± 0.41	3.34E-191	0.187
IFNβ	12.05 ± 0.37	9.06E-229	0.209
IL28β	11.77 ± 0.40	1.33E-186	0.103
CXCL11	11.66 ± 0.34	7.83E-250	0.209
IL28α	11.27 ± 0.42	3.42E-153	0.275
OASL	10.80 ± 0.19	0.00E+00	0.338
CXCL10	10.15 ± 0.13	1.39E-168	0.238
IFNα13	9.89 ± 0.45	0.00E+00	0.072
IFNα8	9.73 ± 0.45	1.17E-106	0.094
HERC5	9.72 ± 0.38	4.91E-102	0.495

Table 2.2 – The 10 most highly virus-suppressed previously-annotated RNAs

Gene	Log₂ fold change	FDR	PhastCons mean score
C10orf71	-3.92 ± 0.17	2.87E-117	0.09
S1PR1	-3.62 ± 0.11	7.56E-219	0.34
WASF3	-3.33 ± 0.38	4.16E-18	0.43
MIR4491	-3.30 ± 0.11	4.80E-183	0.05
BHLHE41	-3.00 ± 0.31	3.84E-22	0.64
HRK	-2.93 ± 0.09	1.06E-223	0.26
MGAT5B	-2.87 ± 0.08	0.00E+00	0.06
VPREB1	-2.84 ± 0.29	3.17E-23	0.22
SEMA7A	-2.82 ± 0.13	6.30E-100	0.48
ARPP21	-2.80 ± 0.13	1.25E-100	0.65

downregulation is necessary for memory CD8⁺ T cell formation (Skon et al., 2013; Teijaro et al., 2016). HRK, activator of apoptosis hara-kiri, encodes a pro-apoptotic protein of the Bcl-2 family targeted by some viruses, though it has not been studied in Sendai virus infection (Kvansakul et al., 2017). SEMA7A, semaphoring 7A, promotes the pathogenesis of West Nile virus, a positive-sense, single-stranded RNA virus (Sultana et al., 2012). The remaining RNAs, including WASF3, which is involved in actin polymerization (Teng et al., 2016), and BHLHE41, a transcription repressor (Azmi et al., 2004; Honma et al., 2002; Nakamura et al., 2008), have not previously been linked to the cellular response to viruses but their identification as virus-regulated makes them likely candidates for antiviral function.

Gene enrichment analysis

To further validate the dataset and verify inclusion of factors involved in the cellular innate immune response to viruses, differentially regulated cellular pathways were identified using DAVID Bioinformatics Resources (Huang da et al., 2009a, b). This is a functional annotation tool that determines statistically significant gene enrichment in biological processes from a given gene list. The previously-annotated virus-induced RNAs mapped to 1608 DAVID gene IDs and were significantly enriched in more than 400 Gene Ontology (GO) biological processes and 40 Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, while the previously-annotated virus-suppressed RNAs mapped to 1530 DAVID gene IDs and were only significantly enriched in less than 70 GO biological processes.

As expected, the top 10 most enriched induced GO processes described cellular functions that enable the cell to respond to virus infection, including type I IFN and cytokine signaling (Fig. 2.3A). The top 10 most enriched induced KEGG pathways also describe pathways in response to various RNA viruses such as influenza A virus, measles virus, and hepatitis C virus, a DNA virus, herpes simplex virus, and several pattern recognition receptor pathways including toll-like receptor signaling, RIG-I-like receptor signaling, and cytosolic-DNA sensing (Fig. 2.3B). Many of the top 10 most suppressed GO processes are related to GTPase signaling (Fig. 2.3C).

To better understand the extent of biological processes that are regulated after virus infection, clustering analysis was carried out on all enriched GO biological processes terms using the heuristic fuzzy multiple-linkage partitioning method (Huang et al., 2007). This method allows genes and GO processes to participate in multiple clusters to account for the diversity and complexity of biological processes. The enriched induced GO biological processes formed 27 clusters with an enrichment score above 3, indicating that a majority of the processes within the cluster were statistically significant (Table E.1). The top five most enriched clusters related to type I IFN signaling, cytokine signaling, type II IFN signaling and the unfolded protein response. The IFN and cytokine signaling pathways are known to induce essential antiviral effectors, and the unfolded protein response is a cellular stress response induced by viral protein translation in the host cell. The enriched suppressed GO biological processes formed 5 clusters with an enrichment score above 3 (Table 2.3). The clusters of suppressed processes are related to cellular

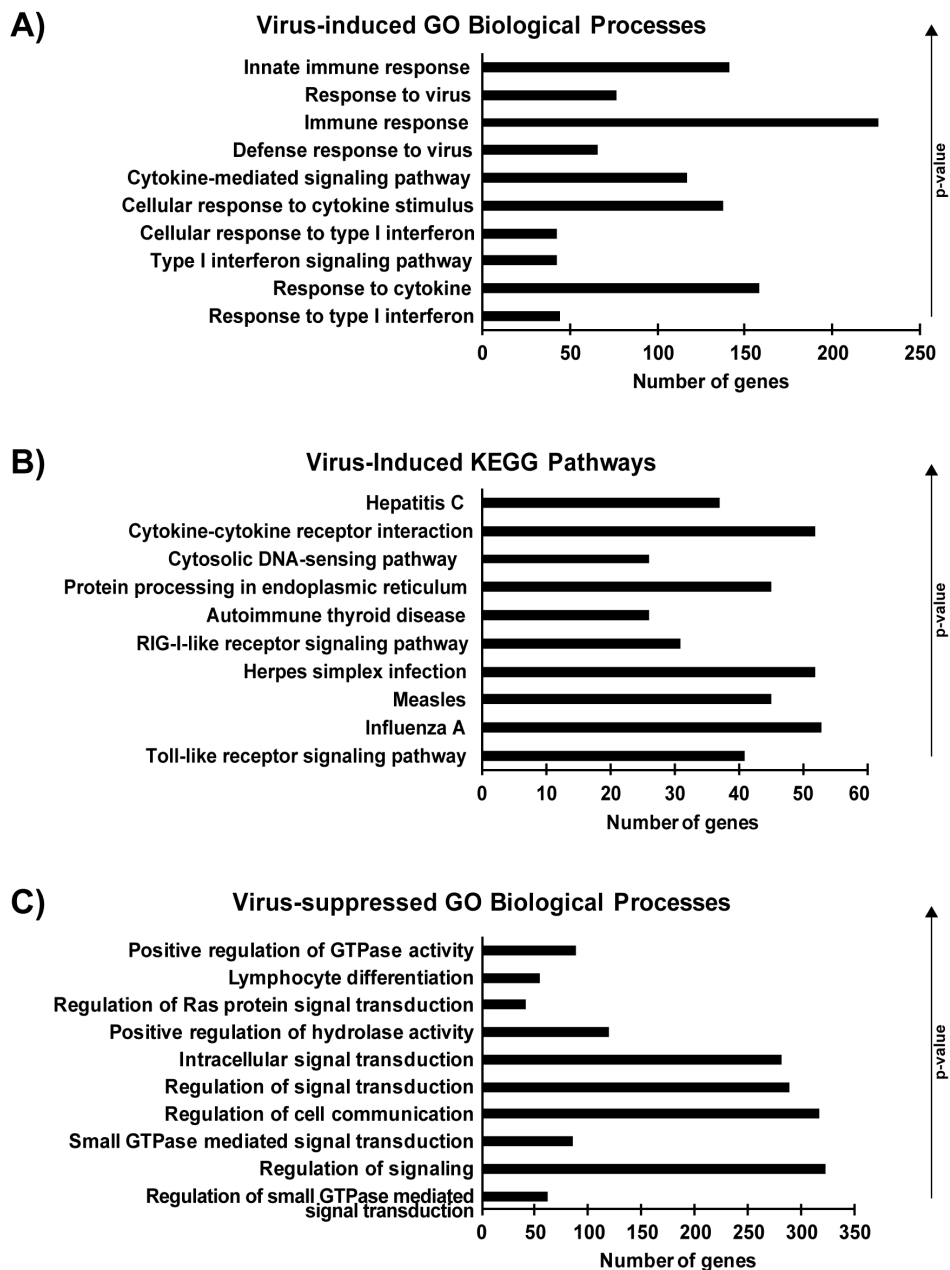


Figure 2.3 – Graphical representation of functional enrichment analysis of Sendai virus-regulated RNAs

Sendai virus-regulated previously-annotated RNAs were analyzed using DAVID to determine enriched virus-induced (A) GO biological processes and (B) KEGG pathways as well as (C) virus-suppressed GO biological processes. The top 10 most significant terms in each analysis are shown here in rank order by p-value. Bars represent number of virus-induced RNAs mapping to each term.

Table 2.3 – Clusters of enriched GO biological processes for virus-suppressed genes

Annotation Cluster 1 - Enrichment Score: 8.605						
Term	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0051056: regulation of small GTPase mediated signal transduction	62	4.05	3.66E-13	2.78	2.66E-09	7.24E-10
GO:0007264: small GTPase mediated signal transduction	85	5.56	2.13E-11	2.16	3.87E-08	4.21E-08
GO:0046578: regulation of Ras protein signal transduction	42	2.75	3.70E-09	2.77	3.37E-06	7.33E-06
GO:0043547: positive regulation of GTPase activity	88	5.75	1.94E-08	1.86	1.41E-05	3.84E-05
GO:0007265: Ras protein signal transduction	54	3.53	3.10E-08	2.26	1.74E-05	6.14E-05
GO:0043087: regulation of GTPase activity	93	6.08	3.22E-08	1.80	1.67E-05	6.37E-05
GO:0035023: regulation of Rho protein signal transduction	27	1.76	1.05E-06	2.94	3.17E-04	2.07E-03
Annotation Cluster 2 - Enrichment Score: 6.993						
Term	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0023051: regulation of signaling	323	21.11	7.64E-12	1.41	2.78E-08	1.51E-08
GO:0010646: regulation of cell communication	317	20.72	1.78E-11	1.40	4.32E-08	3.53E-08
GO:0009966: regulation of signal transduction	290	18.95	4.80E-11	1.42	6.99E-08	9.50E-08
GO:0035556: intracellular signal transduction	282	18.43	1.21E-10	1.42	1.46E-07	2.39E-07
GO:1902531: regulation of intracellular signal transduction	184	12.03	2.30E-07	1.44	7.98E-05	4.56E-04
GO:0009968: negative regulation of signal transduction	120	7.84	1.27E-05	1.47	2.88E-03	2.51E-02
GO:0010648: negative regulation of cell communication	126	8.24	3.61E-05	1.43	7.26E-03	7.13E-02
GO:0048585: negative regulation of response to stimulus	142	9.28	4.14E-05	1.39	7.88E-03	8.18E-02
GO:0023057: negative regulation of signaling	126	8.24	4.34E-05	1.42	8.06E-03	8.58E-02

GO:1902532: negative regulation of intracellular signal transduction	54	3.53	7.78E-04	1.59	7.66E-02	1.53E+00
Annotation Cluster 3 - Enrichment Score: 6.603						
Term	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0051345: positive regulation of hydrolase activity	119	7.78	2.49E-09	1.74	2.59E-06	4.93E-06
GO:0043547: positive regulation of GTPase activity	88	5.75	1.94E-08	1.86	1.41E-05	3.84E-05
GO:0043087: regulation of GTPase activity	93	6.08	3.22E-08	1.80	1.67E-05	6.37E-05
GO:0065009: regulation of molecular function	279	18.24	6.79E-07	1.30	2.24E-04	1.34E-03
GO:0043085: positive regulation of catalytic activity	167	10.92	1.06E-06	1.43	3.09E-04	2.10E-03
GO:0050790: regulation of catalytic activity	235	15.36	1.30E-06	1.33	3.65E-04	2.58E-03
GO:0044093: positive regulation of molecular function	190	12.42	1.98E-06	1.38	5.32E-04	3.91E-03
GO:0051336: regulation of hydrolase activity	144	9.41	5.12E-06	1.44	1.20E-03	1.01E-02
Annotation Cluster 4: Enrichment Score: 5.914						
Term	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0030098: lymphocyte differentiation	54	3.53	1.03E-08	2.33	8.30E-06	2.03E-05
GO:0042113: B cell activation	41	2.68	2.40E-06	2.23	6.24E-04	4.76E-03
GO:0030183: B cell differentiation	22	1.44	7.34E-05	2.63	1.13E-02	1.45E-01
Annotation Cluster 5 - Enrichment Score: 3.668						
Term	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0040011: locomotion	161	10.52	5.02E-06	1.41	1.22E-03	9.93E-03
GO:0051674: localization of cell	141	9.22	1.48E-05	1.42	3.17E-03	2.94E-02
GO:0048870: cell motility	141	9.22	1.48E-05	1.42	3.17E-03	2.94E-02
GO:0016477: cell migration	126	8.24	3.80E-05	1.42	7.44E-03	7.52E-02
GO:0030334: regulation of cell migration	80	5.23	6.77E-05	1.56	1.11E-02	1.34E-01
GO:0040012: regulation of locomotion	86	5.62	1.49E-04	1.50	1.99E-02	2.94E-01
GO:2000145: regulation of cell motility	82	5.36	2.57E-04	1.49	3.23E-02	5.08E-01
GO:0030335: positive regulation of cell migration	48	3.14	1.01E-03	1.63	8.67E-02	1.98E+00

GO:0006928: movement of cell or subcellular component	168	10.98	1.30E-03	1.25	1.00E-01	2.55E+00
GO:0051270: regulation of cellular component movement	84	5.49	1.41E-03	1.40	1.03E-01	2.74E+00
GO:2000147: positive regulation of cell motility	48	3.14	2.07E-03	1.57	1.32E-01	4.02E+00
GO:0051272: positive regulation of cellular component movement	49	3.20	2.12E-03	1.56	1.32E-01	4.11E+00
GO:0040017: positive regulation of locomotion	49	3.20	2.30E-03	1.55	1.38E-01	4.46E+00

signaling, especially small GTPase signaling, cell motility and B-cell differentiation. Though these processes are not well-established direct antiviral processes like the enriched virus-induced processes, they highlight possibly poorly studied aspects of the cellular response to viruses. Small GTPases, such as Ras and Rho, control various cellular signaling pathways for growth, adhesion, differentiation and survival (Bar-Sagi and Hall, 2000). The suppression of molecules involved in these pathways could indicate growth arrest or apoptotic signaling by the cell to restrict the virus infection. Similarly, suppression of cell motility processes could be a strategy to limit the inter- and intracellular transport of virus particles to also restrict the infection (Bohn et al., 1986; Cudmore et al., 1995; Greber and Way, 2006; Vaughan et al., 2009).

VIRUS-REGULATED PATHWAYS THAT PROMOTE VIRUS INFECTION: CHANGES IN METABOLIC GENE EXPRESSION

In addition to revealing virus-regulated genes that enable the host to restrict virus infection, this analysis also identified virus-regulated processes that promote virus infection. Further analysis of the enriched induced GO biological processes clusters showed that in addition to enrichment of direct antiviral processes, there was also an enrichment of metabolic processes in clusters 14, 19, and 20 (Table E.1). Viruses are incapable of replicating on their own and instead use host cellular machinery to propagate. They alter host metabolism to promote their own replication (Sanchez and Lagunoff, 2015). Cancer cells are also known to alter normal cellular metabolism so several successful chemotherapies target these important metabolic processes (Vander

Heiden, 2011). Identification of virus-regulated RNAs encoding proteins with roles in metabolic processes could indicate which pathways are critical to virus survival, and thus identify potential targets to consider for development of antiviral therapy (Gaelings et al., 2017; Gualdoni et al., 2018).

Of 597 genes encoding proteins known to participate in various metabolic processes, 42 were induced (7%) and 30 were suppressed (5%) following Sendai virus infection in Namalwa cells (Fig. 2.4). These virus-regulated proteins are involved in several different metabolic processes including amino acid metabolism, mitochondrial energy metabolism and oxidative stress. The five most virus-induced genes were all induced more than 15-fold following virus infection. The most highly-induced gene, CCL5, encodes a chemokine that has an additional function in regulating glycolysis by mediating glucose uptake and ATP hydrolysis (Chan et al., 2012; Gao et al., 2016). The gene CYP3A4 encodes cytochrome P450 3A4 that metabolizes xenobiotics to remove them from the body (Li et al., 1995; van Herwaarden et al., 2007). BLVRA, biliverdin reductase A, is an enzyme that converts biliverdin to the antioxidant bilirubin (Gibbs et al., 2015). It also has diverse functions in the MAPK and PI3K signaling pathways enabling it to impact cell growth and differentiation. PTGS1 encodes COX-1, a key enzyme in the synthesis of prostaglandins, lipids that have diverse hormone-like effects (Smith et al., 1996). ATP4A encodes the alpha subunit of the H⁺/K⁺ ATPase that catalyzes the hydrolysis of ATP. Though the products of the virus-regulated genes participate in diverse metabolic processes, this

Altered Metabolic Gene Expression

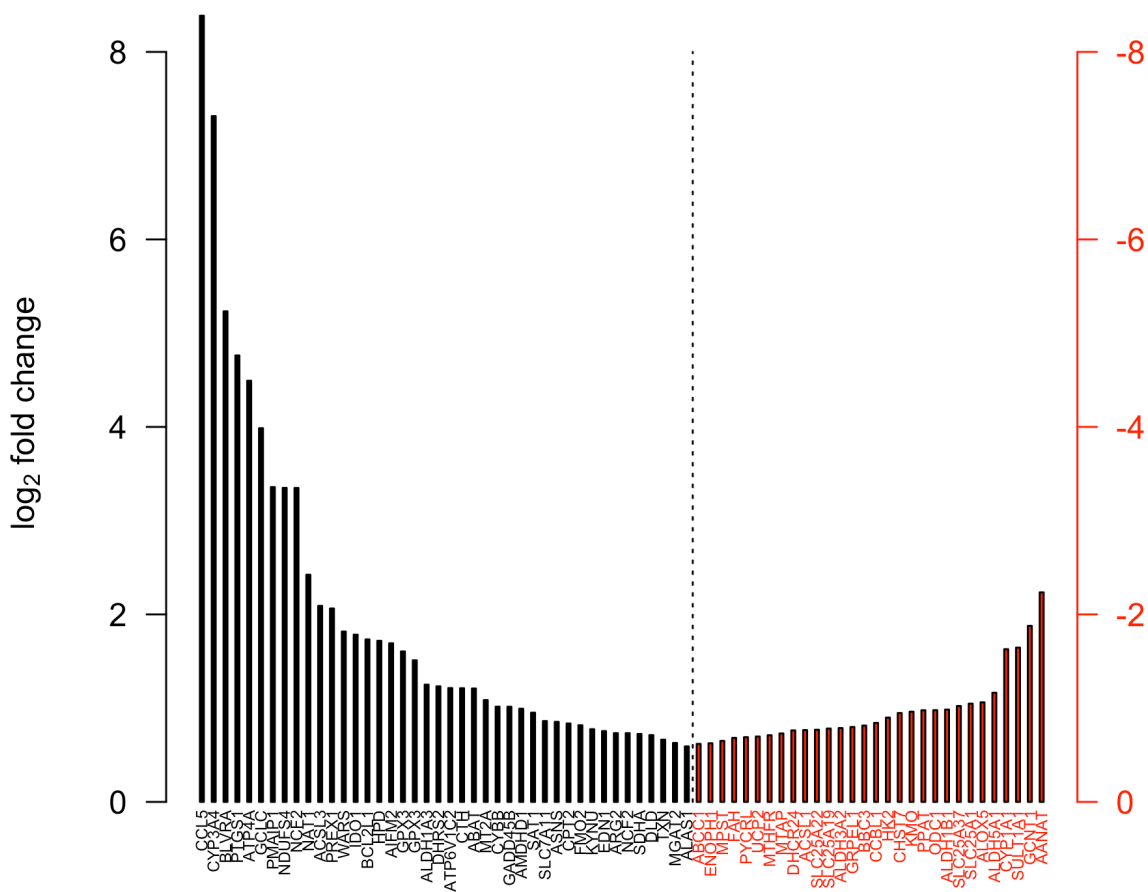


Figure 2.4 – Metabolic gene expression changes following Sendai virus infection
 Expression changes of RNAs encoding metabolic genes following Sendai virus infection for 6 hours in Namalwa cells was analyzed. Bars indicate the log₂ fold change in expression following virus infection as detected by RNA-sequencing. Fold change of virus-induced genes is shown in black and virus-suppressed genes in red.

analysis highlights metabolic processes that should be analyzed further as they may be important in Sendai virus infection of Namalwa cells.

Genomic distribution of annotated and unannotated RNAs

Advances in sequencing technology have not only revealed that genomic regions currently annotated as intergenic and intronic are in fact transcribed, but also that these regions may encode virus-regulated RNAs (Djebali et al., 2012; Freaney et al., 2013). The goal of this study was to more completely understand the virus-regulated transcriptome and identify previously unrecognized virus-regulated RNAs with roles in the cellular response to viruses. The successful identification of RNAs encoding well-established virus-regulated factors demonstrates that the experimental system used in this study is capable of uncovering factors and pathways that were previously not known to be important in the response to viruses.

Among the 3300 virus-induced RNAs there were 1755 previously RefSeq-annotated RNAs and 1545 previously-unannotated RNAs. The 2910 virus-suppressed RNAs included 1850 previously RefSeq-annotated RNAs and 1060 previously-unannotated RNAs. There was a dramatic difference in the expression levels of the various RNAs (Fig. 2.5). The 50th percentile for previously-annotated RNAs, regardless of whether they were virus-induced or virus-suppressed, had higher expression than all the previously-unannotated RNAs. This suggests that at least a subset of the previously-unannotated

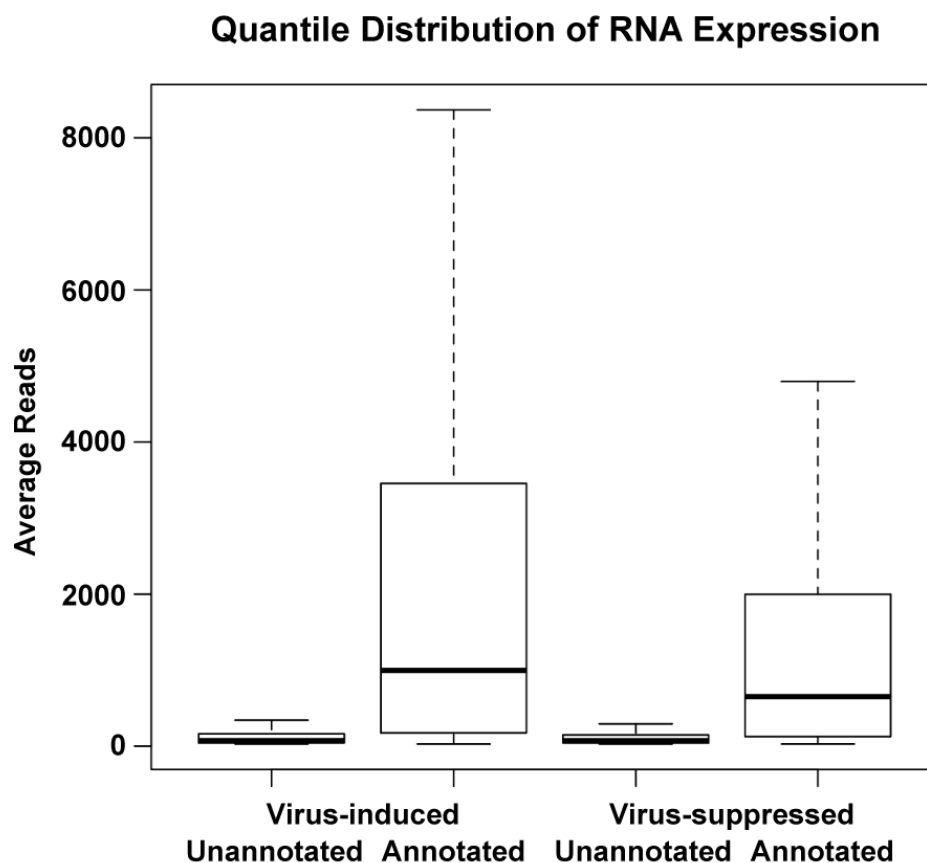


Figure 2.5 – Comparison of expression of Sendai virus-regulated previously-unannotated and previously-annotated RNAs

The quantile distribution of the average reads of virus-induced unannotated and annotated RNAs and virus-suppressed unannotated and annotated RNAs in Namalwa cells infected with Sendai virus for 6 hours was analyzed. Each box indicates the interquartile range or middle 50% of RNAs based on expression of each category with the line marking the median value. The outliers, RNAs with reads more than 1.5 times the interquartile range above the upper quartile, are not displayed for clarity.

RNAs could be noncoding RNAs as they are usually expressed at much lower levels than protein-coding RNAs (Derrien et al., 2012).

All these RNAs were found to be widely distributed among all the chromosomes (Fig. 2.6A-D), indicating that transcriptional regulation by virus infection is not limited to a specific genomic region. However, there were significant differences in the distribution of the virus-regulated RNAs among specific genomic features (Fig. 2.6E-H). The distribution of RNAs was examined across six genomic feature categories, including: promoters, transcriptional termination sites (TTS), coding exons, introns, intergenic regions and untranslated regions (UTRs; includes 5' and 3' UTRs). Most the previously RefSeq-annotated RNAs mapped to coding exons for both the virus-induced (87.6%; Fig. 2.6E) and virus-suppressed (80.2%; Fig. 2.6G) RNAs. Less than 15% of the virus-regulated previously-annotated RNAs mapped to introns and intergenic regions indicating that the current RefSeq annotations largely reflect mRNA-encoding genes. In contrast, more than 90% of the previously-unannotated RNAs mapped to introns or intergenic regions. Only 1.7% of the virus-induced previously-unannotated RNAs (Fig. 2.6F) and 1.3% of the virus-suppressed unannotated RNAs (Fig. 2.6H) mapped to coding exons. These findings support the hypothesis that there is previously unrecognized virus-regulated transcription and are consistent with previous work showing abundant virus-induced RNA polymerase II binding and elongation throughout intronic and intergenic regions (Freaney et al., 2013).

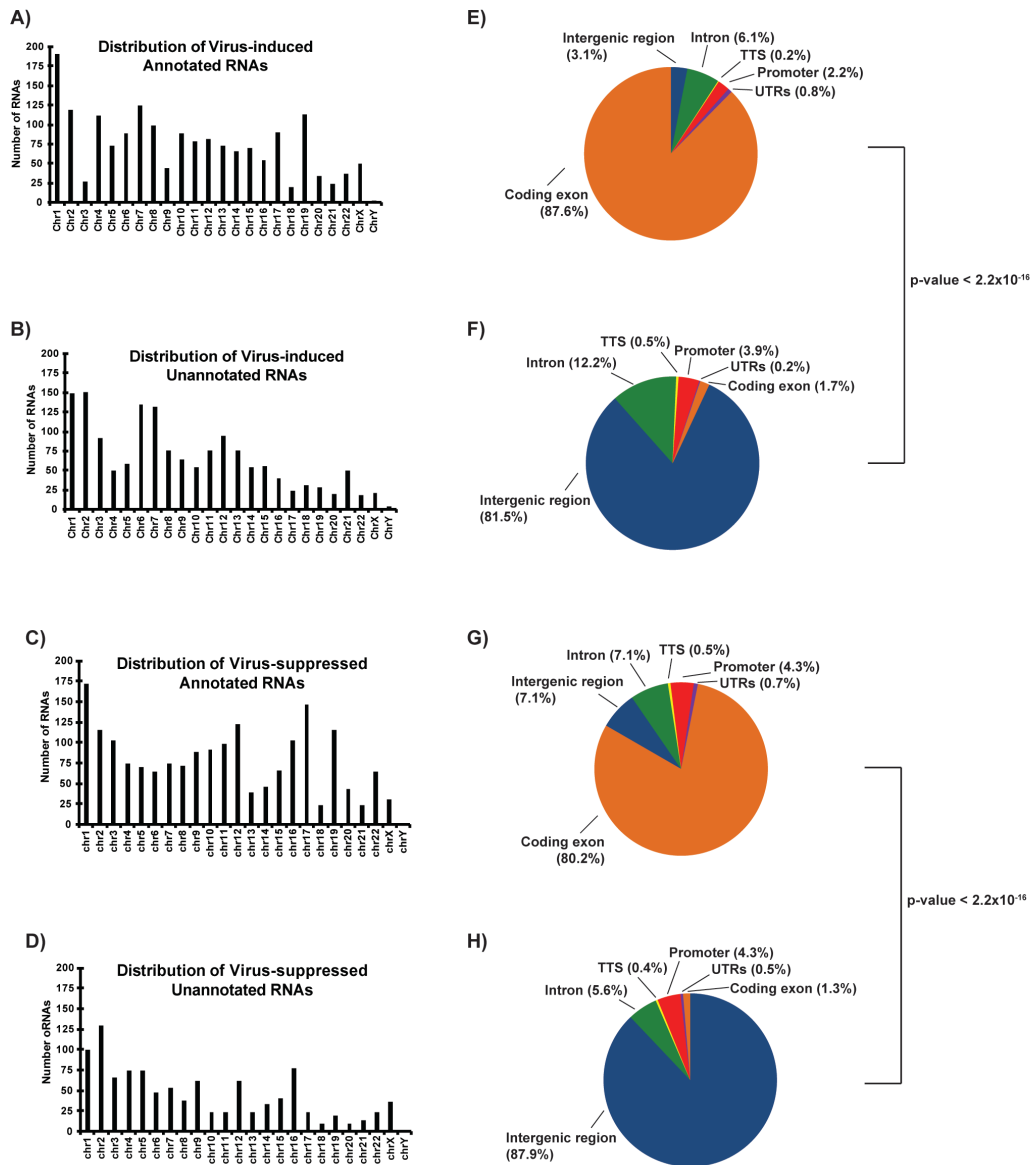


Figure 2.6 – Comparison of genomic distribution of Sendai virus-regulated previously-annotated and previously-unannotated RNAs

Bar graphs represent the number of RNAs identified on each chromosome for virus-induced (A) previously-annotated and (B) previously-unannotated RNAs, and virus-suppressed (C) previously-annotated and (D) previously-unannotated RNAs. Pie charts illustrate genomic feature distribution of virus-induced (E) previously-annotated and (F) previously-unannotated RNAs, and virus-suppressed (G) previously-annotated and (H) previously-unannotated RNAs. RNAs are mapped to one of six annotation categories: promoters, transcriptional termination sites (TTS), exons, intergenic regions, introns, and untranslated regions (UTRs; includes 5' and 3' UTRs), with the percentage of sites corresponding to each category displayed in parentheses near the label. Statistical analysis was done using Fisher's exact test.

Protein coding analysis of annotated and unannotated RNAs

Following identification of virus-regulated previously-unannotated RNAs, a PhyloCSF score was calculated for each RNA to determine protein coding potential and broadly classify the RNAs by possible function. PhyloCSF is a method that uses a multi-species sequence alignment to calculate a score reflecting the likelihood that an open reading frame (ORF) encodes a protein (Lin et al., 2011). Analysis of previously characterized RNAs has shown that known noncoding RNAs generally score below 50 while protein-coding genes score above 50 (Guttman et al., 2010; Guttman et al., 2013).

There were 30 virus-induced RNAs and 4 virus-suppressed RNAs that did not have an ORF and were excluded from this analysis. The remaining 3270 virus-induced RNAs and 2906 virus-suppressed RNAs were analyzed for their protein coding potential. A majority of the previously-annotated RNAs, regardless of whether they were virus-induced (77.6%, Fig. 2.7A) or virus-suppressed (79%, Fig. 2.7D), were predicted to be protein-coding mRNAs. These results highlight that previous studies have largely focused on the identification of proteins and there is a lack of knowledge about noncoding RNAs. This RNA-sequencing analysis was able to fill this gap and identify more of the noncoding RNAs as only 4.2% of the virus-induced unannotated RNAs (Fig. 2.7A) and 5.8% of the virus-suppressed unannotated RNAs were predicted to encode a protein (Fig. 2.7D). The classification of the previously-unannotated RNAs as primarily noncoding confirms the earlier suggestion of their noncoding status based on their low expression levels (Fig. 2.5).

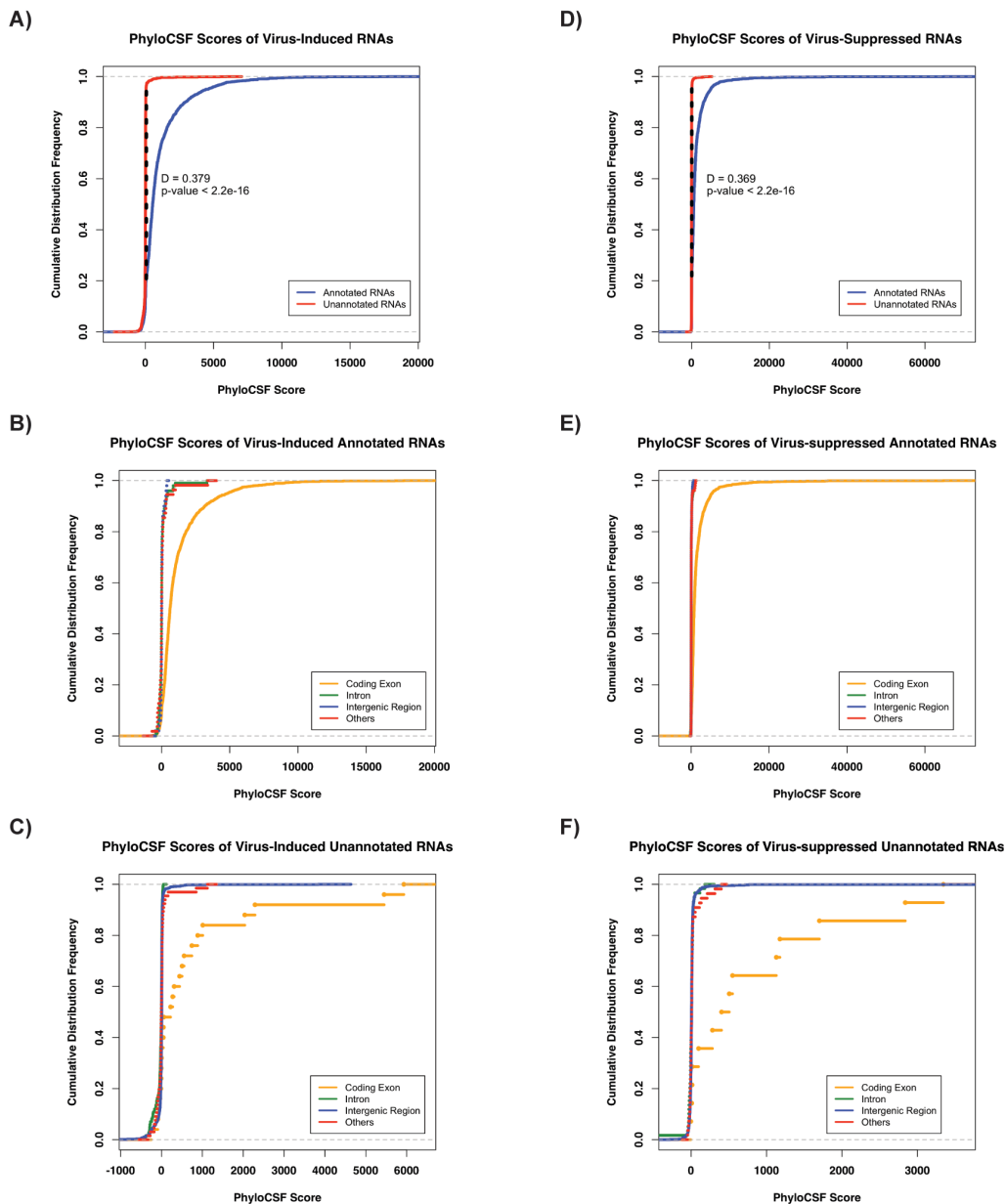


Figure 2.7 – Comparison of protein coding potential of Sendai virus-regulated previously-annotated and previously-unannotated RNAs.

The cumulative distribution frequency of PhyloCSF scores for (A) virus-induced and (D) virus-suppressed previously-annotated (blue) and previously-unannotated RNAs (red). Statistical analysis was done using a K-S test. The cumulative distribution frequency for RNAs mapping to the different genomic feature annotations for virus-induced (B) previously RefSeq-annotated and (C) previously-unannotated RNAs and for virus-suppressed (E) previously-annotated and (F) previously-unannotated RNAs. The different genomic feature annotations include: coding exons (orange), intergenic regions (blue), introns (green), and promoters, transcriptional termination sites, and UTRs (all grouped together; red).

A larger proportion of the RNAs mapping to exonic genomic features had a high PhyloCSF score (Fig. 2.7B-C,E-F). For the virus-induced RNAs, 87% of the previously-annotated RNAs and 52% of the previously-unannotated RNAs mapping to exons had a PhyloCSF score above 50 (Fig. 2.7B-C). Similarly, for the virus-suppressed RNAs mapping to exons, 92.6% of the previously-annotated RNAs and 71.4% of the previously-unannotated RNAs had a PhyloCSF score above 50 (Fig. 2.7E-F). These results indicate that the previously-unannotated RNAs are mostly virus-regulated noncoding RNAs, though there are 64 virus-induced and 10 virus-suppressed previously-unannotated RNAs that were classified as potential protein-coding mRNAs.

Vertebrate conservation analysis of virus-regulated annotated and unannotated RNAs

Though most of the virus-regulated previously-unannotated RNAs are unlikely to code for a protein, they may nevertheless have important regulatory functions in the cellular response to viruses. Many immune genes have undergone positive, diversifying selection to evolve important antiviral functions (Kosiol et al., 2008). Positive selection in protein-coding genes can be analyzed by calculating a ratio of the nonsynonymous mutation rate to the synonymous mutation rate. However, this analysis could not be used for previously-unannotated RNAs as they are predicted to be noncoding RNAs. Instead, their sequence conservation among vertebrates was analyzed. Sequence conservation is frequently used to determine functional capacity of RNAs, but in the case of virus-regulated RNAs, lower conservation may be indicative of an antiviral role.

Sequence conservation was calculated by PhastCons, a method which uses a multiple-species alignment to calculate a score for each nucleotide corresponding to the probability that it is within a conserved element (Siepel et al., 2005). To determine average sequence conservation across vertebrates, the mean PhastCons score for each RNA was calculated. Overall, the previously-annotated RNAs had much higher sequence conservation than the previously-unannotated RNAs for both virus-induced (Fig. 2.8A) and virus-suppressed RNAs (Fig. 2.8D).

For the virus-induced RNAs, while only 50% of the previously-annotated RNAs had a score of 0.413 or less, 98% of the unannotated RNAs had a score of 0.410 or less, indicating that their sequences were much less conserved than the annotated RNAs (Fig. 2.8A). The maximum mean PhastCons score for a previously-unannotated RNA was 0.687, whereas there were 24 previously-annotated RNAs with scores of 0.9 or higher, suggesting that the virus-induced unannotated RNAs on average represent a novel evolving group of RNAs. For the virus-suppressed RNAs, 50% of the previously-annotated RNAs had a score of 0.379 or less, but 97% of the unannotated RNAs had a score less than 0.379 (Fig. 2.8D), again suggesting that the previously-unannotated RNAs may represent a more recently evolved group of RNAs.

Comparison of the PhastCons score with virus-inducibility revealed that the most highly induced previously-annotated RNAs were not highly conserved among vertebrates, in accordance with these genes having evolved under positive selection (Fig. 2.8B, blue

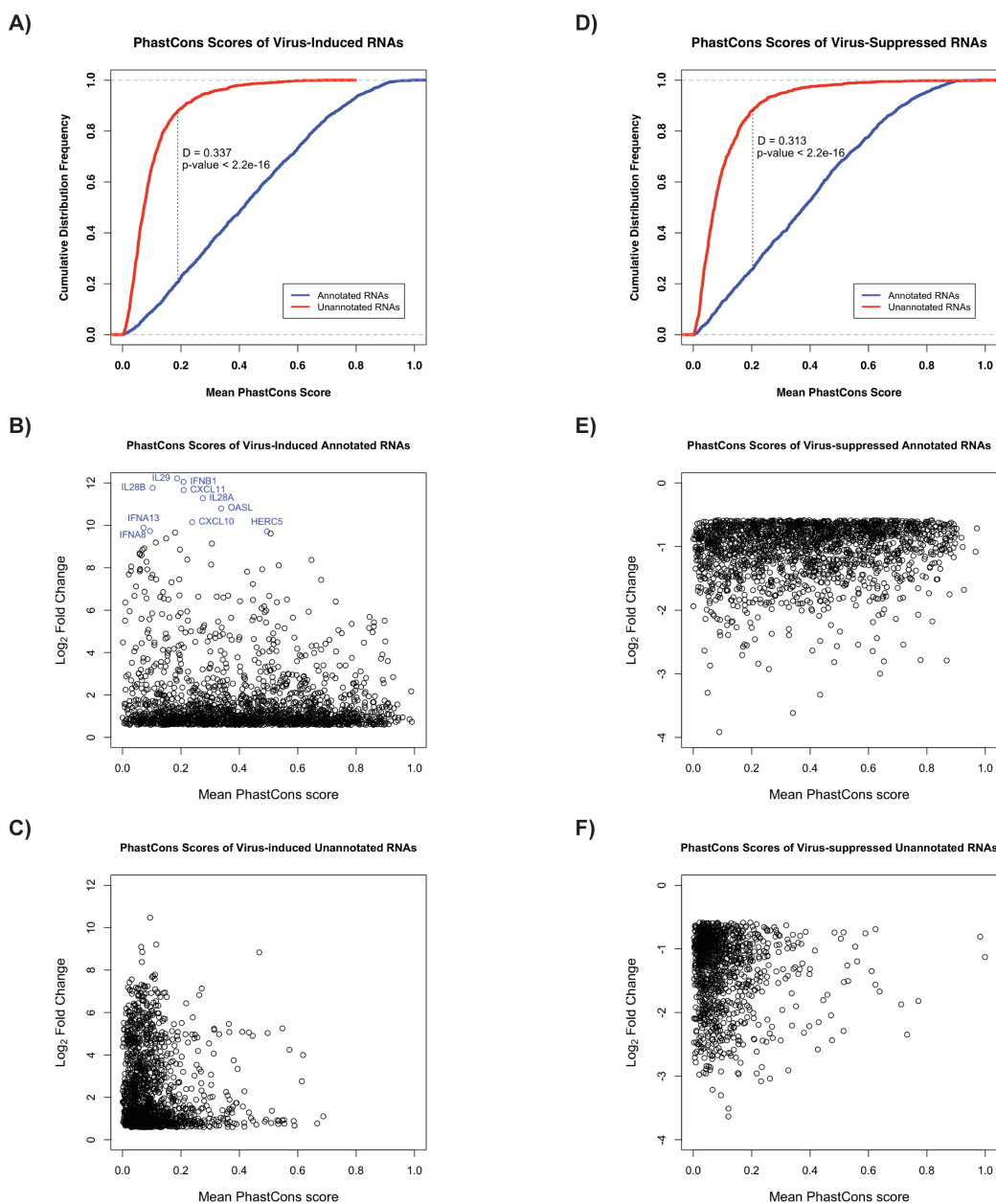


Figure 2.8 – Comparison of vertebrate sequence conservation of Sendai virus-regulated annotated and unannotated RNAs

The cumulative distribution frequency of the mean PhastCons score for the (A) virus-induced and (D) virus-suppressed previously-annotated (blue) and previously-unannotated RNAs (red). Statistical analysis was done using the K-S test. The mean PhastCons score for each RNA is plotted against the log₂ fold change in expression after Sendai virus infection for virus-induced (B) previously-annotated and (C) previously unannotated RNAs and for virus-suppressed (E) annotated and (F) unannotated RNAs. The 10 most highly virus-induced previously-annotated RNAs are labeled in blue in (B).

circles). The top 10 most up-regulated annotated RNAs all had a mean PhastCons score under 0.5, with the top 5 having a score under 0.3 (Fig. 2.8B; Table 2.1). As there were examples of highly virus-induced unannotated RNAs that were similarly poorly conserved (Fig. 2.8C), it is anticipated that at least a fraction of these virus-induced unannotated RNAs could have novel functions in the response to viruses. The most highly virus-suppressed previously-annotated RNAs were more diverse in their PhastCons scores (Fig. 2.8E; Table 2.2), though there were many highly suppressed, poorly conserved unannotated RNAs that may also have functions in the response to viruses (Fig. 2.8F).

Transcription regulation of virus-induced RNAs

Following virus infection, the magnitude of differential expression was much less for the suppressed RNAs compared to the induced RNAs. The expression of the most suppressed RNA, C10orf71, was suppressed only 50-fold (Table 2.2) by virus infection while the most induced RNA, IL-29, was induced almost 20,000-fold (Table 2.1). More than 480 of the virus-induced RNAs were induced by more than 50-fold after virus infection. Due to this discrepancy in range of expression change, further emphasis was placed on analyzing the virus-induced RNAs.

Sendai virus infection induces many cellular immune responses, including activation of the master innate immune transcription factors, IRF3 and NF- κ B (Freaney et al., 2013). These two transcription factors co-regulate transcription of many important immune genes, including type I IFN (Honda and Taniguchi, 2006; Kim and Maniatis, 1997).

Combining ChIP-sequencing and RNA-sequencing analysis of Sendai virus infection in Namalwa cells tested the idea that these transcription factors co-regulate the expression of virus-induced previously unannotated RNAs.

Virus-induced binding of IRF3, the p65 subunit of NF- κ B, and total RNA polymerase II (Pol II) within 5kb of the loci encoding virus-induced RNAs was analyzed. Of the 3300 loci encoding virus-induced RNAs, IRF3 bound near 739 loci, p65 bound near 137 loci and RNA Pol II bound near 770 loci. Approximately two-thirds of the RNAs transcribed from these loci were previously RefSeq-annotated. There were 56 previously-annotated RNAs (Fig. 2.9A; Table 2.4) and 19 previously-unannotated RNAs (Fig. 2.9B; Table 2.5) that may be co-regulated by the innate immune transcription factors. The previously-annotated RNAs identified in this analysis to be co-regulated include RNAs such as IFN β and CCL5 (Lin et al., 1999) that are known to be regulated by both IRF3 and NF- κ B. For most of the co-regulated loci, the ChIP-sequencing peaks for the transcription factors and RNA Pol II overlapped with the transcript, indicated by a distance of 0 (Table 2.4, Table 2.5), suggesting proximal regulation of transcription. This analysis demonstrated that at least a subset of the virus-induced previously-unannotated RNAs are likely to be co-regulated by IRF3 and NF- κ B.

Inducibility of virus-induced unannotated RNAs by diverse stimuli

Most of the well-known antiviral genes, including type I IFNs, are broad-acting effectors that are widely induced by diverse viruses and in most cell types. However, there are

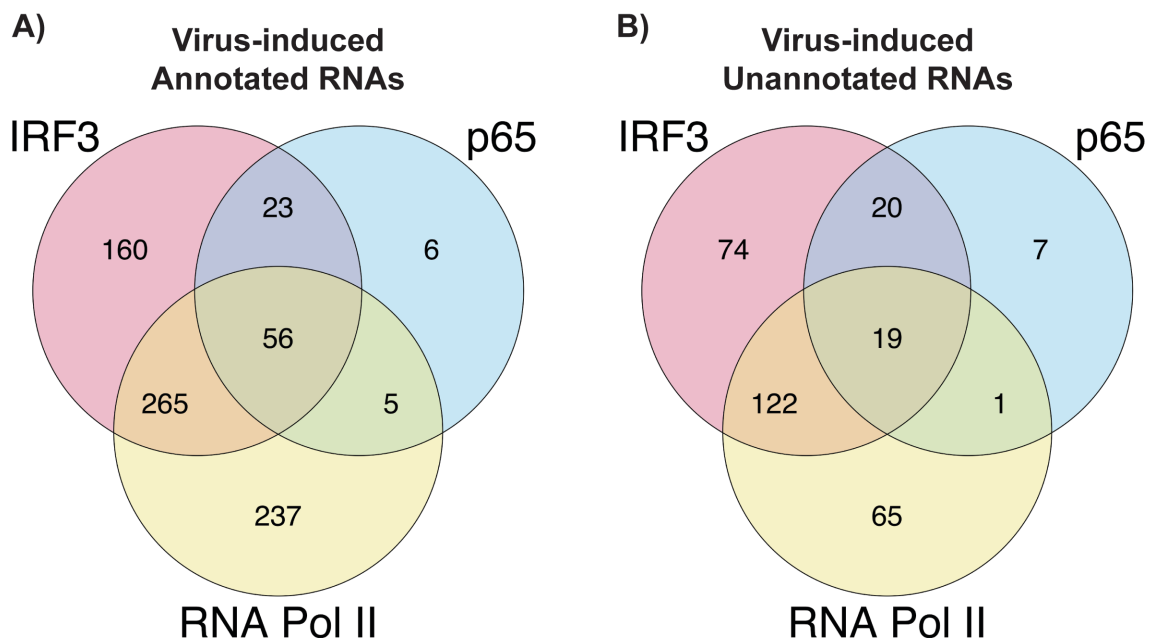


Figure 2.9 – Transcription factor occupancy at Sendai virus-induced loci

ChIP-sequencing analysis of IRF3, p65 subunit of NF- κ B, and total RNA Pol II binding in Namalwa cells infected with Sendai virus for 4 hours was combined with RNA-sequencing analysis of Namalwa cells infected with Sendai virus for 6 hours. Virus-inducible binding of the transcription factors and RNA Pol II within 5kb of loci encoding virus-induced RNAs was analyzed. Venn diagrams illustrate the degree of co-regulation by the transcription factors of (A) previously-annotated and (B) previously-unannotated loci.

Table 2.4 – Factor occupancy in virus-induced previously-annotated RNAs

Gene	Log ₂ fold change	Distance from Factor (bp)		
		RNA Pol II	IRF3	p65
IL29	12.21 ± 0.41	0	-143	0
IFNB1	12.05 ± 0.37	1184	0	0
IFIT2	9.38 ± 0.05	0	0	0
CCL5	8.39 ± 0.07	0	0	0
AFAP1	8.15 ± 0.09	0	0	0
TNF	6.60 ± 0.07	0	0	0
NRG2	5.43 ± 0.53	0	0	0
ZBP1	5.07 ± 0.12	0	0	0
ZC3HAV1	4.89 ± 0.03	0	0	-2269
HERC6	4.72 ± 0.16	-583	0	0
TRAF1	4.71 ± 0.05	0	0	0
RGS8	3.94 ± 0.42	0	0	0
CD70	3.36 ± 0.04	0	3846	-1911
CCL3	3.10 ± 0.06	0	-240	-21
B4GALT5	3.10 ± 0.04	0	0	0
LOC100506801	3.00 ± 0.21	0	0	0
SRPK2	2.72 ± 0.05	0	0	0
LINC00309	2.53 ± 0.16	0	0	0
ABLIM2	2.45 ± 0.19	4178	520	0
ICAM1	2.42 ± 0.04	0	0	0
RRAS2	2.37 ± 0.07	0	0	0
LITAF	2.13 ± 0.04	0	-47	-60
NFE2L3	2.07 ± 0.07	0	0	0
BMF	1.99 ± 0.06	0	0	0
DIP2B	1.96 ± 0.03	0	0	0
DHTKD1	1.92 ± 0.05	0	0	0
MNAT1	1.88 ± 0.07	0	0	0
LRCH1	1.87 ± 0.04	0	0	0
ALCAM	1.83 ± 0.06	0	0	0
BCL2L1	1.73 ± 0.05	0	0	0
TRIM67	1.65 ± 0.40	-924	-1398	-1427
IL2RA	1.65 ± 0.21	0	0	0

Gene	Log ₂ fold change	Distance from Factor (bp)		
		RNA Pol II	IRF3	p65
NFKB1	1.64 ± 0.05	0	0	0
RELB	1.62 ± 0.06	0	-2036	-1977
CHEK2	1.62 ± 0.05	0	0	0
NIPA1	1.60 ± 0.05	0	0	0
SNX11	1.54 ± 0.07	0	0	0
LOC678655	1.53 ± 0.11	2573	0	0
HIVEP3	1.46 ± 0.06	0	0	0
KMT2C	1.42 ± 0.04	0	0	0
SLC9A8	1.39 ± 0.08	0	-104	-172
METTL19	1.29 ± 0.05	-554	-119	0
KDM2B	1.19 ± 0.03	0	3518	0
SUSD1	1.19 ± 0.07	0	0	0
ARHGAP24	1.11 ± 0.07	0	0	0
ADAP2	1.04 ± 0.39	2205	2245	2154
PLCG2	0.90 ± 0.03	0	-4718	0
GALNT2	0.90 ± 0.04	0	0	0
FAM18A	0.85 ± 0.30	0	0	0
MLLT10	0.80 ± 0.05	0	0	0
FAM53B	0.78 ± 0.05	0	0	0
COPG1	0.77 ± 0.04	0	0	-3684
IFNAR2	0.77 ± 0.06	0	0	0
ELOVL5	0.76 ± 0.04	0	0	0
CPNE5	0.59 ± 0.06	0	0	0
IL4I1	0.59 ± 0.04	0	0	0

Table 2.5 – Factor occupancy in virus-induced previously-unannotated RNAs

Gene	Log ₂ fold change	Distance from Factor (bp)		
		RNA Pol II	IRF3	p65
XLOC_033609	8.38 ± 0.42	0	0	-37
XLOC_026525	6.90 ± 0.34	0	-117	0
XLOC_013749	5.22 ± 0.42	0	0	-48
XLOC_038794	4.92 ± 0.54	0	0	-42
XLOC_031255	4.36 ± 0.47	0	0	0
XLOC_039397	4.07 ± 0.52	0	73	0
XLOC_014649	3.99 ± 0.30	0	0	-3166
XLOC_008525	3.69 ± 0.45	-42	-525	0
XLOC_003873	3.30 ± 0.19	0	0	0
XLOC_037556	2.94 ± 0.43	0	-3263	270
XLOC_007528	2.85 ± 0.21	4418	4113	0
XLOC_019499	2.14 ± 0.34	813	0	0
XLOC_026021	2.09 ± 0.05	0	0	0
XLOC_038381	1.96 ± 0.05	0	0	-730
XLOC_038891	1.71 ± 0.30	0	0	0
XLOC_018524	1.67 ± 0.21	0	0	0
XLOC_030300	1.17 ± 0.34	0	0	0
XLOC_001315	0.96 ± 0.32	0	0	0
XLOC_015311	0.95 ± 0.25	2993	45	-4279

some antiviral genes that have more restricted antiviral activity (Schoggins et al., 2011). To determine whether virus-induced unannotated RNAs are activated in contexts other than Sendai virus infection, a subset of the virus-induced unannotated RNAs were screened in Namalwa cells to determine whether their expression could also be induced by infection with two viruses identified through the KEGG pathway analysis (Fig. 2.3B): a distinct RNA-genome virus, influenza A virus (A/Udorn/72) and a DNA-genome virus, HSV-1. In addition, cells were subjected to direct stimulation with type I IFN (IFN α). Cells were either infected with 5 pfu/cell of virus for 10 hours or treated with 1000 U/mL IFN α for 6 hours prior to RNA isolation for gene expression analysis by RT-qPCR with gene-specific primers. The known virus-induced genes IFIT1 (Pichlmair et al., 2011), IFN β (Stark et al., 1998), CSAG3 and USP18 (Porritt and Hertzog, 2015) were used as controls.

This analysis revealed great variation in the inducibility of the virus-induced unannotated RNAs, not only in which stimuli induced them, but also in the level of induction (Fig. 2.10; Table 2.6). Using a 2-fold increase in expression as a minimum for induction, the virus-induced unannotated RNAs were classified into one of eight clusters based on their behavior (Fig. 2.10): (1) inducible by viruses and IFN α , (2) inducible by viruses only, (3) inducible by RNA-viruses and IFN α , (4) inducible by RNA viruses only, (5) inducible by Sendai virus, HSV-1 and IFN α , (6) inducible by both Sendai virus and HSV-1, (7) inducible by Sendai virus and IFN α , and (8) inducible by influenza A virus or HSV-1. The virus-induced unannotated RNAs that were inducible by multiple viruses and IFN α may

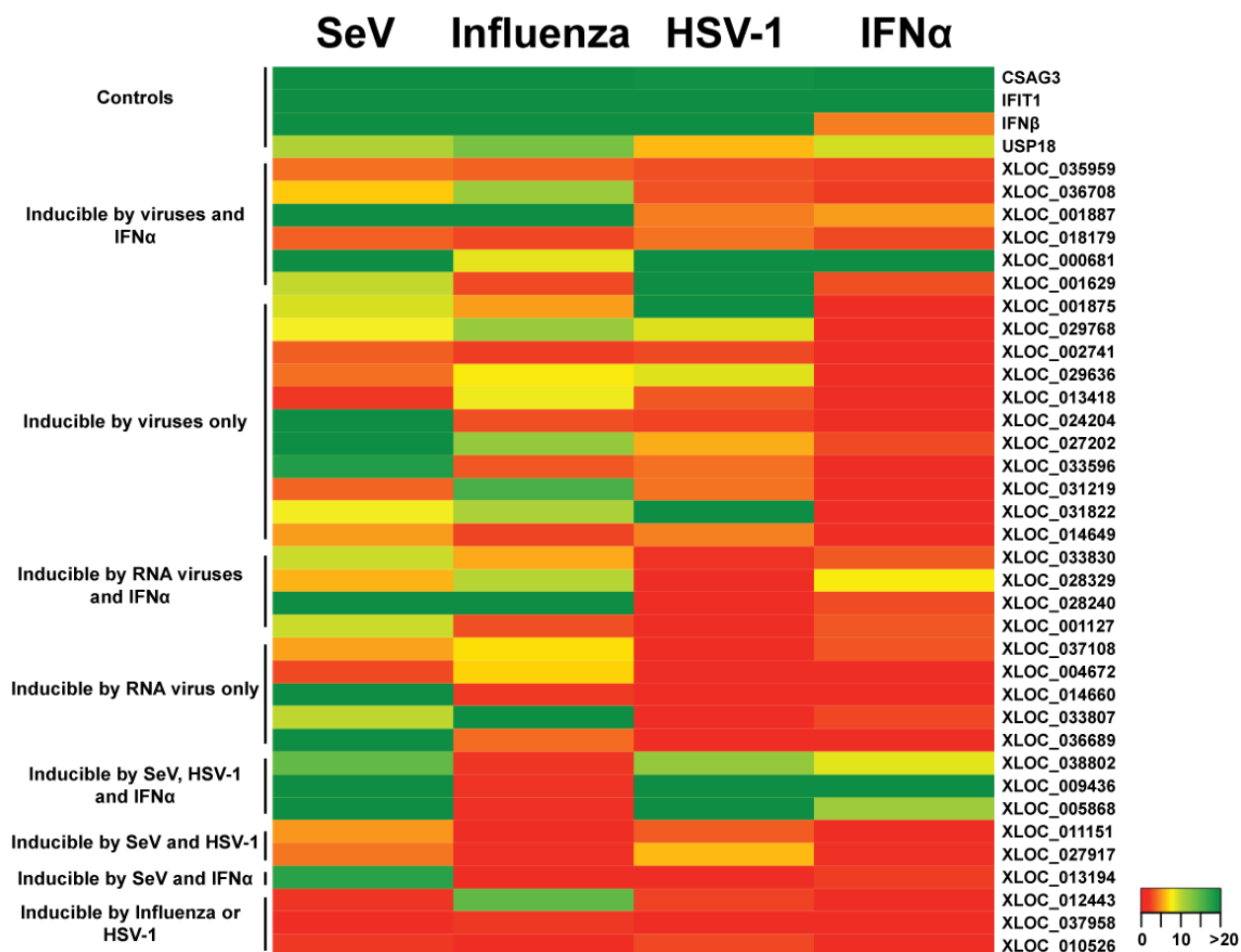


Figure 2.10 – Classification of virus-induced previously-unannotated RNA expression in Namalwa cells from virus infections and IFN treatment

Total RNA from Namalwa cells infected with 5 pfu/cell for 10 hours of Sendai virus (SeV), influenza A virus, or HSV-1 or directly treated with 1000 U/mL of IFN α for 6 hours was analyzed by RT-qPCR. Heat map indicates expression of each RNA after infection or treatment with IFN α . Average values (n=3) of fold change are reported normalized to GAPDH expression.

Table 2.6 – Details of virus-induced previously-unannotated RNA expression in Namalwa cells from virus infections and IFN α treatment

Table shows fold change measured by RT-qPCR. Statistical analysis was done using a two-tailed Student's t-test (* p-value < 0.05, ** p-value < 0.05, (n.s.) not significant)

RNA	Sendai virus	Influenza A virus	HSV-1	IFN α	
CSAG3	2,804.38 \pm 977.55**	34.10 \pm 13.36**	19.35 \pm 7.18**	20.74 \pm 4.90**	Controls
IFIT1	91,925.51 \pm 44,707.01*	35.38 \pm 8.19**	3,124.12 \pm 637.53**	4,648.01 \pm 2,184.67**	
IFNβ	5,963.05 \pm 4,968.36*	261.70 \pm 35.80*	88.95 \pm 0.00**	4.35 \pm 3.16*	
USP18	10.32 \pm 4.67**	13.39 \pm 2.50**	6.08 \pm 1.89**	9.24 \pm 1.53**	
XLOC_035959	3.97 \pm 1.20**	3.55 \pm 1.48**	2.87 \pm 0.46**	2.41 \pm 0.58**	Inducible by viruses and IFNα
XLOC_036708	6.44 \pm 4.27*	11.57 \pm 6.05**	2.97 \pm 1.24**	2.22 \pm 1.06*	
XLOC_001887	159.68 \pm 22.05**	29.56 \pm 2.07**	4.29 \pm 0.15**	5.19 \pm 0.74**	
XLOC_018179	3.42 \pm 1.48**	2.57 \pm 0.66**	4.04 \pm 1.44**	2.68 \pm 0.54**	
XLOC_000681	141.64 \pm 104.63*	8.56 \pm 2.26**	32.05 \pm 15.83**	49.52 \pm 23.44**	
XLOC_001629	9.71 \pm 6.17*	2.75 \pm 0.43**	20.29 \pm 12.71*	2.89 \pm 1.46*	
XLOC_001875	9.08 \pm 5.07**	5.24 \pm 1.87**	261.49 \pm 147.00**	0.23 \pm 0.03(n.s.)	Inducible by viruses only
XLOC_029768	7.78 \pm 2.86*	11.65 \pm 4.79**	8.88 \pm 0.44**	0.49 \pm 0.29*	
XLOC_002741	3.35 \pm 2.01*	2.30 \pm 0.77*	2.67 \pm 1.56*	0.81 \pm 0.32(n.s.)	
XLOC_029636	3.85 \pm 1.54**	7.33 \pm 4.13**	8.77 \pm 4.86**	-	
XLOC_013418	2.09 \pm 0.30**	8.21 \pm 3.53**	3.15 \pm 0.23*	0.70 \pm 0.12**	
XLOC_024204	32.14 \pm 8.09**	2.83 \pm 0.98**	2.48 \pm 0.72**	1.37 \pm 0.34*	
XLOC_027202	439.61 \pm 262.66**	11.91 \pm 5.37**	5.68 \pm 2.62**	2.71 \pm 2.03(n.s.)	
XLOC_033596	18.15 \pm 2.97**	3.13 \pm 0.44*	3.98 \pm 0.72*	0.59 \pm 0.29*	

RNA	Sendai virus	Influenza A virus	HSV-1	IFN α	
XLOC_031219	3.58 \pm 1.43**	15.91 \pm 2.97**	3.99 \pm 0.44**	0.47 \pm 0.16**	
XLOC_031822	8.12 \pm 5.05*	10.55 \pm 3.44**	47.50 \pm 19.01**	1.58 \pm 0.42*	
XLOC_014649	5.20 \pm 2.65*	2.51 \pm 0.82*	4.39 \pm 2.42*	1.44 \pm 1.71(n.s.)	
XLOC_033830	9.48 \pm 5.35**	5.50 \pm 1.45**	1.85 \pm 1.08(n.s.)	3.25 \pm 0.82**	Inducible by RNA viruses and IFN α
XLOC_028329	5.82 \pm 1.75**	10.08 \pm 6.31*	1.34 \pm 0.43(n.s.)	7.38 \pm 0.72*	
XLOC_028240	4,282.34 \pm 2,303.71**	36.67 \pm 20.16**	-	2.79 \pm 0.14*	
XLOC_001127	9.48 \pm 1.51**	2.84 \pm 0.58**	1.60 \pm 0.30*	3.13 \pm 0.92**	
XLOC_037108	5.42 \pm 0.42**	7.10 \pm 1.43**	1.59 \pm 0.30**	3.03 \pm 1.69(n.s.)	Inducible by RNA viruses only
XLOC_004672	2.67 \pm 1.56*	6.69 \pm 4.27*	0.94 \pm 0.12(n.s.)	0.23 \pm 0.14**	
XLOC_014660	22.49 \pm 6.39**	2.16 \pm 0.78*	1.64 \pm 0.67*	0.35 \pm 0.11**	
XLOC_033807	9.79 \pm 4.25**	47.69 \pm 32.29*	1.31 \pm 0.56(n.s.)	2.54 \pm 1.43(n.s.)	
XLOC_036689	57.55 \pm 7.87*	3.82 \pm 0.11**	-	-	
XLOC_038802	14.50 \pm 11.28*	1.92 \pm 0.17**	11.95 \pm 5.08**	8.70 \pm 2.24**	Inducible by Sendai virus, HSV-1 and IFN α
XLOC_009436	839.46 \pm 591.13*	1.86 \pm 0.14*	28.64 \pm 20.24*	59.79 \pm 20.39**	
XLOC_005868	21.14 \pm 16.58*	1.76 \pm 1.27(n.s.)	22.06 \pm 14.57*	11.35 \pm 4.38**	
XLOC_011151	5.03 \pm 1.86**	1.18 \pm 0.25(n.s.)	3.34 \pm 1.08**	0.57 \pm 0.16**	Inducible by Sendai virus and HSV-1 only
XLOC_027917	4.14 \pm 2.66*	1.75 \pm 0.93(n.s.)	6.04 \pm 1.28**	1.77 \pm 0.69*	
XLOC_013194	17.65 \pm 5.32*	0.32 \pm 0.01(n.s.)	-	2.30 \pm 0.20*	Inducible by Sendai virus and IFN α
XLOC_012443	1.95 \pm 0.59*	14.58 \pm 3.19**	2.38 \pm 0.87**	0.74 \pm 0.13*	Inducible by Influenza or HSV-1
XLOC_037958	1.52 \pm 0.48*	2.07 \pm 0.43**	0.86 \pm 0.48(n.s.)	0.64 \pm 0.17**	
XLOC_010526	2.07 \pm 0.93(n.s.)	1.63 \pm 0.50*	2.64 \pm 1.29*	0.94 \pm 0.20(n.s.)	

represent type I ISGs that are broad-acting novel antiviral effectors, while the RNAs induced only by a subset of stimuli may represent restricted antiviral effectors.

To investigate the cell-specificity of these virus-inducible previously-unannotated RNAs transcriptional responses, the inducibility after virus infection and IFN α stimulation was also examined in several other human cell types. The epithelial 2fTGH cells and monocyte THP-1 cells were infected with 5 pfu/cell of Sendai virus (for 4 hours due to cytopathic effects) or influenza A virus (10 hours) or directly treated with IFN α (6 hours). THP-1 cells were also infected with 5 pfu/cell of HSV-1 (10 hours). In 2fTGH cells, transfection of the synthetic dsRNA, poly(I:C), an IFN activator that can robustly induce the expression of IFNs and ISGs, was included as a stimulus. The known virus-induced genes IFN β (Stark et al., 1998), IFIT1 (Pichlmair et al., 2011), CSAG3 and USP18 (Porritt and Hertzog, 2015) were expressed in both 2fTGH cells (Fig. 2.11A) and THP-1 cells (Fig. 2.12A) by various stimuli.

The 6 virus-inducible previously-unannotated RNAs identified as most likely to be broad-acting antiviral effectors in the Namalwa cells screen were also induced in both 2fTGH cells (Fig. 2.11B) and THP-1 cells (Fig. 2.12B). In 2fTGH cells, all 6 of these previously-unannotated RNAs were induced by poly(I:C), though XLOC_035959 was induced less than 2-fold (Fig. 2.11B). Expression of the other 5 previously-unannotated RNAs was also up-regulated more than 2-fold by influenza A virus. In THP-1 cells, influenza A virus

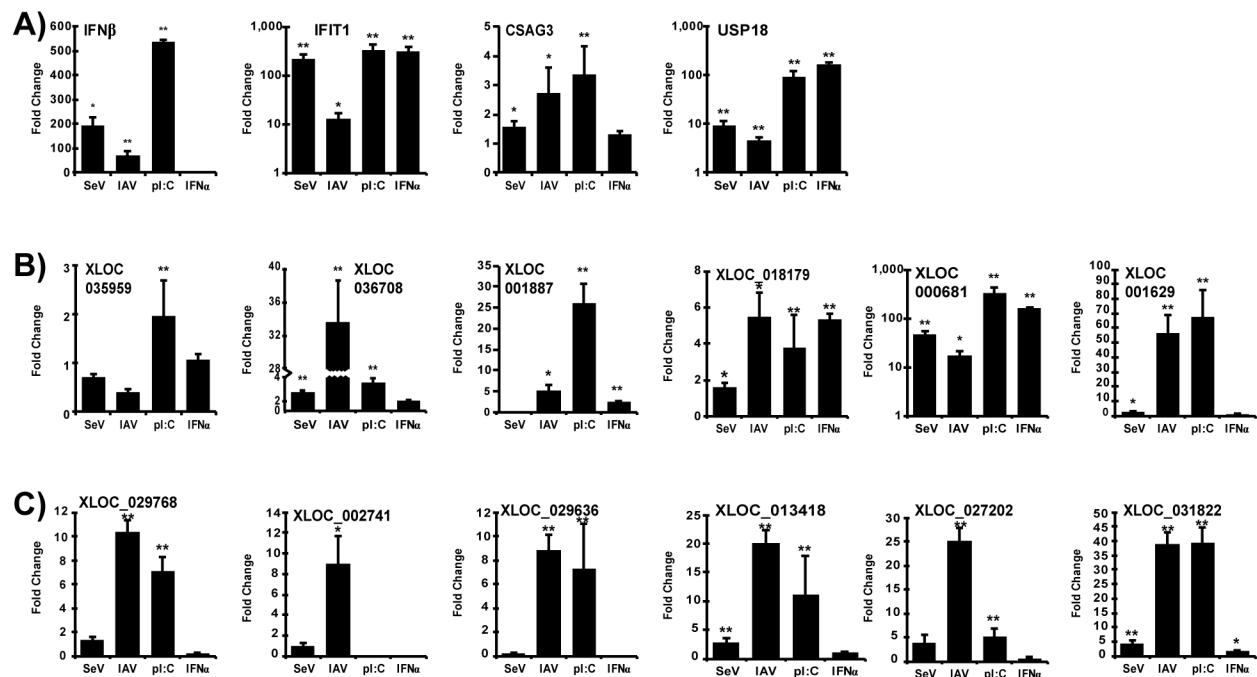


Figure 2.11 – Expression of virus-induced previously-unannotated RNAs in 2fTGH cells from various stimuli

Total RNA from 2fTGH cells infected with 5 pfu/cell Sendai virus (4 hours) or influenza A virus (IAV; 10 hours), or transfected with synthetic dsRNA polyI:C (pl:C; 6 hours), or directly treated with IFN α (6 hours) was analyzed by RT-qPCR. Gene-specific primers were used for A) control genes, B) virus-inducible previously-unannotated RNAs inducible by viruses and IFN α in Namalwa cells and C) virus-inducible previously-unannotated RNAs only inducible by viruses in Namalwa cells. Data are representative of ≥ 2 replicate experiments and is shown normalized to GAPDH expression. Bars indicate average values of technical replicates ($n=3$) with error bars representing standard deviation. Statistical analysis was done using a two-tailed Student's *t*-test (* p -value < 0.05 , ** p -value < 0.005).

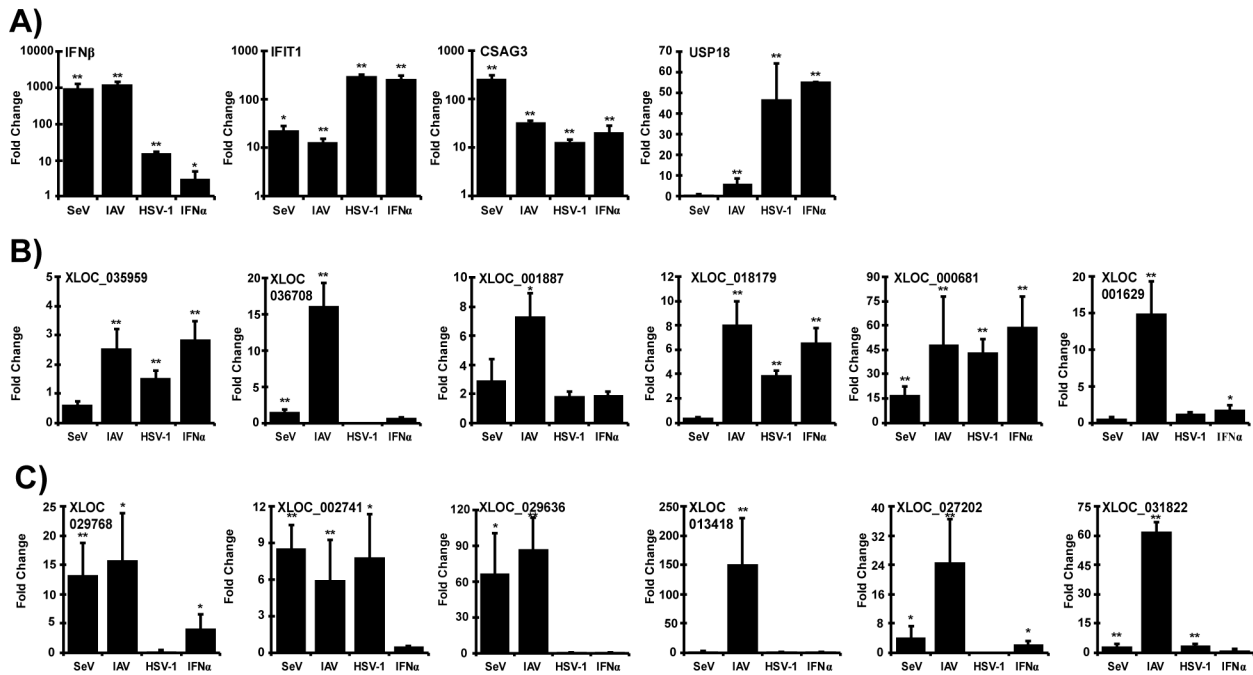


Figure 2.12 - Expression of virus-induced previously-unannotated RNAs in THP-1 cells from various stimuli

Total RNA from THP-1 cells infected with 5 pfu/cell Sendai virus (4 hours), influenza A virus (IAV; 10 hours) or HSV-1 (10 hours) or directly treated with IFN α (6 hours) was analyzed by RT-qPCR. Gene-specific primers were used for A) control genes, B) virus-inducible previously-unannotated RNAs inducible by viruses and IFN α in Namalwa cells and C) virus-inducible previously-unannotated RNAs only inducible by viruses in Namalwa cells. Data are representative of ≥ 2 replicate experiments and is shown normalized to GAPDH expression. Bars indicate average values of technical replicates (n=3) with error bars representing standard deviation. Statistical analysis was done using a two-tailed Student's *t*-test (* p-value < 0.05, ** p-value < 0.005).

robustly induced all 6 previously-unannotated RNAs (Fig. 2.12B), with HSV-1 and IFN α also inducing half of the tested RNAs.

The expression of the virus-inducible previously-unannotated RNAs in the cluster only inducible by viruses in Namalwa cells was also tested in other cell types. There was greater variability in the inducibility of these previously-unannotated RNAs. There were 6 previously-unannotated RNAs induced in both 2fTGH cells (Fig. 2.11C) and THP-1 cells (Fig. 2.12C) by 10 hours of influenza A virus infection, while the other 5 previously-unannotated RNAs from this cluster showed more varied induction (Table 2.7, 2.8).

The expression level and induction of each virus-inducible previously-unannotated RNA varied greatly depending on the inducing stimulus, likely due to differences in the activation kinetics or intensity of each stimulus. Of the 12 virus-inducible previously-unannotated RNAs expressed in both 2fTGH cells and THP-1 cells, 10 RNAs were also expressed in the epithelial HeLa and A549 cells (Table 2.7, 2.8). The expression of these previously-unannotated RNAs by at least 1 stimulus, and in most cases more than 1 stimulus, in multiple cell types demonstrates that these previously-unannotated RNAs are not cell-specific and are candidates for antiviral function.

Identification of negative regulators of the cellular innate immune response

In addition to uncovering previously-unannotated RNAs that may have roles in the cellular response to viruses, the RNA-sequencing data also revealed previously-annotated

Table 2.8 – Details of virus-induced previously-unannotated RNA expression in multiple cell lines cells from virus infections, polyI:C and IFN α treatment

Table shows fold change measured by RT-qPCR. Statistical analysis was done using a two-tailed Student's t-test (* p-value < 0.05, ** p-value < 0.05, (n.s.) not significant)

RNA	2fTGH			
	IFN α (6h)	SeV (4h)	IAV (10h)	pl:C (6h)
CSAG3	1.31 \pm 0.11(n.s.)	1.59 \pm 0.20*	2.75 \pm 0.87*	3.35 \pm 1.00**
IFIT1	315.44 \pm 62.56**	218.75 \pm 44.70**	12.47 \pm 4.37*	336.84 \pm 79.96**
IFN β	0.27 \pm 0.05*	193.34 \pm 32.07*	74.09 \pm 16.09**	535.62 \pm 7.88**
USP18	163.83 \pm 18.13**	9.61 \pm 1.67**	4.58 \pm 0.65**	91.89 \pm 29.21**
XLOC_035959	1.08 \pm 0.10(n.s.)	0.71 \pm 0.08(n.s.)	0.40 \pm 0.07**	1.97 \pm 0.72**
XLOC_036708	1.12 \pm 0.16(n.s.)	2.25 \pm 0.20**	34.63 \pm 4.98**	3.33 \pm 0.48**
XLOC_001887	2.47 \pm 0.05**	-	5.25 \pm 1.33*	26.11 \pm 4.66**
XLOC_018179	5.34 \pm 0.34**	1.59 \pm 0.26*	5.47 \pm 1.37*	3.76 \pm 1.83**
XLOC_021911	-	-	-	-
XLOC_000681	162.11 \pm 6.75**	45.96 \pm 8.01**	17.03 \pm 3.76*	329.40 \pm 99.07**
XLOC_001629	1.27 \pm 0.04(n.s.)	2.66 \pm 0.68*	56.69 \pm 12.67**	67.47 \pm 18.79**
XLOC_001875	-	-	-	-
XLOC_029768	0.28 \pm 0.03(n.s.)	1.41 \pm 0.16(n.s.)	10.35 \pm 1.00**	7.12 \pm 1.11**
XLOC_002741	-	1.02 \pm 0.23(n.s.)	8.98 \pm 2.73*	-
XLOC_029636	-	0.21 \pm 0.10*	8.84 \pm 1.27**	7.33 \pm 3.66**
XLOC_013418	1.06 \pm 0.08(n.s.)	2.80 \pm 0.70**	20.11 \pm 2.33**	11.24 \pm 6.72**
XLOC_024204	1.16 \pm 0.30(n.s.)	-	-	35.54 \pm 9.72**
XLOC_027202	0.66 \pm 0.23(n.s.)	3.89 \pm 1.61(n.s.)	25.22 \pm 2.57**	5.13 \pm 1.62**
XLOC_033596	0.74 \pm 0.15(n.s.)	-	-	-

XLOC_031219	1.14 ± 0.06*	1.31 ± 0.03*	8.87 ± 1.02**	2.23 ± 0.36**
XLOC_031822	1.71 ± 0.26*	4.38 ± 1.00**	38.90 ± 3.92**	39.19 ± 5.66**
XLOC_014649	-	-	16.06 ± 6.50(n.s.)	-
THP-1				
RNA	IFNα (6h)	SeV (4h)	IAV (10h)	HSV-1 (10h)
CSAG3	20.62 ± 7.33**	264.29 ± 40.41**	32.99 ± 2.75**	13.05 ± 1.65**
IFIT1	268.26 ± 39.06**	22.93 ± 5.89*	13.36 ± 2.16**	296.72 ± 23.67**
IFNβ	3.06 ± 1.75*	935.47 ± 325.76**	1,196.60 ± 270.65**	15.88 ± 1.20**
USP18	55.35 ± 1.66**	1.10 ± 0.25(n.s.)	6.22 ± 2.50**	47.13 ± 16.96**
XLOC_035959	2.86 ± 0.62**	0.61 ± 0.11**	2.54 ± 0.68**	1.52 ± 0.27**
XLOC_036708	0.64 ± 0.09**	1.52 ± 0.30**	16.09 ± 3.25**	-
XLOC_001887	1.91 ± 0.22(n.s.)	2.94 ± 1.45(n.s.)	7.32 ± 1.60*	1.83 ± 0.34(n.s.)
XLOC_018179	6.63 ± 1.18**	0.39 ± 0.06**	8.04 ± 1.90**	3.88 ± 0.41**
XLOC_021911	1.28 ± 0.49(n.s.)	0.90 ± 0.63(n.s.)	11.74 ± 4.80**	0.78 ± 0.26*
XLOC_000681	59.14 ± 18.96**	17.30 ± 5.05**	48.46 ± 29.24**	43.19 ± 8.11**
XLOC_001629	1.80 ± 0.62*	0.60 ± 0.09**	14.90 ± 4.42**	1.23 ± 0.25(n.s.)
XLOC_001875	-	-	-	-
XLOC_029768	4.13 ± 2.49*	13.21 ± 5.45**	15.83 ± 7.97*	0.26 ± 0.13**
XLOC_002741	0.46 ± 0.11**	8.52 ± 1.92**	5.96 ± 3.24**	7.80 ± 3.59*
XLOC_029636	0.48 ± 0.23*	66.69 ± 33.56*	87.46 ± 26.31**	0.34 ± 0.11**
XLOC_013418	0.68 ± 0.13(n.s.)	1.32 ± 0.59(n.s.)	150.79 ± 78.42**	0.52 ± 0.29*
XLOC_024204	-	-	-	-
XLOC_027202	2.07 ± 0.98*	4.14 ± 2.98*	24.68 ± 11.89**	-
XLOC_033596	0.35 ± 0.13**	5.97 ± 3.25*	29.05 ± 5.50**	2.05 ± 0.73*
XLOC_031219	0.71 ± 0.23*	0.37 ± 0.09**	-	-

XLOC_031822	1.21 ± 0.57(n.s.)	3.03 ± 1.15**	62.25 ± 4.85**		3.48 ± 0.98**
XLOC_014649	-	-	-		-
HeLa					
RNA	IFNα (6h)	SeV (4h)	IAV (10h)	pI:C (6h)	HSV-1 (10h)
CSAG3	6.88 ± 0.87**	1.61 ± 0.08**	16.72 ± 4.54**	43.36 ± 7.75**	1.74 ± 0.38*
IFIT1	73.29 ± 14.20**	10.91 ± 1.60**	2.14 ± 0.49**	73.29 ± 19.37**	7.77 ± 2.11**
IFNβ	4.79 ± 0.35**	71.81 ± 12.66**	91.89 ± 16.95**	2,198.22 ± 1,271.88**	6.61 ± 1.80**
USP18	64.25 ± 8.97**	3.15 ± 0.11**	1.02 ± 0.09(n.s.)	53.68 ± 9.19**	2.40 ± 0.41**
XLOC_035959	1.94 ± 0.15**	0.55 ± 0.08**	0.41 ± 0.11**	6.97 ± 1.22**	1.01 ± 0.16(n.s.)
XLOC_036708	0.82 ± 0.22(n.s.)	0.63 ± 0.17**	9.82 ± 2.88**	25.93 ± 5.18**	2.32 ± 0.49**
XLOC_001887	-	-	-	-	-
XLOC_018179	1.34 ± 0.26(n.s.)	0.86 ± 0.17(n.s.)	1.22 ± 0.42(n.s.)	12.87 ± 0.99**	1.30 ± 0.43(n.s.)
XLOC_021911	-	-	-	-	-
XLOC_000681	21.43 ± 2.22**	4.65 ± 1.15**	3.05 ± 0.11**	311.73 ± 29.20**	4.76 ± 0.75**
XLOC_001629	7.74 ± 1.73*	0.93 ± 0.45(n.s.)	316.73 ± 81.83*	30.87 ± 5.87**	3.58 ± 0.82*
XLOC_001875	-	-	-	-	-
XLOC_029768	6.59 ± 0.73**	-	4.75 ± 1.16*	12.17 ± 3.33*	11.59 ± 2.61**
XLOC_002741	1.60 ± 0.24*	0.70 ± 0.34(n.s.)	3.48 ± 0.25**	22.61 ± 7.50*	0.70 ± 0.03(n.s.)
XLOC_029636	1.38 ± 0.36(n.s.)	0.97 ± 0.06(n.s.)	1.26 ± 0.31(n.s.)	105.01 ± 37.60**	0.47 ± 0.09**
XLOC_013418	0.79 ± 0.21(n.s.)	0.31 ± 0.08(n.s.)	18.54 ± 3.57**	110.21 ± 30.29**	2.41 ± 0.50**
XLOC_024204	-	-	-	-	-
XLOC_027202	5.76 ± 0.11**	2.21 ± 0.57*	10.92 ± 2.70**	33.56 ± 16.06**	1.15 ± 0.09(n.s.)
XLOC_033596	-	-	-	-	-
XLOC_031219	1.14 ± 0.14(n.s.)	0.38 ± 0.17(n.s.)	0.64 ± 0.19(n.s.)	12.51 ± 1.29**	1.16 ± 0.42(n.s.)
XLOC_031822	2.62 ± 0.69*	0.77 ± 0.33(n.s.)	23.51 ± 15.43*	141.39 ± 29.88**	3.55 ± 0.37**
XLOC_014649	-	-	-	19.65 ± 13.45**	-

	A549			
RNA	IFNα (6h)	SeV (4h)	IAV (10h)	pl:C (6h)
CSAG3	5.89 \pm 0.46**	35.61 \pm 8.82**	20.03 \pm 4.67**	2,284.87 \pm 527.87**
IFIT1	2,584.11 \pm 373.37**	6,857.20 \pm 407.32**	564.45 \pm 97.98**	52,113.31 \pm 25,134.14*
IFNβ	2.02 \pm 0.33*	2,791.36 \pm 733.09**	1,388.24 \pm 544.44**	77,375.77 \pm 7,657.04**
USP18	30.21 \pm 6.83**	13.42 \pm 1.52**	6.83 \pm 3.57**	195.27 \pm 48.79**
XLOC_035959	2.47 \pm 0.33**	0.88 \pm 0.15(n.s.)	10.49 \pm 5.17*	65.81 \pm 6.12**
XLOC_036708	0.76 \pm 0.18(n.s.)	1.30 \pm 0.36(n.s.)	78.36 \pm 24.48**	120.74 \pm 26.50**
XLOC_001887	-	-	8.53 \pm 1.51**	1,103.92 \pm 302.33**
XLOC_018179	71.80 \pm 4.67**	-	63.84 \pm 18.43**	485.67 \pm 126.56**
XLOC_021911	0.60 \pm 0.11(n.s.)	8.04 \pm 2.61**	64.89 \pm 25.07**	42.10 \pm 20.26*
XLOC_000681	158.93 \pm 25.85**	171.19 \pm 44.41**	607.55 \pm 124.78**	8,909.40 \pm 1,702.90**
XLOC_001629	0.94 \pm 0.25(n.s.)	2.23 \pm 0.38**	44.81 \pm 18.04**	322.81 \pm 50.50**
XLOC_001875	-	-	-	-
XLOC_029768	-	-	-	-
XLOC_002741	1.70 \pm 0.22(n.s.)	0.93 \pm 0.25(n.s.)	10.09 \pm 6.18*	268.08 \pm 96.39**
XLOC_029636	0.80 \pm 0.36(n.s.)	4.84 \pm 1.18**	11.53 \pm 7.76*	219.82 \pm 59.34**
XLOC_013418	0.70 \pm 0.17(n.s.)	0.69 \pm 0.09**	69.33 \pm 21.73**	121.39 \pm 35.46**
XLOC_024204	1.16 \pm 0.16(n.s.)	-	-	-
XLOC_027202	-	19.96 \pm 9.47**	294.05 \pm 75.01**	2,548.71 \pm 1,017.17*
XLOC_033596	-	-	-	-
XLOC_031219	0.81 \pm 0.18(n.s.)	0.92 \pm 0.23(n.s.)	15.84 \pm 9.16**	55.31 \pm 16.77**
XLOC_031822	1.57 \pm 0.20*	4.42 \pm 0.55**	79.57 \pm 27.30**	9,099.13 \pm 2,258.32**
XLOC_014649	-	-	-	-

lncRNAs as inducible by virus infection. While there has been a recent increase in the identification of lncRNAs through deep-sequencing analyses, very few have so far been linked to virus infections. This dataset is uniquely poised to characterize both unrecognized previously-annotated and previously-unannotated RNAs as relevant in the cellular response to viruses.

AFF1-AS1 (LOC100506746) and ZBED5-AS1 (LOC729013) are lncRNAs that were identified in the RNA-sequencing analysis but have not previously been described as virus-inducible. These lncRNAs were inducible in independent Sendai virus infections in Namalwa cells analyzed by RT-qPCR (Fig. 2.13). Their noncoding status was verified by PhyloCSF analysis. Both lncRNAs had a score well below the minimum score of 50 for most protein-coding RNAs: AFF1-AS1 had a PhyloCSF score of 8.55 and ZBED5-AS1 had a score of -42.95. AFF1-AS1 and ZBED5-AS1 also had poor sequence conservation in vertebrates with a mean PhastCons score of 0.1 for both. Though p65 virus-inducible binding was not detected at these loci, both AFF1-AS1 and ZBED5-AS1 may be regulated by IRF3 as both IRF3 and RNA Pol II bound within 5 kb of the loci in Namalwa cells after a 4 hour Sendai virus infection (Freaney et al., 2013).

The inducibility of these lncRNAs in Namalwa cells by diverse viruses was analyzed to determine if these lncRNAs could have broad-acting or more restricted effects. The kinetics of their induction by Sendai virus varied as AFF1-AS1 expression was

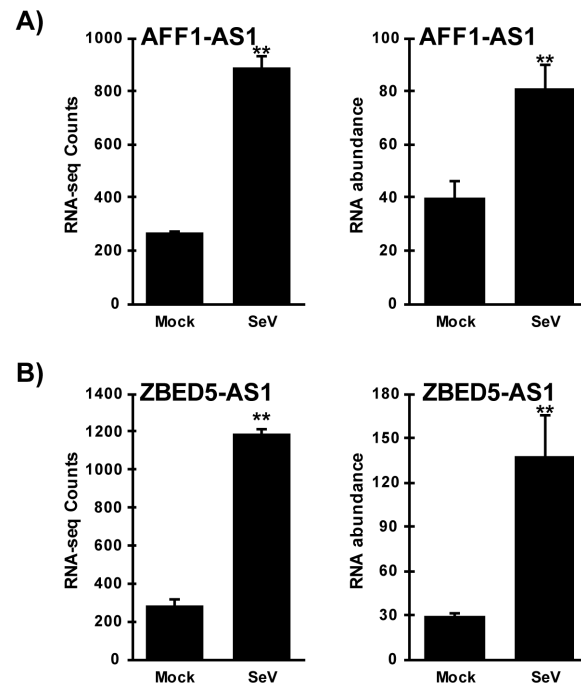


Figure 2.13 - Validation of AFF1-AS1 and ZBED5-AS1 expression induced by Sendai virus infection

Namalwa cells were infected with Sendai virus for 6 hours and total RNA was analyzed by RNA-sequencing or by RT-qPCR in independent samples. For each indicated RNA, RNA-sequencing counts are plotted on the graph on the left and RNA abundance from RT-qPCR on the right. Expression of (A) AFF1-AS1 and (B) ZBED5-AS1 was validated. RNA abundance data are representative of ≥ 3 replicate experiments and are shown normalized to GAPDH expression. Bars indicate average values of technical replicates ($n=3$) with error bars representing standard deviation. Statistical analysis was done using a two-tailed Student's t-test for RT-qPCR measurements (* p-value < 0.05, ** p-value < 0.005).

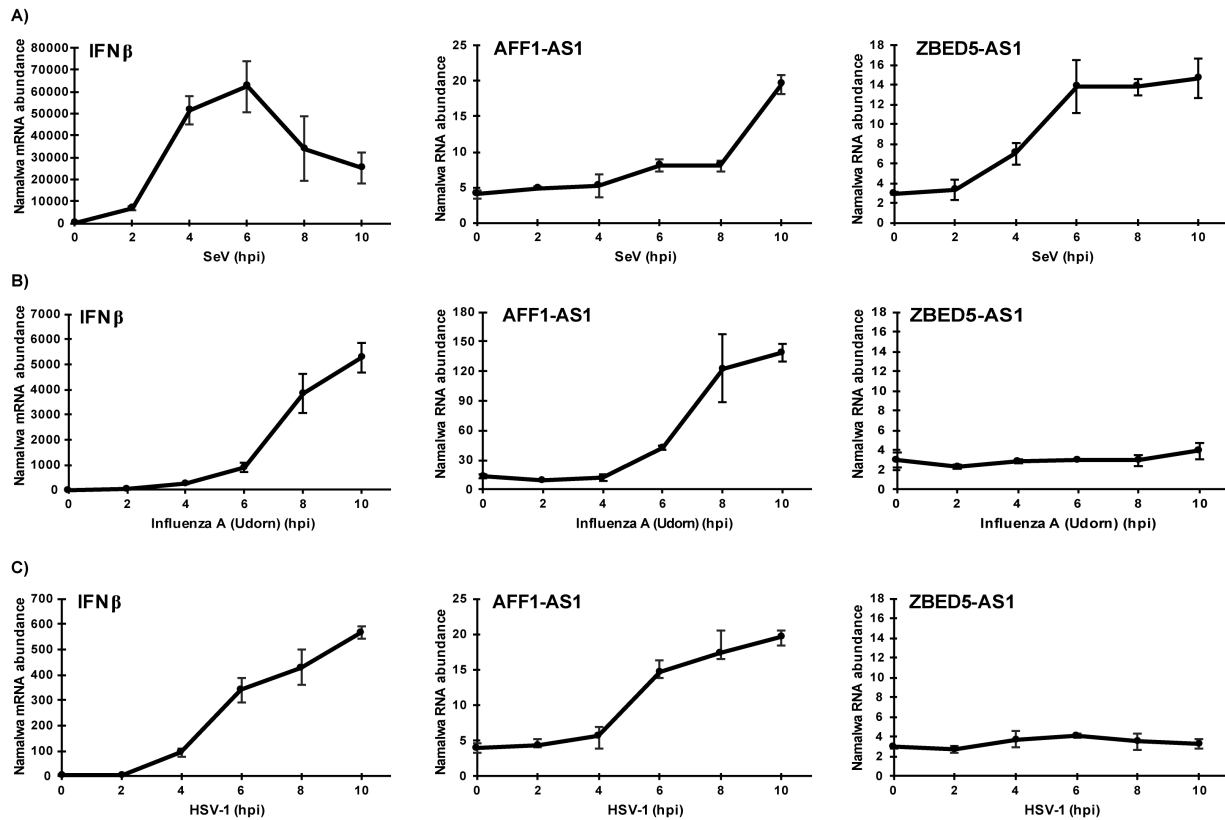


Figure 2.14 – Expression of AFF1-AS1 and ZBED5-AS1 in Namalwa cells infected with diverse viruses

Total RNA from Namalwa cells infected with A) Sendai virus (SeV), B) Influenza A virus, or C) HSV-1 was analyzed by RT-qPCR at the indicated times. Data are representative of ≥ 2 replicate experiments and are shown normalized to GAPDH expression.

upregulated even more by infections longer than 6 hours, while ZBED5-AS1 was potently induced by 6 hours (Fig. 2.14A). Influenza A virus (Fig. 2.14B) and HSV-1 (Fig. 2.14C) were able to induce AFF1-AS1 expression in Namalwa cells, but neither virus induced ZBED5-AS1 expression. The differences in their inducibility suggests that AFF1-AS1 and ZBED5-AS1 could have difference potential antiviral functions with AFF1-AS1 possibly having more broad effects since it is induced by multiple viruses in Namalwa cells.

The cell-specificity of AFF1-AS1 and ZBED5-AS1 was also tested by looking at their expression in 2fTGH cells. In these cells, Sendai virus only induced AFF1-AS1 (Fig. 2.15A). Influenza A virus (Fig. 2.15B) modestly induced ZBED5-AS1 but greatly induced AFF1-AS1 expression. The synthetic dsRNA poly(I:C) induced both lncRNAs (Fig. 2.15C). Though there was variability in their levels of induction, these results demonstrate that AFF1-AS1 and ZBED5-AS1 are not restricted to Namalwa cells in their expression.

Antiviral function potential of both lncRNAs was further assessed by loss of function analyses in 2fTGH cells. Expression of AFF1-AS1 was efficiently knocked down by lentiviral shRNA transduction (Fig. 2.16A). The loss of AFF1-AS1 expression resulted in increased basal IFN β expression, which diminished over time after transcription induction by poly(I:C) (Fig. 2.16B). This increased basal IFN β mRNA expression was seen in other independent experiments as well (Fig. C.1A-B). Loss of AFF1-AS1 expression did not affect expression of the neighboring protein-coding gene AFF1 (Fig. C.1B). AFF1-AS1 also negatively regulated IRF3 activation in multiple independent experiments (Fig. 2.16C

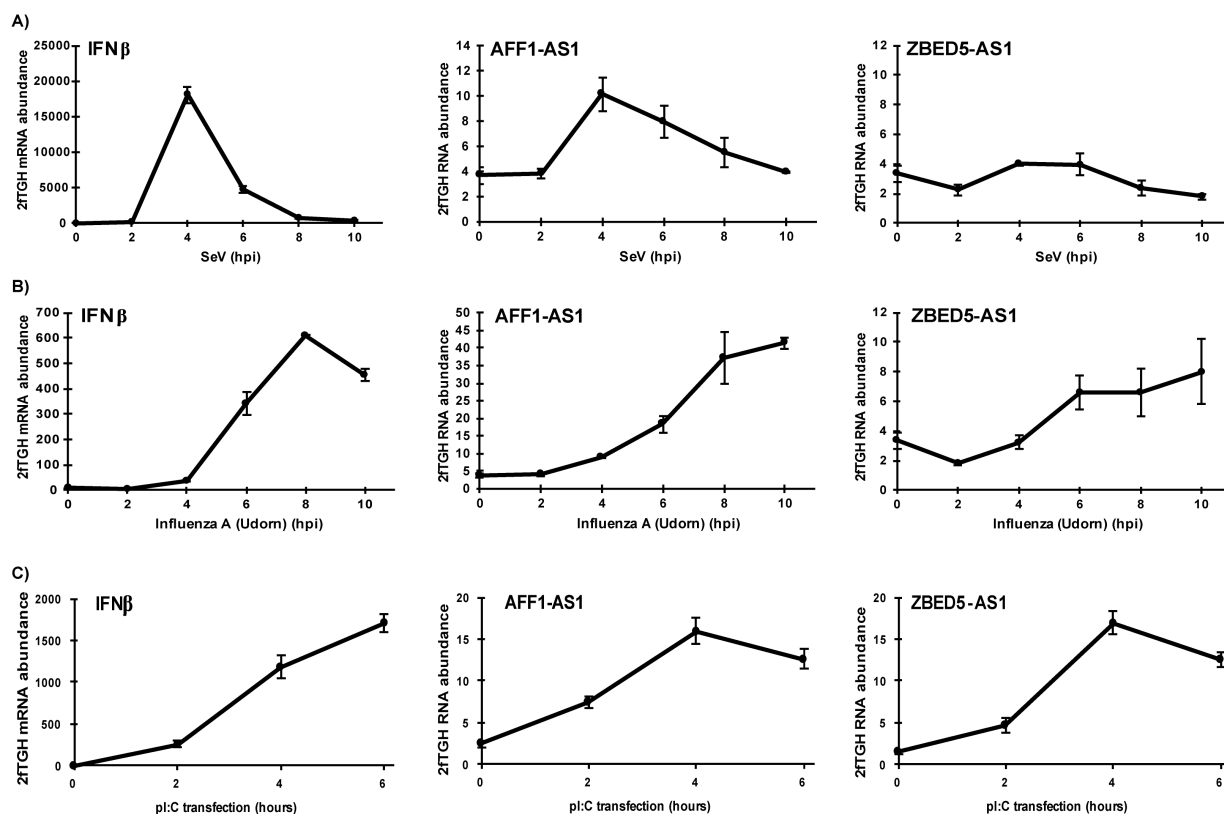


Figure 2.15 – Expression of AFF1-AS1 and ZBED5-AS1 in 2fTGH cells by diverse stimuli

Total RNA from 2fTGH cells infected with A) Sendai virus (SeV) or B) Influenza A virus, or C) transfected with poly(I:C) (pi:C) was analyzed by RT-qPCR at the indicated times. Data are representative of ≥ 2 replicate experiments and are shown normalized to GAPDH expression.

and Fig. C.2A-B), indicating that its effect on the cellular response to virus infections is prior to IRF3 signaling. ZBED5-AS1 expression was knocked down by siRNA transfection (Fig. 2.17A). Loss of ZBED5-AS1 expression also affected IFN β expression, resulting in more rapid and increased IFN β mRNA accumulation (Fig. 2.17B, Fig. C.3). Loss of ZBED5-AS1 expression did not affect expression of the neighboring protein-coding gene ZBED5 (Fig. C.3). However, unlike AFF1-AS1, ZBED5-AS1 does not affect IRF3 activation (Fig. 2.17C), suggesting that it may have a different role in the response to viruses than AFF1-AS1.

The analysis of AFF1-AS1 and ZBED5-AS1 demonstrates the power of the RNA-sequencing study to uncover novel factors that are involved in the cellular response to viruses. The data can identify regulators with subtle effects such as AFF1-AS1 and ZBED5-AS1.

DISCUSSION

While it is well-established that virus infection induces widespread gene expression changes in the host, recent technological advances have revealed a major gap in the field. Deep-sequencing studies have identified previously-unannotated RNAs transcribed from intergenic and intronic genomic regions (Djebali et al., 2012). Though very few of these RNAs have been linked to the cellular response to viruses so far, a ChIP-sequencing study showed virus-inducible binding of transcription factors in intergenic and

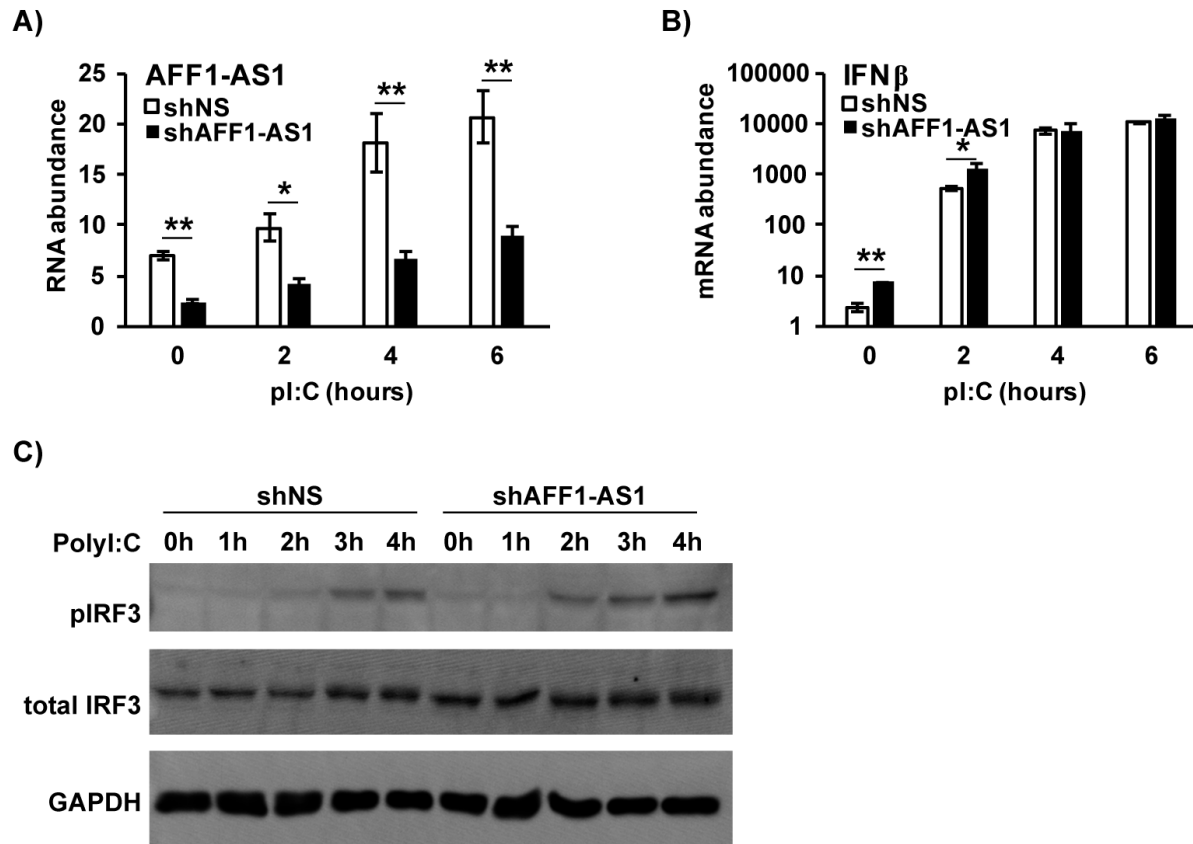


Figure 2.16 – AFF1-AS1 knockdown by shRNA in 2fTGH cells

Cells transduced with non-silencing (shNS) or AFF1-AS1 targeting shRNA (shAFF1-AS1) were transfected with poly(I:C) for the indicated times. Total RNA was harvested and expression of (A) AFF1-AS1 and (B) IFN β was analyzed by RT-qPCR. Data are representative of ≥ 3 replicate experiments and are shown normalized to GAPDH expression. Bars indicate average values of technical replicates (n=3) with error bars representing standard deviation. Statistical analysis was done using a two-tailed Student's t-test (* p-value < 0.05, ** p-value < 0.005). (C) Whole cell lysates were collected at the indicated time points and analyzed by immunoblotting with antibodies against phospho-IRF3 (pIRF3), total IRF3 and GAPDH.

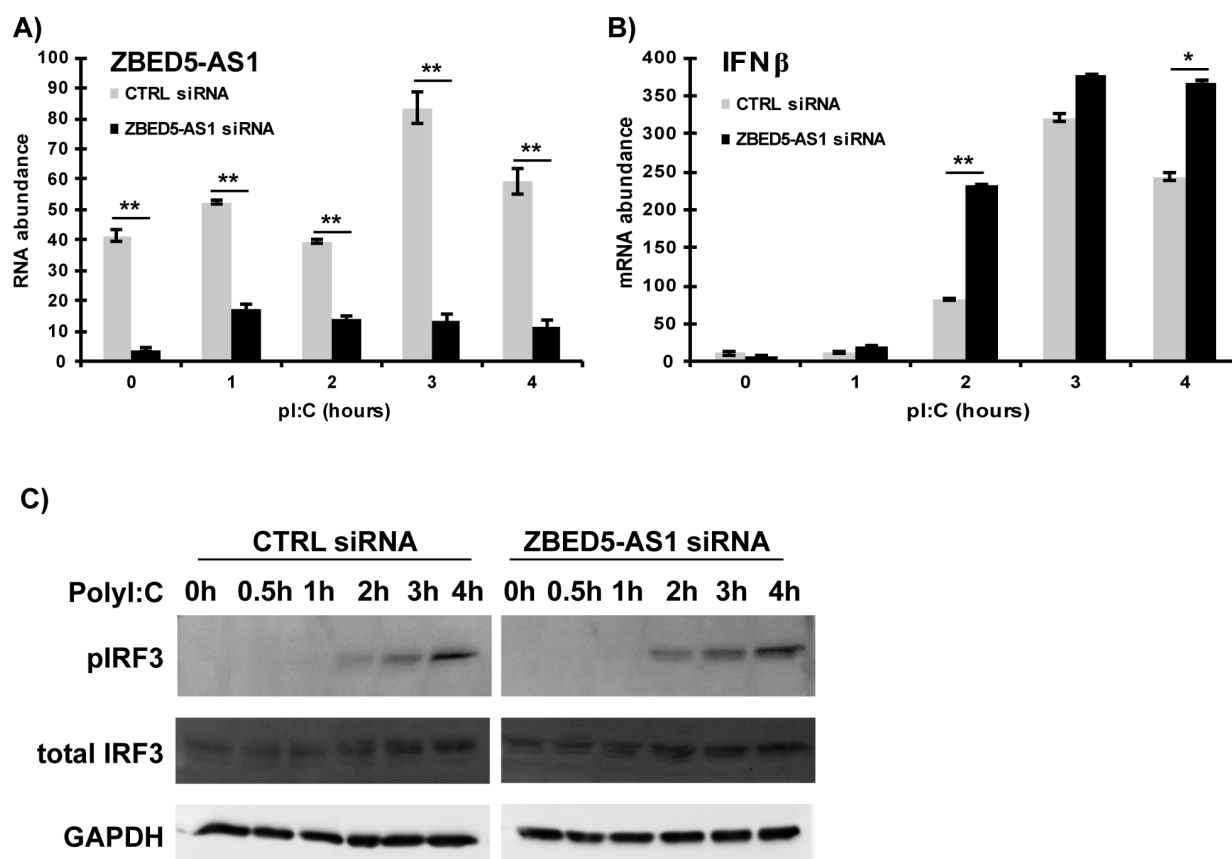


Figure 2.17 – ZBED5-AS1 knockdown by siRNA in 2fTGH cells

Cells were transfected with non-targetting control siRNA or ZBED5-AS1 targeting siRNA for 48 hours, followed by poly(I:C) transfection for 4 hours. Total RNA was harvested at the indicated time points and expression of (A) ZBED5-AS1 and (B) IFN β was analyzed by RT-qPCR. Data are representative of ≥ 3 replicate experiments and are shown normalized to GAPDH expression. Bars indicate average values of technical replicates ($n=3$) with error bars representing standard deviation. Statistical analysis was done using a two-tailed Student's t-test (* p-value < 0.05, ** p-value < 0.005). (C) Whole cell lysates were collected at the indicated time points and analyzed by immunoblotting with antibodies against pphospho-IRF3 (pIRF3), total IRF3 and GAPDH.

intronic regions genome-wide (Freaney et al., 2013). Since the expression of most antiviral effectors and regulators of the cellular response to viruses is controlled by virus infection, these findings suggested that there are many virus-regulated previously-unannotated RNAs that may have a role in the cellular response to viruses. Identification and characterization of these previously unrecognized RNAs would not only begin to fill the gap in knowledge about the cellular response to viruses, but also drive the development of new medically relevant tools such as diagnostics, antivirals, and other immune therapies.

RNA-sequencing of mock-infected and Sendai virus-infected Namalwa cells revealed the previously unrecognized depth and breadth of the virus-regulated transcriptome. The dataset identified 3605 previously-annotated and 2605 previously-unannotated virus-regulated RNAs. The ChIP-sequencing study of virus-induced binding of IRF3, NF- κ B, and RNA Pol II had predicted the expression of 1450 virus-induced previously-unannotated RNAs (Freaney et al., 2013). Though 501 of these predicted RNAs were partially matched to RNAs identified in the RNA-sequencing analysis, most were excluded from further analysis due to their low expression levels. The small number of virus-induced previously-unannotated RNAs identified as co-regulated by these factors is likely due to the limitations of this analysis (Fig. 2.9; Table 2.5). The ChIP-sequencing analysis was done after 4 hours of infection whereas the RNA-sequencing was done after 6 hours of infection to allow for RNA accumulation. These static experiments may not be able to

fully capture the dynamics of virus-induced transcriptional regulation and there may be even more previously-unannotated RNAs that are co-regulated by IRF3 and NF- κ B.

Though the virus-regulated previously-annotated and previously-unannotated RNAs were distributed similarly across the chromosomes (Fig. 2.6A-D), there was a significant difference in the genomic features associated with these different loci (Fig. 2.6E-H). More than 80% of the previously-unannotated RNAs were transcribed from intergenic regions whereas more than 80% of the previously-annotated RNAs were transcribed from exons (Fig. 2.6E-H). This discrepancy highlights that current annotations of genomic features are based on identification of a relatively small number of RNAs and the labeling of most of the genome as intergenic “junk” DNA. The identification of thousands of previously-unannotated RNAs indicates a need for a corresponding refinement in the annotations of genomic features.

In order to assess the possible functional roles of these virus-regulated previously-unannotated RNAs, their protein coding potential and sequence conservation among vertebrates was calculated. Unlike the previously RefSeq-annotated RNAs which were predominantly identified as protein coding mRNAs, very few of the previously-unannotated RNAs were likely to encode a protein (Fig. 2.7A,D). Most of these RNAs are predicted to be noncoding RNAs, suggesting that they may have regulatory functions in the response to viruses.

Sequence conservation, calculated by a mean PhastCons score, was also used to assess possible functionality of the previously-unannotated RNAs. The most highly inducible, well-known RNAs that encode antiviral factors had a mean PhastCons score below 0.5 (Fig. 2.8B; Table 2.1). While most genes have undergone negative selection to be conserved through evolution, many immune genes have been positively selected for diversity (Daugherty and Malik, 2012; Meyerson and Sawyer, 2011). Virus evolution to overcome host antiviral immunity creates a potent selective force for diversification of key antiviral effectors. The cyclical virus-host interplay has resulted in immune genes being among the more recently evolved and more diversified genes in the human genome (Kosiol et al., 2008). The poor sequence conservation of the highly virus-induced unannotated RNAs may indicate that they are previously unidentified antiviral factors that evolved under positive selection (Fig. 2.8C).

In further support of the concept of previously-unannotated RNA functionality, expression of a subset of the RNAs was analyzed by infection with different RNA and DNA viruses, and antiviral stimulation with poly(I:C) and IFN α . Differences in both expression and temporal regulation were observed among the virus-induced unannotated RNAs tested due to specific virus or treatment, or cellular host cell line. It is reasonable to speculate that at least some of the virus-induced unannotated RNAs could represent novel antiviral factors with unique roles in the cellular innate immune response to viruses. This expression analysis does not assign any specific function to these RNAs, but their

identification and classification is the first step in understanding their roles in the cellular response to virus infections.

In addition to revealing virus-regulated previously-unannotated RNAs, the RNA-sequencing data also identified previously-annotated RNAs that had not been directly linked to virus infection prior to this analysis. Though viruses are known to generally alter host metabolism, the pathways critical to infection have not been identified. Identification of specific virus-regulated RNAs that encode proteins involved in metabolic processes can highlight which processes are important for virus infection and provide novel targets for the development of antivirals and diagnostics (Gaelings et al., 2017). Similarly, the suppression of cell motility and widespread small GTPase signaling has not previously been directly shown to be involved in the cellular response to viruses. However, the enrichment of these pathways in the RNA-sequencing dataset suggests that they may be an integral component of the cellular response to viruses and should be explored in future studies.

Of the 1755 previously-annotated virus-induced RNAs, 328 were not assigned to any GO biological process, indicating that they are poorly characterized functionally. AFF1-AS1 and ZBED5-AS1 are two lncRNAs from this subset that were now identified as negative regulators of the cellular innate immune response. These previously-annotated, virus-regulated RNAs might represent additional new mediators whose functions are of interest to more fully understand the cellular response to viruses.

Both lncRNAs were induced in the two human cell lines tested, but they likely have different functions as negative regulators. AFF1-AS1 may function prior to IRF3 signaling to inhibit IFN β transcription, while ZBED5-AS1 has another method of suppressing IFN β mRNA accumulation. Though these lncRNAs represent minor factors in the negative regulation of the cellular immune response, they demonstrate the power of this analysis to identify novel factors in the cellular innate immune response.

**CHAPTER III: INHIBITION OF INTERFERONOPATHIC MDA5 BY PARAMYXOVIRUS
V PROTEINS**

INTRODUCTION

Note: This chapter contains work accepted for publication at Journal Interferon & Cytokine Research

Type I interferons (IFNs) are the primary antiviral cytokines expressed following virus infection. They exert antiviral activity by inducing the expression of hundreds of antiviral effectors that restrict virus replication by degrading host and viral RNAs, inhibiting DNA and RNA synthesis, inducing growth arrest, and activating apoptosis in addition to engaging innate and adaptive immune cells (McNab et al., 2015; Stark et al., 1998). The antiviral activities of the IFN response are usually rapid and transient, and unregulated IFN production can lead to cellular, immunological, and developmental defects (Gresser et al., 1980). Increased type I IFN production and signaling and a signature of chronically upregulated interferon stimulated gene (ISG) expression characterizes several diseases including systemic lupus erythematosus (SLE) (Baechler et al., 2003; Bennett et al., 2003; Crow and Wohlgemuth, 2003), rheumatoid arthritis (van der Pouw Kraan et al., 2007), type I diabetes (Ferreira et al., 2014), and Aicardi-Goutieres Syndrome (AGS) (Crow et al., 2015; Rice et al., 2013). While many mechanisms have been proposed to account for the IFN signature, recent characterization of monogenic inborn diseases, collectively known as type I interferonopathies, have elucidated several heritable mechanisms for enhanced IFN production and signaling (Crow, 2011).

Interferonopathies caused by monogenic mutations

Defects in cellular machinery involved in nucleic acid metabolism and homeostasis have been identified in interferonopathic diseases (Fig. 3.1). TREX1, a 3'-5' DNA exonuclease,

is mutated in SLE and AGS patients causing a loss of exonuclease activity and increase in cellular ssDNA and dsDNA (Crow et al., 2006a; Fye et al., 2011). The accumulation of DNA results in chronic activation of DNA damage checkpoint signaling (Crow et al., 2006a; Fye et al., 2011; Yang et al., 2007). Mutations in the DNA endonucleases DNase I (Martinez Valle et al., 2008) and DNase II (Rodero et al., 2017) have been identified in patients. Reduced activity prevents clearing of chromatin breakdown byproducts and stimulates production of anti-DNA nucleoprotein antibodies. SAMHD1, a triphosphohydrolase, normally degrades deoxyribonucleoside triphosphates (dNTPs) (Kretschmer et al., 2015). Mutations in SAMHD1 in AGS patients result in elevated cellular dNTP concentrations that trigger cellular stress responses and senescence (Kretschmer et al., 2015; Rice et al., 2009). The RNA endonuclease RNASE H2 is essential for removing ribonucleotides from the genome and degrading RNA:DNA hybrids (Crow et al., 2006b; Hiller et al., 2012). Loss of function mutations in RNASE H2 in patients with AGS and SLE (Crow et al., 2006b; Gunther et al., 2015) result in chronic DNA damage signaling and inflammatory signaling through the DNA sensor cGAS and its adaptor STING (Hiller et al., 2012; Mackenzie et al., 2016). Mutations to the adenosine deaminase acting on RNA 1, ADAR1, are linked to AGS (Rice et al., 2012). The mutant ADAR1 loses its ability to modify cellular dsRNAs such as retroelements which normally prevents immune recognition (Liddicoat et al., 2016; Nishikura, 2010). Instead, accumulated unmodified endogenous dsRNAs trigger antiviral signaling and IFN production through the RLR pathway (Zhao et al., 2018). Increased expression of endogenous retroelements such as LINE1 is also thought to stimulate IFN production

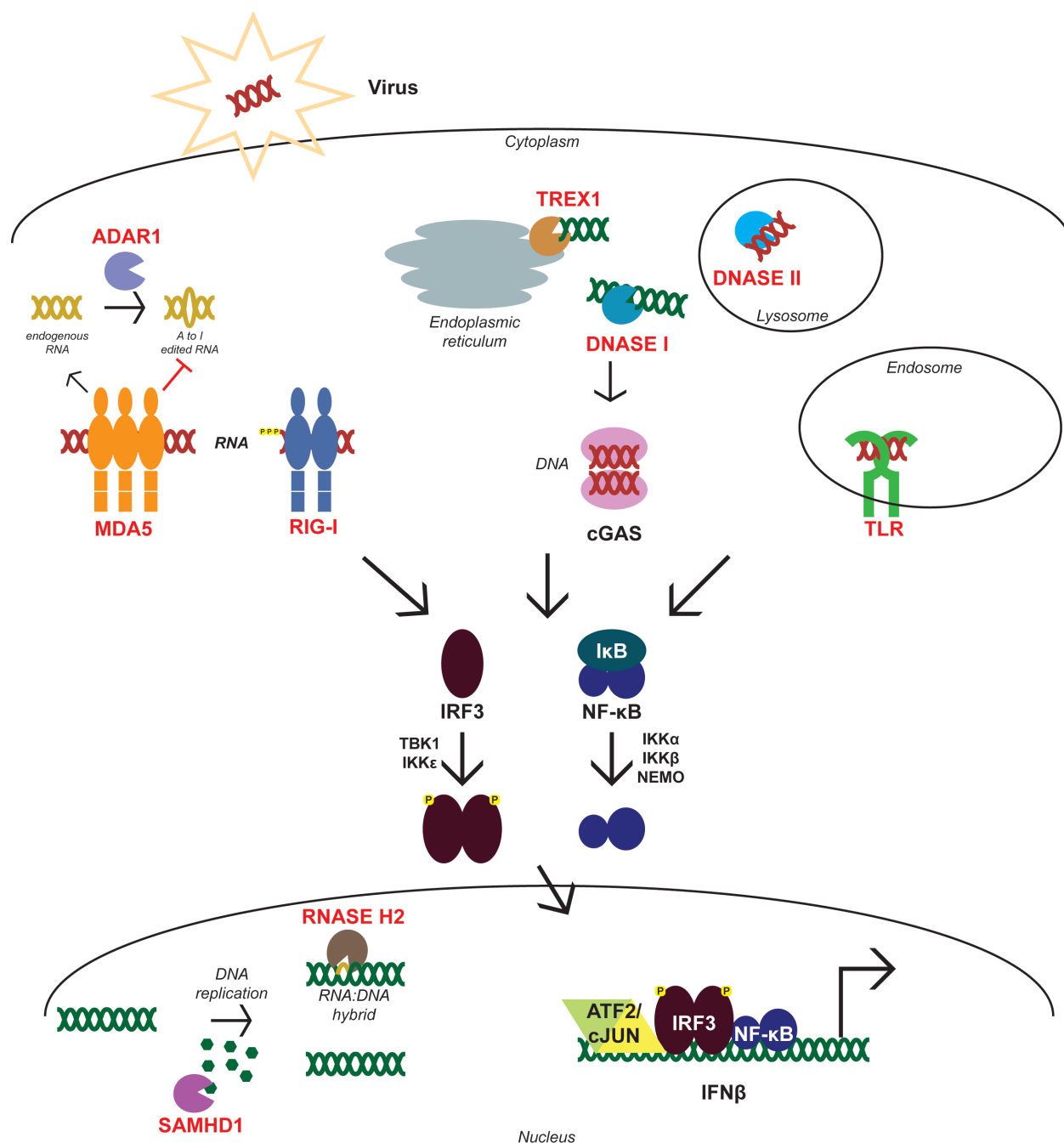


Figure 3.1 – Cellular machinery implicated in type I interferonopathies

Several enzymes are responsible for preventing accumulation of cellular DNA (TREX1, Dnase I, Dnase II, SAMHD1) and editing nucleic acids (ADAR1, Rnase H2) to prevent activation of the PRRs under homeostatic conditions. Mutations in any of these enzymes or the PRRs themselves (all marked in red) induce inappropriate expression of type I IFNs such as IFN β , causing type I interferonopathies in patients.

(Mavragani et al., 2016). LINE1 expression and activity are regulated by pathways involving ADAR1, TREX1 and SAMHD1 (Choi et al., 2018; Orecchini et al., 2017; Zhao et al., 2013; Zhao et al., 2018).

In addition to defects that contribute to imbalances in endogenous nucleic acid accumulation and processing, nucleic acid sensing PRRs themselves have also been implicated in type I interferonopathies (Kato et al., 2017) (Fig. 3.1). Analysis of SNPs and genetic linkage have implicated TLRs (Laska et al., 2014; Wang et al., 2014; Wu et al., 2015b), and gain of function mutations were identified in RIG-I (Jang et al., 2015). Mutations to MDA5 have been identified repeatedly in patients, and gain of MDA5 function is not only more prevalent, but has also been mechanistically characterized (Funabiki et al., 2014; Oda et al., 2014; Rice et al., 2014; Rutsch et al., 2015; Van Eyck et al., 2015).

MDA5 and its role in chronic interferon production

MDA5 is comprised of a central conserved DECH-box helicase region, an N-terminal caspase activation and recruitment domain (CARD) region, a pincer domain and a C-terminal domain (Fig. 3.2A) (Cordin et al., 2006; Wu et al., 2013a). The helicase domain is divided into domain 1 (Hel1) containing the conserved motifs Q, I, II and III, the insertion domain (Hel2i), and domain 2 (Hel2) containing the conserved motifs IV, V, and VI (Cordin et al., 2006; Wu et al., 2013a). The helicase domain has intrinsic RNA binding and ATP hydrolysis activities, enabling MDA5 to recognize non-self dsRNAs and form oligomers in an ATP-dependent process (Bruns et al., 2014; Peisley et al., 2012; Peisley et al., 2011;

Wu et al., 2013a). Upon oligomerization the CARD region contacts the downstream antiviral signaling protein, MAVS, to induce the activation of IRF3 and NF κ B transcription factors that lead to the production of IFN.

In Type I interferonopathies, including SLE and AGS, several mutations have been identified within the helicase domain of MDA5 (Funabiki et al., 2014; Oda et al., 2014; Rice et al., 2014; Rutsch et al., 2015; Van Eyck et al., 2015). These mutations result in more stable dsRNA binding and oligomerization that leads to both constitutive and hyperactive IFN production and signaling activity (Rice et al., 2014). The interferonopathic MDA5 mutations may thereby result in a loss of tolerance to endogenous retroelements and stimulate constitutive IFN production (Ahmad et al., 2018). Most patient-isolated MDA5 variants retain ATP hydrolysis activity, but at least one mutation, R337G, leads to a decrease in ATP hydrolysis (Rice et al., 2014). These properties resemble those of constitutively active MDA5 (CA-MDA5) variants generated in the laboratory during systematic investigations of RNA binding, ATP hydrolysis, and IFN stimulation by all the RLRs (Bamming and Horvath, 2009). Specific alterations to MDA5 helicase domain motif I (MI) and motif III (MIII) were found that lead to constitutive and hyperactive IFN production despite the loss of ATP hydrolysis.

Paramyxovirus evasion of antiviral signaling

In the course of their evolution, RNA viruses have developed a number of strategies to escape, avoid, or antagonize host immune surveillance and antiviral response (Gerlier

and Lyles, 2011; Gotoh et al., 2002). In many cases, viral strategies have been described that impinge on IFN production or responses. One large family of enveloped negative-strand RNA viruses, the Paramyxoviridae, is known to interfere with both IFN production and IFN signaling (Ramachandran and Horvath, 2009; Versteeg and Garcia-Sastre, 2010). For many viruses in this family, IFN evasion and antagonism is mediated in part or whole by non-structural proteins known as V proteins. The Paramyxovirus V proteins are recognizable by the presence of a highly conserved cysteine-rich C-terminal zinc-binding finger domain that forms a structural module that is linked to suppression of IFN production. V proteins use this conserved domain to directly interfere with cellular MDA5. Despite high amino acid sequence conservation of the RLRs within a defined minimal V protein binding region (MVBR), V proteins show specificity in their ability to bind to and antagonize MDA5 and its associate, LGP2, but they do not bind to the analogous region of RIG-I (Parisien et al., 2009; Rodriguez and Horvath, 2013, 2014). Evidence regarding the mechanism of V protein-mediated MDA5 suppression suggests that V protein interferes with MDA5 dsRNA interaction (Childs et al., 2009) and prevents MDA5 ATP hydrolysis (Parisien et al., 2009).

This specific MDA5 targeting capacity of Paramyxovirus V proteins suggested a hypothesis that they could be used to antagonize or suppress interferonopathic MDA5 signaling. To test this concept, both laboratory and patient derived CA-MDA5 proteins were generated, characterized for IFN and ISG expression, and assessed for sensitivity to V proteins from PIV5, mumps virus, measles virus, Nipah virus and Hendra virus.

Results indicate that Paramyxovirus V proteins are effective inhibitors of CA-MDA5 proteins that can significantly suppress interferonopathic activity.

RESULTS

Construction and Characterization of Constitutively Active (CA)-MDA5 Proteins

MDA5 mutations that confer hyperactive signaling profiles have been identified in patients with interferonopathic diseases, as well as in laboratory experiments designed to characterize the MDA5 catalytic domains. Seven constitutively active MDA5 (CA-MDA5) mutations were generated using a well-characterized N-terminal FLAG epitope tagged wild type (WT) human MDA5 cDNA as the template for site directed mutagenesis (Fig. 3.2A) (Bamming and Horvath, 2009; Bruns et al., 2014; Rodriguez and Horvath, 2013). Specifically, five of these CA-MDA5 mutants (R337G, D393V, G495R, R720Q, and R779C), were identified in patients characterized with an IFN signature (Oda et al., 2014; Rice et al., 2014). All five of the patient-derived CA-MDA5 proteins can induce the IFN β promoter reporter gene in the absence of virus infection or dsRNA stimulation, and each bind RNA, but one of the mutations, R337G, is defective for ATP hydrolysis activity. The other two CA-MDA5 mutations, a point mutation to helicase domain MI (K335A) or a double mutation to helicase domain MIII (T488A/S490A) are defective in ATP hydrolysis activity, yet are able to strongly induce IFN gene expression in human and murine cell lines in the absence of virus infection or dsRNA stimulation (Bamming and Horvath, 2009).

To establish a baseline to compare the activities of the mutant MDA5 proteins, their expression and activity were compared to WT MDA5 following transfection into HEK293T cells. All of the mutant proteins accumulated to similar levels as the expressed WT MDA5, as demonstrated in immunoblotting experiments with antibodies for MDA5 as well as FLAG epitope tag (Fig. 3.2B).

The signaling capacity of the MDA5 proteins was evaluated using an IFN β promoter-driven reporter gene assay. Protein expression vectors were co-transfected into cells along with the IFN β firefly luciferase reporter gene and a normalization control Renilla luciferase. Expression of WT MDA5 stimulated reporter gene expression, giving rise to a 77-fold activation. Strikingly, all of the MDA5 mutations conferred hyperactive IFN β promoter activity, ranging from 7.5-fold increase over WT MDA5 for the most active mutant (MI) to 2.6-fold increase over WT MDA5 for the least active mutant (D393V) (Fig. 3.2C). The relative hyperactivity of these MDA5 proteins is highly reproducible and establishes a rank-order of constitutive activity: MI \cong MIII \gg R337G \cong R779C $>$ D393V \cong G495R \cong R720Q \gg WT.

Additional characterization of these mutant MDA5 proteins was obtained from analysis of endogenous IFN β gene expression. Total RNA was prepared from WT and CA-MDA5 expressing cells and IFN β mRNA was measured by RT-qPCR (Fig. 3.3). Transfection efficiency in these samples was assessed to be greater than 90% by GFP expression in parallel samples. MDA5 expression activates the expression of IFN β , leading to a 5-fold

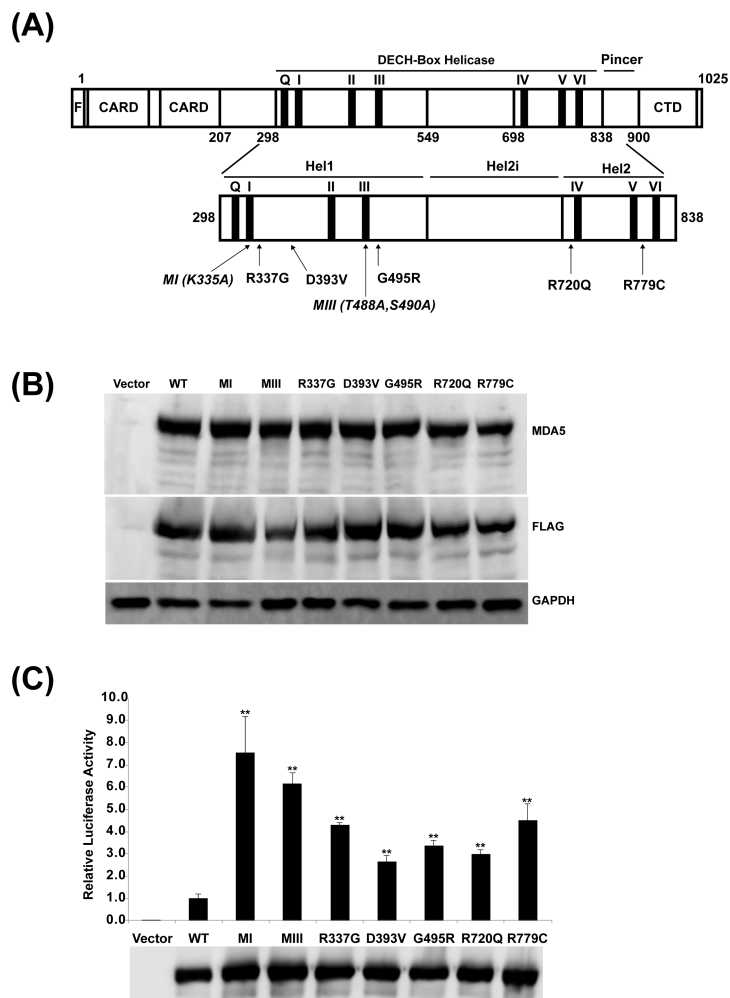


Figure 3.2 – Construction and characterization of mutant MDA5 proteins

(A) A diagram illustrating MDA5 domain structure including FLAG (F) epitope tag, CARDS, DECH-box helicase domain, pincer domain and C-terminal domain (CTD). The MDA5 minimal V protein binding region coincides with the Hel2 domain within the helicase domain. The expanded view of the helicase domain below shows the relative locations of the naturally arising activating mutations (R337G, D393V, G495R, R720Q, and R779C), as well as the laboratory generated mutations (italicized, MI (K335A) and MIII (T488A, S490A)).

(B) HEK293T cells were transfected with empty vector, WT or CA-MDA5 and assayed by immunoblotting with antibodies for MDA5, FLAG epitope tag and GAPDH.

(C) HEK293T cells were transfected with empty vector, WT or CA-MDA5 as indicated along with an IFN β promoter-driven luciferase reporter and a Renilla luciferase vector for normalization. After 24 hours cell were lysed and assayed for luciferase activity. Data are reported relative to WT MDA5 signal normalized to 1. Bars indicate average values (n=3) with error bars representing standard deviation. Statistical analysis was done using a two-tailed Student's t-test. (**p-value <0.005) Immunoblot analysis indicates similar expression between WT MDA5 and CA-MDA5.

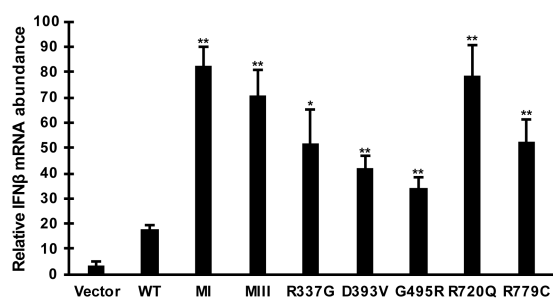


Figure 3.3 – CA-MDA5 mutants increase endogenous IFN β mRNA expression

Control empty vector, WT or CA-MDA5 were expressed in HEK293T cells. Total RNA was harvested and analyzed by RT-qPCR. Data are representative of ≥ 3 replicate experiments and are shown normalized to GAPDH expression. Bars indicate average values of technical replicates (n=3) with error bars representing standard deviation. Statistical analysis was done using a two-tailed Student's t-test (* p-value <0.05, ** p-value <0.005).

stimulation over cells without MDA5 expression. All of the CA-MDA5 proteins exhibited hyperactivity in endogenous gene induction, ranging from 4.6-fold over WT MDA5 for the MI mutant to 1.9-fold for the G495R mutant, which was equal to the D393V mutant in exhibiting the lowest level of hyperactivity in IFN β mRNA stimulation. Unexpectedly, the differences among the CA-MDA5 proteins in endogenous gene regulation were not as dramatic or variable compared to the reporter gene assay. This was most dramatically for R720Q, possibly indicating the influence of endogenous cellular mechanisms that function to attenuate or suppress the IFN β response.

V Proteins Bind to CA-MDA5 Proteins

Paramyxovirus V protein suppression of MDA5 relies on direct protein-protein interaction through contact with the helicase domain C-terminal lobe (Andrejeva et al., 2004; Parisien et al., 2009). To test the ability of V proteins to associate with CA-MDA5 proteins, V proteins from diverse Paramyxoviruses were used in co-immunoprecipitation assays. Specifically, five V proteins from the Rubulaviruses, PIV5 and mumps virus; the Morbilliviruses, measles virus; and the Henipaviruses, Nipah virus and Hendra virus, were expressed in cells from vectors containing in-frame N-terminal HA epitope tags along with each of the FLAG epitope-tagged MDA5 proteins (Ramachandran et al., 2008; Rodriguez et al., 2002; Rodriguez et al., 2003; Ulane et al., 2005). Whole cell lysates were prepared and subjected to FLAG/M2 antibody immunoaffinity purification followed by immunoblotting with HA antibody to detect co-precipitation. All five of the tested V proteins

were found to co-precipitate with all of the MDA5 proteins irrespective of their mutations (Fig. 3.4).

V Proteins Suppress CA-MDA5 Signal Transduction

To confirm that the co-precipitation results in MDA5 suppression, signaling activity was measured with the IFN β promoter-luciferase reporter gene assay. WT or CA-MDA5 proteins were co-expressed with the Paramyxovirus V proteins and luciferase activity determined. In all cases, MDA5 expression was able to activate the reporter gene, with similar relative potency as noted in Figure 3.2C, and this MDA5-mediated transcriptional activity was decreased by the presence of V proteins (Fig. 3.5). A specific order of inhibitory impact was found, with the PIV5 and Nipah virus V proteins most effective at inhibiting MDA5, irrespective of the activating mutation. Measles virus V protein was equally effective as PIV5 and Nipah virus V protein in suppressing WT MDA5 (Fig. 3.5A), as well as CA-MDA5 proteins MI, MIII, and R337G (Fig. 3.5B-D) but was less effective against D393V, G495R, R720Q, and R779C (Fig. 3.5E-H). Hendra virus and Mumps virus V proteins were the poorest inhibitors in this assay, but still mediated significant degree of interference with all the CA-MDA5 proteins.

PIV5 V Protein Suppresses CA-MDA5 Mediated IFN β Gene Expression

Of the V proteins tested, the PIV5 V protein was found to be uniformly effective at suppressing CA-MDA5 proteins in the reporter assay. To verify the observed suppression of CA-MDA5 protein activity, the ability of PIV5 V protein to interfere with MDA5 induction

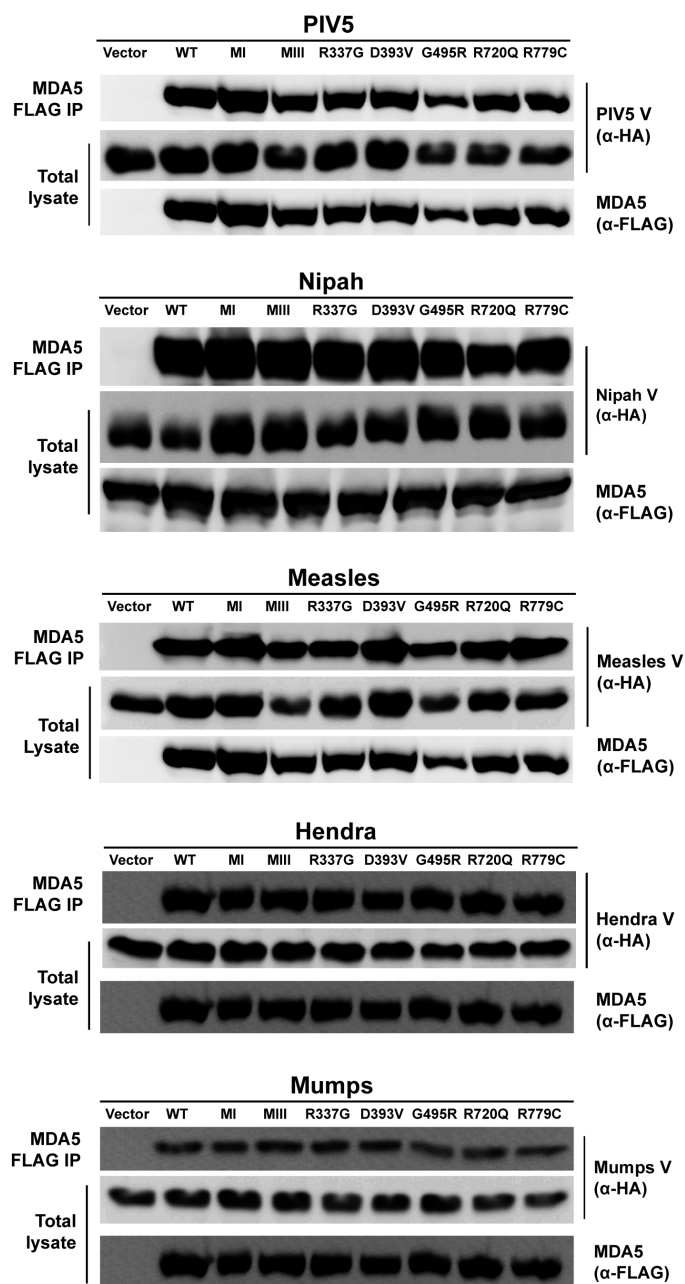


Figure 3.4 – V Proteins bind CA-MDA5 mutants

HA-epitope tagged V protein from PIV5, Nipah virus, measles virus, Hendra virus or mumps virus was co-expressed with FLAG-epitope tagged WT or CA-MDA5 or the empty vector as indicated. Cell lysates were subjected to FLAG immunoprecipitation and assayed by immunoblot with antibody for the HA epitope tag to determine if the V protein co-immunoprecipitated with MDA5.

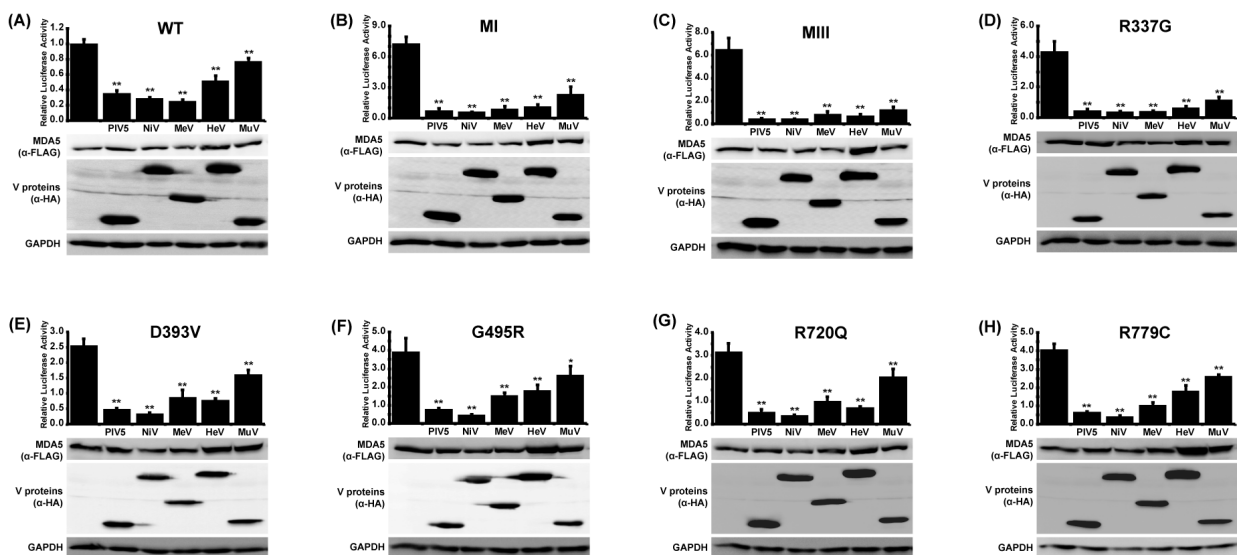


Figure 3.5 – V proteins suppress CA-MDA5 mutant signaling

(A) WT or (B-H) CA-MDA5 were co-transfected with an IFN β promoter-driven luciferase reporter and a control Renilla luciferase in the presence and absence of PIV5, Nipah virus (NiV), measles virus (MeV), Hendra virus (HeV) or mumps virus (MuV) V protein. After 24 hours cells were lysed and assayed for luciferase activity. Data are reported relative to WT MDA5 signal in the absence of V protein normalized to 1. Bars indicate average values ($n \geq 3$) with error bars representing standard deviation. Statistical analysis was done using a two-tailed Student's t-test. (*p-value < 0.05, ** p-value < 0.005) Corresponding immunoblots of cell lysates assayed with anti-FLAG antibody for MDA5 expression and anti-HA antibody for V protein expression show similar protein expression for all samples.

of endogenous IFN β gene expression was tested. PIV5 V protein was co-expressed along with WT or CA-MDA5, and total RNA was subjected to RT-qPCR analysis of IFN β mRNA. The PIV5 V protein interfered with the activity of all the CA-MDA5 proteins as indicated by greatly reduced IFN β gene expression (Fig. 3.6).

PIV5 V Protein Suppresses Interferonopathic Gene Expression

Patients with type I interferonopathies are characterized by elevated antiviral gene expression signatures even in the absence of infection, as diagnosed by small panels of ISG expression (Baechler et al., 2003; Bennett et al., 2003). To test the impact of V proteins on ISG responses, a panel of ISGs was measured in cells expressing WT or CA-MDA5 proteins alone and in combination with PIV5 V protein. Following protein expression, RNA was prepared for RT-qPCR using primers for the well-known ISGs CCL5, RSAD2, IFIT1, IFI6, and ISG15 (Schoggins et al., 2011). All ISGs tested were activated by WT or CA-MDA5 expression, with similar extent of hyperactivity as observed for IFN β mRNA induction. The ISG activation induced by all MDA5 proteins was suppressed robustly by co-expressed PIV5 V protein (Fig. 3.7).

PIV5 V Protein Suppresses Hyperactive Antiviral Response

To further validate the suppression of CA-MDA5, a biological endpoint assay was tested. MDA5-mediated activation of IFN production is a first step in establishing the cellular antiviral state, and this can be evaluated using an antiviral cytopathic effect (CPE) assay (Bamming and Horvath, 2009). Conditioned media was harvested from cells

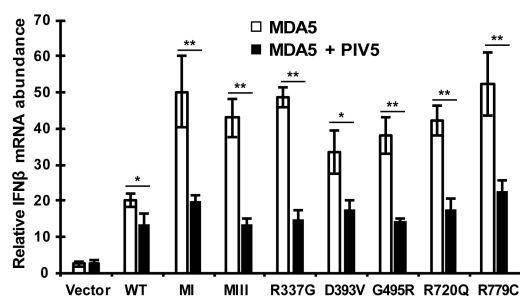


Figure 3.6 – PIV5 V protein suppresses CA-MDA5-induced IFN β expression

Empty vector, WT or CA-MDA5 were expressed in HEK293T cells in the presence and absence of PIV5 V protein. Total RNA was collected and analyzed for IFN β mRNA expression by RT-qPCR. Data are representative of ≥ 3 replicate experiments and are shown normalized to GAPDH expression. Bars indicate average values of technical replicates (n=3) with error bars representing standard deviation. Statistical analysis was done using a two-tailed Student's t-test (* p-value < 0.05, ** p-value < 0.005).

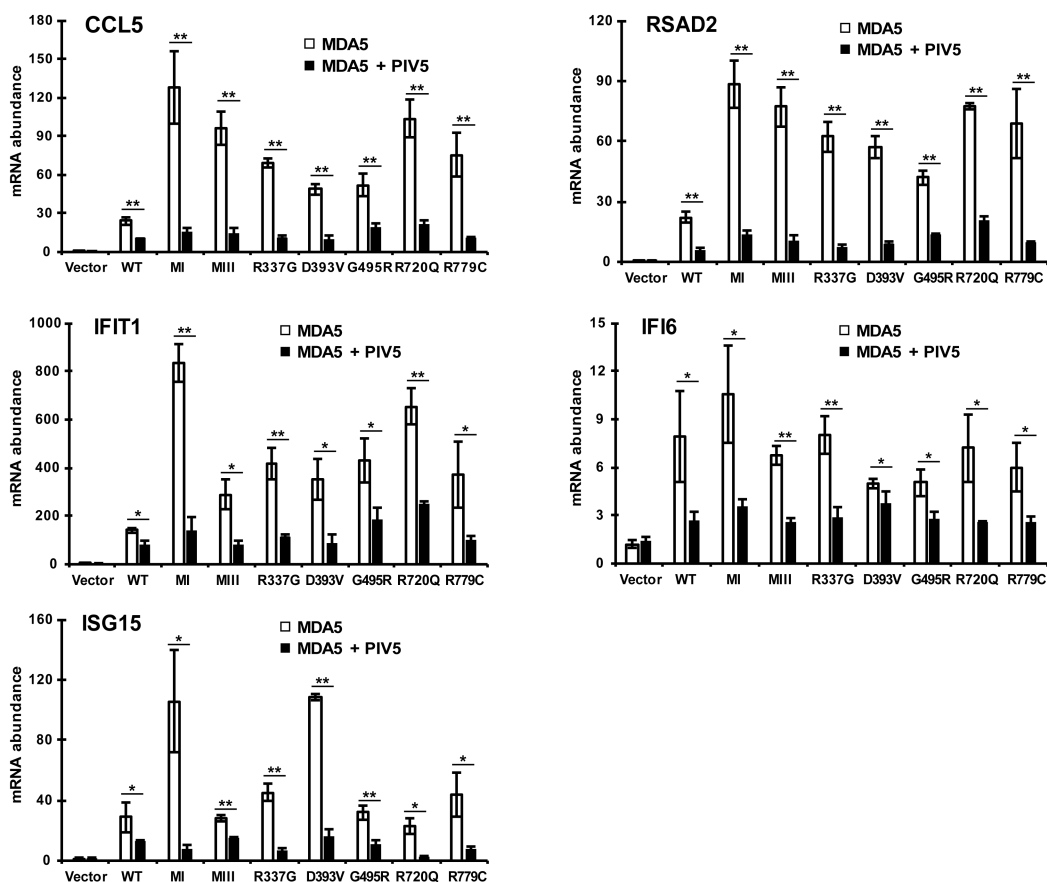


Figure 3.7 – PIV5 V protein suppresses CA-MDA5-induced ISG expression

Empty vector, WT or CA-MDA5 were expressed in HEK293T cells in the presence and absence of PIV5 V protein. Total RNA was collected and analyzed for CCL5, RSAD2, IFIT1, IFI6 and ISG15 mRNA expression by RT-qPCR. Data are representative of ≥ 3 replicate experiments and are shown normalized to GAPDH expression. Bars indicate average values of technical replicates ($n=3$) with error bars representing standard deviation. Statistical analysis was done using a two-tailed Student's t-test (* p-value < 0.05, ** p-value < 0.005).

expressing WT or CA-MDA5 proteins both in the presence and in the absence of PIV5 V protein, then applied to freshly-plated 2fTGH cells for 8 hours. These cells were then inoculated with VSV for 18 hours. Media from the MDA5 expressing cells, containing secreted IFN, protected the 2fTGH cells from VSV-induced cell death, and the CA-MDA5 mutants induced greater antiviral protection than WT MDA5 with similar rank order efficacy as observed in prior assays (Fig. 3.8). Comparison to a standard titration of IFN α -mediated CPE protection under these conditions shows that WT MDA5 produced equivalent of 156.25 U/mL activity, while the most active CA-MDA5, MI, produced equivalent of 2500 U/mL of activity. The least protective CA-MDA5 D393V still produced 312.5 U/mL of activity. Co-expression of PIV5 V protein decreased protection from VSV-induced cytopathicity irrespective of the specific MDA5 mutation, consistent with the ability to interfere with MDA5 signal transduction.

DISCUSSION

Hosts and viruses have co-evolved, with hosts developing innate immune strategies to combat viruses and viruses developing mechanisms to evade and antagonize the host innate immune system. As we look to develop treatments for human immune diseases, we should exploit these natural antagonists to benefit human health. One particularly interesting viral protein is the Paramyxovirus V protein that has evolved to prevent IFN antiviral responses (Ramachandran and Horvath, 2009; Versteeg and Garcia-Sastre, 2010). The V proteins share a common target to inhibit IFN production, the antiviral sensor MDA5 (Childs et al., 2007; Ramachandran and Horvath, 2009). In this study, the ability

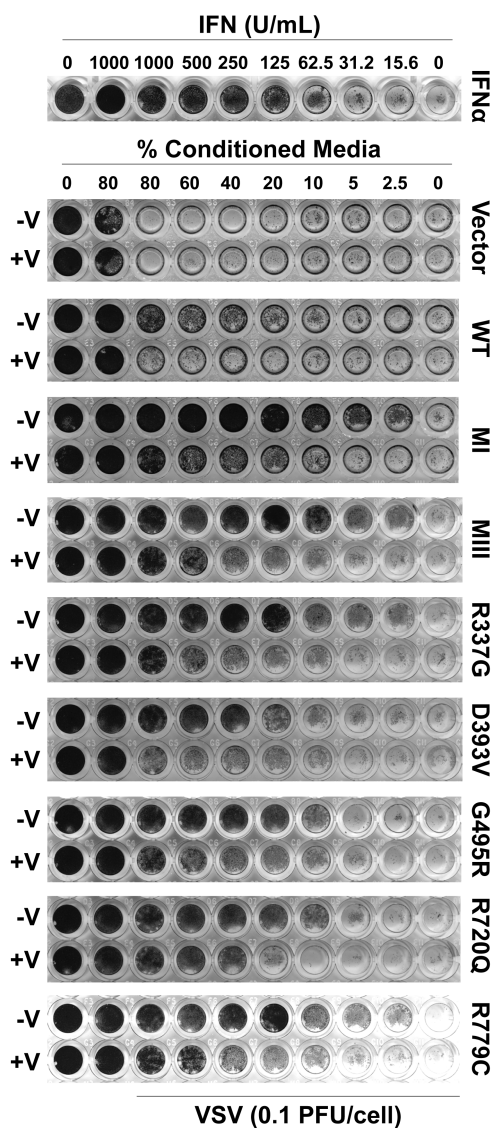


Figure 3.8 – PIV5 V protein suppresses antiviral response of CA-MDA5 mutants

Conditioned media from HEK293T cells expressing empty vector, WT or CA-MDA5 in the presence and absence of PIV5 V was collected and used to treat 2fTGH cells for 8 hours. Cells were infected with VSV for 18 hours, fixed and stained to assay protection from virus-induced cytopathicity. As a control, 2fTGH cells were directly treated with IFN α for 8 hours prior to virus infection.

of Paramyxovirus V proteins to antagonize constitutively active MDA5 variants was tested as a potential means to control chronic IFN production. Results indicate that diverse Paramyxovirus V proteins were able to interact with both patient-derived and laboratory-generated CA-MDA5 proteins and suppress their chronic signaling and antiviral activities. The V proteins not only reduced IFN and ISG expression levels, but they also suppressed heightened antiviral responses observed in cells harboring CA-MDA5. These results indicate that V proteins are an attractive system for developing new means to suppress MDA5-induced interferonopathic syndromes. As V proteins not only target MDA5-mediated IFN production, but also IFN-JAK-STAT signaling, they have the ability to counteract toxic effects of excessive IFN production and chronic IFN-stimulated gene expression (Ramachandran and Horvath, 2009). This feature of V proteins can be seen in the increased suppression of ISG expression by PIV5 V protein (Fig. 3.7) compared to the suppression of IFN β expression (Fig. 3.6). PIV5 V protein not only targets MDA5 but it also induces degradation of STAT1, likely leading to increased ISG suppression (Didcock et al., 1999).

Interestingly, the diverse V proteins exhibited differential inhibition of MDA5 signaling (Fig. 3.5). Though the V protein C-terminal domains are highly conserved, previous work has determined that there are universal and virus-specific requirements for residues in MDA5 engagement (Ramachandran and Horvath, 2010). PIV5 and Nipah virus V proteins do not require the first conserved histidine or the first conserved cysteine for MDA5 interference, while the mumps virus and measles virus V proteins require both of these

residues (Ramachandran and Horvath, 2010). These virus-specific requirements could indicate a difference in affinity for MDA5 and explain why PIV5 and Nipah virus V proteins are more potent inhibitors of MDA5 signaling than the V proteins of measles virus, Hendra virus and mumps virus.

In this study we used full-length V proteins to analyze their potential in antagonizing interferonopathic MDA5. Previous research has shown that the conserved V protein C-terminal domain is sufficient to interact with MDA5 and inhibit signaling (Andrejeva et al., 2004; Childs et al., 2007; Childs et al., 2009; Parisien et al., 2009). All the V proteins require the first conserved arginine in the C-terminal domain for MDA5 interaction and interference (Ramachandran and Horvath, 2010), suggesting that peptides surrounding this region may retain MDA5 suppression activity. Dissection of minimal inhibitory regions is expected to define short peptides for inhibition of CA-MDA5 to be used as a starting point for rational design of compounds and assays that could yield small molecule inhibitors of MDA5 signaling.

In addition to providing proof of concept for V protein interference with CA-MDA5, this study refines our understanding of the mechanisms underlying V protein inhibition of MDA5. It was previously shown that V protein can interfere with MDA5 dsRNA binding (Childs et al., 2009) as well as ATP hydrolysis (Parisien et al., 2009). Three CA-MDA5 proteins analyzed (MI, MIII, and R337G) are defective for ATP hydrolysis yet they are still suppressed by V proteins, suggesting that the V protein interference with ATP hydrolysis

is a consequence of interaction that may not be directly linked to signaling inhibition. Further work is required to address both the mechanistic basis for hyperactivity in CA-MDA5 proteins, and the roles that ATP hydrolysis plays in MDA5 signaling.

This study provides an initial proof of the concept that viral IFN evasion proteins could be harnessed as effective antagonists of type I interferonopathies. Moreover, such a therapeutic strategy would not be limited for use only in patients with CA-MDA5 mutations. Interferonopathies caused by ADAR1 mutations develop due to increased recognition of unmodified endogenous retroelements by WT MDA5 (Ahmad et al., 2018), and mutations in TREX1, SAMHD1 and RNase H2 lead to overexpression of endogenous LINE1 RNA that triggers signaling through WT MDA5 (Choi et al., 2018; Li et al., 2017; Zhao et al., 2013; Zhao et al., 2018). Even in these cases, targeting of MDA5 by V proteins would counteract overactive IFN and ISG expression.

CHAPTER IV: PERSPECTIVES

Though there have been significant advances in medicine over the years, there are still many established and emerging virus infections, and immune and autoimmune disorders that lack an effective and specific treatment. There are a few FDA-approved antiviral drugs against viruses such as influenza that have limited efficacy. However, the rise in globalization has resulted in an increase in pathogenic viruses infecting humans and creating a pressing need for the development of new antiviral drugs. A number of chronic viral infections and diseases such as multiple sclerosis are currently treated with type I interferon (IFN), even though type I IFN itself can drive autoimmune disorders and cause toxic side effects in patients (Gresser et al., 1980; Trinchieri, 2010). In the case of many other autoimmune disorders that currently have no treatment available, immunosuppressive drugs are used to at least limit the symptoms of the disease. This lack of effective treatments for viral infections, and immune and autoimmune disorders is in part due to the complexity of the innate immune system.

The innate immune system has co-evolved with viruses over thousands of years, creating an intricate signaling system that must survive infections from not only diverse viruses but other pathogens as well (Daugherty and Malik, 2012; Kosiol et al., 2008). Viruses are able to mutate, pressuring the innate immune system to adapt so it can successfully counteract the new mutated virus. This adaptation in turn forces the virus to evolve once again to mount a productive infection. This continued arms race has shaped modern viruses as well as the cellular innate immune system. Though viruses evolve on a much shorter time scale than humans, they are limited by the many layers of the innate immune system.

The complexity of the cell-intrinsic innate immune response enables humans to successfully compete against rapidly mutating viruses. The viruses survive only if they can evade or antagonize several layers of immune effectors.

Development of new specific and effective antiviral drugs, therapies for immune and autoimmune disorders and diagnostics for these diseases requires a more thorough understanding not only of the innate immune response but also of the mechanisms employed by viruses that can infect humans. Once more is known, the endogenous innate immune response could be enhanced to become potent antiviral drugs or used as markers of a disease. Similarly, virus-encoded inhibitors targeting the innate immune response could also be harnessed to design therapies for autoimmune diseases.

This thesis project aimed to further the insight about the response to virus infections from both of these perspectives. Though many factors involved in the innate immune system have been identified and characterized, recent advances in deep-sequencing approaches are revealing a major gap in the current understanding of the innate immune response (Djebali et al., 2012; Freaney et al., 2013). The results presented in this thesis demonstrate that there are thousands of previously-unannotated virus-regulated noncoding RNAs as well as annotated RNAs not previously recognized as virus-regulated that may have roles in antiviral immunity. Recognizing successful virus evasion strategies, we demonstrated a proof of the concept that viral IFN-evasion proteins could be used to suppress excessive immune signaling found in human disease. Together

these findings not only expand the current knowledge about the cellular innate immune response to viruses, but also provide insight into the therapeutic application of previously untapped regulators of the antiviral response.

ENDOGENOUS REGULATORS

The cell-intrinsic innate immune response is tightly controlled by various signaling components to ensure a rapid and transient response to viruses. Much of the current understanding of this response comes from studies using conventional molecular biology techniques and microarray technologies. Though these assays revealed many factors involved in the response, they could not facilitate the discovery of low-expressing *de novo* transcripts.

The use of RNA-sequencing to study virus-regulated host gene expression changes enabled the identification of thousands of previously-unannotated RNAs. In addition to detecting 3605 RefSeq-annotated RNAs, this analysis uncovered 2605 previously-unannotated RNAs that were differentially regulated by virus infection. These novel RNAs were primarily transcribed from intergenic and intronic genomic regions (Fig. 2.6E-H) and predicted to be noncoding RNAs (Fig. 2.7A,D), unlike the virus-regulated previously RefSeq-annotated RNAs which were primarily exonic mRNAs encoding proteins. The stark difference in classification between virus-regulated previously-annotated and previously-unannotated RNAs highlights how narrowly focused the analysis of the cellular innate immune response has been upto now.

These results demonstrate that noncoding RNAs are a poorly investigated component of the cellular innate immune response. The innate immune response is a highly-regulated system and the noncoding RNAs could exert the subtle regulation that contributes to differences in the induction and attenuation of immune genes. Further characterization of these RNAs is needed to determine their functions more specifically.

Bioinformatics analyses and approaches based on inducibility were employed to broadly analyze these previously unrecognized RNAs for function. Expression of a subset of the previously-unannotated RNAs in five different cell lines and their low sequence conservation among vertebrates suggests that they may have regulatory or direct antiviral roles (Fig. 2.8, 2.10, 2.11B-C, 2.12B-C, and Table 2.6, 2.7, 2.8). A consequence of the arms race between hosts and viruses is that the immune genes are some of the most recently and rapidly evolving mammalian genes and therefore have poor sequence conservation among vertebrates. The newly identified RNAs could have low sequence conservation because they have been positively selected over time for a specific antiviral function. Though specific functions were not assigned to all the previously unrecognized RNAs, two noncoding RNAs were further characterized.

In addition to the previously-unannotated RNAs, this RNA-sequencing analysis also discovered many previously-annotated RNAs that had not yet been characterized as virus-regulated or as having an antiviral function. The virus-induced lncRNAs AFF1-AS1 and ZBED5-AS1 were identified as negative regulators of the innate immune response

through this RNA-sequencing pipeline. Though both were found to suppress the expression of IFN β , they likely have different mechanisms of action.

Results suggest that AFF1-AS1 affects IFN β expression by negatively regulating an upstream process, perhaps directly regulating the activation of IRF3 (Fig. 2.16). AFF1-AS1 may prevent IRF3 phosphorylation by directly binding the transcription factor. At basal conditions, it may only modestly limit IRF3 availability for phosphorylation but when it is induced by virus infection, it may be an effective component of the IRF3 activation attenuation processes. The lncRNA NKILA has a similar negative feedback function. NKILA, which is induced by NF- κ B, binds to the I κ B α /NF- κ B complex and prevents phosphorylation of I κ B α and subsequent release of NF- κ B (Liu et al., 2015).

ZBED5-AS1, on the other hand, seems to directly regulate IFN β mRNA accumulation (Fig. 2.17), though further experimentation is required to determine if its negatively regulating the transcription or mRNA decay of IFN β . The lncRNA NRAV has been shown to affect histone modifications of target gene promoters (Ouyang et al., 2014) so it is possible that ZBED5-AS1 has a similar inhibitory mechanism that prevents RNA Pol II access to the IFN β gene. Alternately, ZBED5-AS1 could act as a scaffold that directs degradation machinery to its target transcripts.

Due to the large number of newly uncovered virus-regulated RNAs, assigning specific functions to all of them will require the development of high-throughput screens. A high-

throughput shRNA or CRISPR/Cas9 knockout screen followed by infection with a fluorescent virus would facilitate the identification of RNAs that inhibit or promote virus replication. However, there would be major challenges in such a study. Noncoding RNAs have low expression levels (Fig. 2.5) and can have cell-specific expression (Derrien et al., 2012), limiting the number of RNAs that could be successfully analyzed in a broad screen in a single cell type. Another layer of complexity would be that cellular antiviral effectors are not always broad-acting and must be tested against various possible targets (Schoggins et al., 2011). Overcoming these limitations requires screening the function of these RNAs in multiple cell lines against multiple viruses.

Further characterization of these previously unrecognized RNAs would yield an in-depth understanding of the complex regulation involved in keeping the innate immune response functioning in a timely and appropriate manner. This level of detail about the system would provide novel directions for drug development.

VIRUS-ENCODED REGULATORS

Modern viruses have evolved unique properties that enable them to effectively infect humans. Some of these properties are already exploited for the development of targeted therapies for immune and autoimmune disorders. For example, oncolytic viruses are used to specifically target and kill cancer cells in patients (Fukuhara et al., 2016). Viral vectors also allow for effective transgene delivery for gene therapy (Kay et al., 2001). However,

a property of viruses that has not been utilized yet in drug development is their specific targeting of innate immune system components.

The proof of concept study in this thesis demonstrates that Paramyxovirus V proteins can be used to design drugs that inhibit hyperactive signaling of the RNA sensor MDA5. This treatment strategy could prove to be a very potent therapeutic because 1) Paramyxovirus V proteins target both type I IFN production and signaling (Versteeg and Garcia-Sastre, 2010), treating both drivers of the disease and 2) targeting of MDA5 could help patients with defects in other components that still result in increased RNA accumulation and detection. These findings motivate the inquiry into other virus-encoded regulators that could aid in the treatment of immune and autoimmune diseases.

Viruses have evolved several methods to evade and antagonize the innate immune system at many different stages of the cellular response (Fig. 1.4). As research into immune and autoimmune disorders increasingly identifies specific disease-causing components, the viral strategies of targeting these components should be used as the basis of rational drug design. For instance, the mutations in RIG-I that have been identified in Singleton-Merten syndrome (Jang et al., 2015) could be targeted by drugs based on poliovirus protease 3C (Barral et al., 2009). The interferonopathic mutations in DNase I and II (Martinez Valle et al., 2008; Rodero et al., 2017) and TREX1 (Crow et al., 2006a; Fye et al., 2011), which likely result in accumulation of DNA detected by cGAS,

could be treated with a drug based on the cGAS-inhibiting KSHV ORF52 protein (Wu et al., 2015a).

The proof of concept study presented in this thesis was done using very ideal parameters. The patient-derived CA-MDA5 mutations investigated specifically cause monogenic autoimmune disorders and the interaction between V proteins and MDA5 has been extensively characterized (Parisien et al., 2009; Ramachandran and Horvath, 2010). Though there are viral proteins known to inhibit the activities of other innate immune signaling components that cause monogenic diseases, the mechanism of targeting must be further characterized to determine if the viral protein can be successfully used for drug design. The bigger challenge at this time to developing specific treatments for many autoimmune disorders is that the genetic component of many diseases has not been identified. However, as more information is elucidated about the causes of diseases, hopefully the strategy of using virus-encoded inhibitors as the basis of drug design will be utilized more. Even if major autoimmune disorders are found not to be monogenic, the drugs designed from virus-encoded proteins could at least be a part of a combination therapy.

CONCLUSIONS

Through the trial and error of evolution, powerful signaling systems have emerged in the human innate immune system and in viruses. These should be used as inspiration in drug design as they have already been optimized through positive selection cycles. The

increase in high-throughput and deep sequencing technologies is not only facilitating the identification of the depth and breadth of the innate immune system, but also the discovery of disease-causing mutations. This increase in knowledge should enable the development of treatments that target the causes of disease, rather than just the symptoms. The results presented in this thesis expand the current understanding about the extent of virus-regulated gene expression changes and highlight unrecognized factors that could potentially be used for the development of future diagnostics and therapeutics.

REFERENCES

Ablasser, A., Goldeck, M., Cavlar, T., Deimling, T., Witte, G., Rohl, I., Hopfner, K.P., Ludwig, J., and Hornung, V. (2013). cGAS produces a 2'-5'-linked cyclic dinucleotide second messenger that activates STING. *Nature* 498, 380-384.

Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Cech, M., Chilton, J., Clements, D., Coraor, N., Eberhard, C., *et al.* (2016). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res* 44, W3-W10.

Agalioti, T., Lomvardas, S., Parekh, B., Yie, J., Maniatis, T., and Thanos, D. (2000). Ordered recruitment of chromatin modifying and general transcription factors to the IFN-beta promoter. *Cell* 103, 667-678.

Ahmad, S., Mu, X., Yang, F., Greenwald, E., Park, J.W., Jacob, E., Zhang, C.Z., and Hur, S. (2018). Breaching Self-Tolerance to Alu Duplex RNA Underlies MDA5-Mediated Inflammation. *Cell* 172, 797-810 e713.

Akira, S., Uematsu, S., and Takeuchi, O. (2006). Pathogen recognition and innate immunity. *Cell* 124, 783-801.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166-169.

Andrejeva, J., Childs, K.S., Young, D.F., Carlos, T.S., Stock, N., Goodbourn, S., and Randall, R.E. (2004). The V proteins of paramyxoviruses bind the IFN-inducible RNA helicase, mda-5, and inhibit its activation of the IFN-beta promoter. *Proc Natl Acad Sci U S A* 101, 17264-17269.

Ank, N., West, H., and Paludan, S.R. (2006). IFN-lambda: novel antiviral cytokines. *J Interferon Cytokine Res* 26, 373-379.

Azmi, S., Ozog, A., and Taneja, R. (2004). Sharp-1/DEC2 inhibits skeletal muscle differentiation through repression of myogenic transcription factors. *J Biol Chem* 279, 52643-52652.

Baechler, E.C., Batliwalla, F.M., Karypis, G., Gaffney, P.M., Ortmann, W.A., Espe, K.J., Shark, K.B., Grande, W.J., Hughes, K.M., Kapur, V., *et al.* (2003). Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. *Proc Natl Acad Sci U S A* 100, 2610-2615.

Baeuerle, P.A., and Henkel, T. (1994). Function and activation of NF-kappa B in the immune system. *Annu Rev Immunol* 12, 141-179.

Bamming, D., and Horvath, C.M. (2009). Regulation of signal transduction by enzymatically inactive antiviral RNA helicase proteins MDA5, RIG-I, and LGP2. *J Biol Chem* 284, 9700-9712.

- Bar-Sagi, D., and Hall, A. (2000). Ras and Rho GTPases: a family reunion. *Cell* 103, 227-238.
- Barber, G.N. (2001). Host defense, viruses and apoptosis. *Cell Death Differ* 8, 113-126.
- Barral, P.M., Sarkar, D., Fisher, P.B., and Racaniello, V.R. (2009). RIG-I is cleaved during picornavirus infection. *Virology* 391, 171-176.
- Barriocanal, M., Carnero, E., Segura, V., and Fortes, P. (2014). Long Non-Coding RNA BST2/BISPR is Induced by IFN and Regulates the Expression of the Antiviral Factor Tetherin. *Front Immunol* 5, 655.
- Baum, A., and Garcia-Sastre, A. (2010). Induction of type I interferon by RNA viruses: cellular receptors and their substrates. *Amino Acids* 38, 1283-1299.
- Beachboard, D.C., and Horner, S.M. (2016). Innate immune evasion strategies of DNA and RNA viruses. *Curr Opin Microbiol* 32, 113-119.
- Bennett, L., Palucka, A.K., Arce, E., Cantrell, V., Borvak, J., Banchereau, J., and Pascual, V. (2003). Interferon and granulopoiesis signatures in systemic lupus erythematosus blood. *J Exp Med* 197, 711-723.
- Blomstrom, D.C., Fahey, D., Kutny, R., Korant, B.D., and Knight, E., Jr. (1986). Molecular characterization of the interferon-induced 15-kDa protein. Molecular cloning and nucleotide and amino acid sequence. *J Biol Chem* 261, 8811-8816.
- Bohn, W., Rutter, G., Hohenberg, H., Mannweiler, K., and Nobis, P. (1986). Involvement of actin filaments in budding of measles virus: studies on cytoskeletons of infected cells. *Virology* 149, 91-106.
- Bowie, A.G., and Unterholzner, L. (2008). Viral evasion and subversion of pattern-recognition receptor signalling. *Nat Rev Immunol* 8, 911-922.
- Brady, G., Haas, D.A., Farrell, P.J., Pichlmair, A., and Bowie, A.G. (2015). Poxvirus Protein MC132 from *Molluscum Contagiosum Virus* Inhibits NF- κ B Activation by Targeting p65 for Degradation. *J Virol* 89, 8406-8415.
- Bruns, A.M., and Horvath, C.M. (2014). Antiviral RNA recognition and assembly by RLR family innate immune sensors. *Cytokine Growth Factor Rev* 25, 507-512.
- Bruns, A.M., Leser, G.P., Lamb, R.A., and Horvath, C.M. (2014). The innate immune sensor LGP2 activates antiviral signaling by regulating MDA5-RNA interaction and filament assembly. *Mol Cell* 55, 771-781.
- Chakrabarti, A., Jha, B.K., and Silverman, R.H. (2011). New insights into the role of RNase L in innate immunity. *J Interferon Cytokine Res* 31, 49-57.

Chan, O., Burke, J.D., Gao, D.F., and Fish, E.N. (2012). The chemokine CCL5 regulates glucose uptake and AMP kinase signaling in activated T cells to facilitate chemotaxis. *J Biol Chem* 287, 29406-29416.

Childs, K., Stock, N., Ross, C., Andrejeva, J., Hilton, L., Skinner, M., Randall, R., and Goodbourn, S. (2007). mda-5, but not RIG-I, is a common target for paramyxovirus V proteins. *Virology* 359, 190-200.

Childs, K.S., Andrejeva, J., Randall, R.E., and Goodbourn, S. (2009). Mechanism of mda-5 Inhibition by paramyxovirus V proteins. *J Virol* 83, 1465-1473.

Choi, J., Hwang, S.Y., and Ahn, K. (2018). Interplay between RNASEH2 and MOV10 controls LINE-1 retrotransposition. *Nucleic Acids Res* 46, 1912-1926.

Coccia, E.M., Uze, G., and Pellegrini, S. (2006). Negative regulation of type I interferon signaling: facts and mechanisms. *Cell Mol Biol (Noisy-le-grand)* 52, 77-87.

Colonna, R.J., and Pang, R.H. (1982). Induction of unique mRNAs by human interferons. *J Biol Chem* 257, 9234-9237.

Cordin, O., Banroques, J., Tanner, N.K., and Linder, P. (2006). The DEAD-box protein family of RNA helicases. *Gene* 367, 17-37.

Crow, M.K., and Wohlgemuth, J. (2003). Microarray analysis of gene expression in lupus. *Arthritis Res Ther* 5, 279-287.

Crow, Y.J. (2011). Type I interferonopathies: a novel set of inborn errors of immunity. *Ann N Y Acad Sci* 1238, 91-98.

Crow, Y.J., Chase, D.S., Lowenstein Schmidt, J., Szykiewicz, M., Forte, G.M., Gornall, H.L., Oojageer, A., Anderson, B., Pizzino, A., Helman, G., *et al.* (2015). Characterization of human disease phenotypes associated with mutations in TREX1, RNASEH2A, RNASEH2B, RNASEH2C, SAMHD1, ADAR, and IFIH1. *Am J Med Genet A* 167A, 296-312.

Crow, Y.J., Hayward, B.E., Parmar, R., Robins, P., Leitch, A., Ali, M., Black, D.N., van Bokhoven, H., Brunner, H.G., Hamel, B.C., *et al.* (2006a). Mutations in the gene encoding the 3'-5' DNA exonuclease TREX1 cause Aicardi-Goutieres syndrome at the AGS1 locus. *Nat Genet* 38, 917-920.

Crow, Y.J., Leitch, A., Hayward, B.E., Garner, A., Parmar, R., Griffith, E., Ali, M., Semple, C., Aicardi, J., Babul-Hirji, R., *et al.* (2006b). Mutations in genes encoding ribonuclease H2 subunits cause Aicardi-Goutieres syndrome and mimic congenital viral brain infection. *Nat Genet* 38, 910-916.

Cudmore, S., Cossart, P., Griffiths, G., and Way, M. (1995). Actin-based motility of vaccinia virus. *Nature* 378, 636-638.

Darnell, J.E., Jr., Kerr, I.M., and Stark, G.R. (1994). Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science* 264, 1415-1421.

Dastur, A., Beaudenon, S., Kelley, M., Krug, R.M., and Huibregtse, J.M. (2006). Herc5, an interferon-induced HECT E3 enzyme, is required for conjugation of ISG15 in human cells. *J Biol Chem* 281, 4334-4338.

Daugherty, M.D., and Malik, H.S. (2012). Rules of engagement: molecular insights from host-virus arms races. *Annu Rev Genet* 46, 677-700.

de Weerd, N.A., and Nguyen, T. (2012). The interferons and their receptors--distribution and regulation. *Immunol Cell Biol* 90, 483-491.

Decker, C.J., and Parker, R. (1994). Mechanisms of mRNA degradation in eukaryotes. *Trends Biochem Sci* 19, 336-340.

Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., Guernec, G., Martin, D., Merkel, A., Knowles, D.G., *et al.* (2012). The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res* 22, 1775-1789.

Didcock, L., Young, D.F., Goodbourn, S., and Randall, R.E. (1999). The V protein of simian virus 5 inhibits interferon signalling by targeting STAT1 for proteasome-mediated degradation. *J Virol* 73, 9928-9933.

Diner, E.J., Burdette, D.L., Wilson, S.C., Monroe, K.M., Kellenberger, C.A., Hyodo, M., Hayakawa, Y., Hammond, M.C., and Vance, R.E. (2013). The innate immune DNA sensor cGAS produces a noncanonical cyclic dinucleotide that activates human STING. *Cell Rep* 3, 1355-1361.

Djebali, S., Davis, C.A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., *et al.* (2012). Landscape of transcription in human cells. *Nature* 489, 101-108.

Dong, J., Xu, S., Wang, J., Luo, R., Wang, D., Xiao, S., Fang, L., Chen, H., and Jiang, Y. (2015). Porcine reproductive and respiratory syndrome virus 3C protease cleaves the mitochondrial antiviral signalling complex to antagonize IFN-beta expression. *J Gen Virol* 96, 3049-3058.

Fearon, D.T., and Locksley, R.M. (1996). The instructive role of innate immunity in the acquired immune response. *Science* 272, 50-53.

Feng, Q., Langereis, M.A., Lork, M., Nguyen, M., Hato, S.V., Lanke, K., Emdad, L., Bhoopathi, P., Fisher, P.B., Lloyd, R.E., *et al.* (2014). Enterovirus 2Apro targets MDA5 and MAVS in infected cells. *J Virol* 88, 3369-3378.

Ferreira, R.C., Guo, H., Coulson, R.M., Smyth, D.J., Pekalski, M.L., Burren, O.S., Cutler, A.J., Doecke, J.D., Flint, S., McKinney, E.F., *et al.* (2014). A type I interferon transcriptional signature precedes autoimmunity in children genetically at risk for type 1 diabetes. *Diabetes* 63, 2538-2550.

Fitzgerald, K.A., and Caffrey, D.R. (2014). Long noncoding RNAs in innate and adaptive immunity. *Curr Opin Immunol* 26, 140-146.

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

Freaney, J.E., Zhang, Q., Yigit, E., Kim, R., Widom, J., Wang, J.P., and Horvath, C.M. (2014). High-density nucleosome occupancy map of human chromosome 9p21-22 reveals chromatin organization of the type I interferon gene cluster. *J Interferon Cytokine Res* 34, 676-685.

Friedman, R.L., Manly, S.P., McMahon, M., Kerr, I.M., and Stark, G.R. (1984). Transcriptional and posttranscriptional regulation of interferon-induced gene expression in human cells. *Cell* 38, 745-755.

Fu, X.Y., Kessler, D.S., Veals, S.A., Levy, D.E., and Darnell, J.E., Jr. (1990). ISGF3, the transcriptional activator induced by interferon alpha, consists of multiple interacting polypeptide chains. *Proc Natl Acad Sci U S A* 87, 8555-8559.

Fukuhara, H., Ino, Y., and Todo, T. (2016). Oncolytic virus therapy: A new era of cancer treatment at dawn. *Cancer Sci* 107, 1373-1379.

Funabiki, M., Kato, H., Miyachi, Y., Toki, H., Motegi, H., Inoue, M., Minowa, O., Yoshida, A., Deguchi, K., Sato, H., *et al.* (2014). Autoimmune disorders associated with gain of function of the intracellular sensor MDA5. *Immunity* 40, 199-212.

Fye, J.M., Orebaugh, C.D., Coffin, S.R., Hollis, T., and Perrino, F.W. (2011). Dominant mutation of the TREX1 exonuclease gene in lupus and Aicardi-Goutieres syndrome. *J Biol Chem* 286, 32373-32382.

Gaelings, L., Soderholm, S., Bugai, A., Fu, Y., Nandania, J., Schepens, B., Lorey, M.B., Tynell, J., Vande Ginste, L., Le Goffic, R., *et al.* (2017). Regulation of kynurenine biosynthesis during influenza virus infection. *FEBS J* 284, 222-236.

Gao, D., Rahbar, R., and Fish, E.N. (2016). CCL5 activation of CCR5 regulates cell metabolism to enhance proliferation of breast cancer cells. *Open Biol* 6.

- Gao, P., Ascano, M., Wu, Y., Barchet, W., Gaffney, B.L., Zillinger, T., Serganov, A.A., Liu, Y., Jones, R.A., Hartmann, G., *et al.* (2013). Cyclic [G(2',5')pA(3',5')p] is the metazoan second messenger produced by DNA-activated cyclic GMP-AMP synthase. *Cell* 153, 1094-1107.
- Gerlier, D., and Lyles, D.S. (2011). Interplay between innate immunity and negative-strand RNA viruses: towards a rational model. *Microbiol Mol Biol Rev* 75, 468-490, second page of table of contents.
- Gibbs, P.E., Miralem, T., and Maines, M.D. (2015). Biliverdin reductase: a target for cancer therapy? *Front Pharmacol* 6, 119.
- Goodbourn, S., Zinn, K., and Maniatis, T. (1985). Human beta-interferon gene expression is regulated by an inducible enhancer element. *Cell* 41, 509-520.
- Gotoh, B., Komatsu, T., Takeuchi, K., and Yokoo, J. (2002). Paramyxovirus strategies for evading the interferon response. *Rev Med Virol* 12, 337-357.
- Greber, U.F., and Way, M. (2006). A superhighway to virus infection. *Cell* 124, 741-754.
- Gresser, I., Morel-Maroger, L., Riviere, Y., Guillon, J.C., Tovey, M.G., Woodrow, D., Sloper, J.C., and Moss, J. (1980). Interferon-induced disease in mice and rats. *Ann N Y Acad Sci* 350, 12-20.
- Groom, J.R., and Luster, A.D. (2011). CXCR3 ligands: redundant, collaborative and antagonistic functions. *Immunol Cell Biol* 89, 207-215.
- Gualdoni, G.A., Mayer, K.A., Kapsch, A.M., Kreuzberg, K., Puck, A., Kienzl, P., Oberndorfer, F., Fruhwirth, K., Winkler, S., Blaas, D., *et al.* (2018). Rhinovirus induces an anabolic reprogramming in host cell metabolism essential for viral replication. *Proc Natl Acad Sci U S A* 115, E7158-E7165.
- Gunther, C., Kind, B., Reijns, M.A., Berndt, N., Martinez-Bueno, M., Wolf, C., Tungler, V., Chara, O., Lee, Y.A., Hubner, N., *et al.* (2015). Defective removal of ribonucleotides from DNA promotes systemic autoimmunity. *J Clin Invest* 125, 413-424.
- Guttman, M., Garber, M., Levin, J.Z., Donaghey, J., Robinson, J., Adiconis, X., Fan, L., Koziol, M.J., Gnirke, A., Nusbaum, C., *et al.* (2010). Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. *Nat Biotechnol* 28, 503-510.
- Guttman, M., and Rinn, J.L. (2012). Modular regulatory principles of large non-coding RNAs. *Nature* 482, 339-346.

Guttman, M., Russell, P., Ingolia, N.T., Weissman, J.S., and Lander, E.S. (2013). Ribosome profiling provides evidence that large noncoding RNAs do not encode proteins. *Cell* 154, 240-251.

Haller, O., Staeheli, P., Schwemmle, M., and Kochs, G. (2015). Mx GTPases: dynamin-like antiviral machines of innate immunity. *Trends Microbiol* 23, 154-163.

Hiller, B., Achleitner, M., Glage, S., Naumann, R., Behrendt, R., and Roers, A. (2012). Mammalian RNase H2 removes ribonucleotides from DNA to maintain genome integrity. *J Exp Med* 209, 1419-1426.

Honda, K., and Taniguchi, T. (2006). IRFs: master regulators of signalling by Toll-like receptors and cytosolic pattern-recognition receptors. *Nat Rev Immunol* 6, 644-658.

Honma, S., Kawamoto, T., Takagi, Y., Fujimoto, K., Sato, F., Noshiro, M., Kato, Y., and Honma, K. (2002). Dec1 and Dec2 are regulators of the mammalian molecular clock. *Nature* 419, 841-844.

Hooks, J.J., Moutsopoulos, H.M., Geis, S.A., Stahl, N.I., Decker, J.L., and Notkins, A.L. (1979). Immune interferon in the circulation of patients with autoimmune disease. *N Engl J Med* 301, 5-8.

Hu, M.M., Liao, C.Y., Yang, Q., Xie, X.Q., and Shu, H.B. (2017). Innate immunity to RNA virus is regulated by temporal and reversible sumoylation of RIG-I and MDA5. *J Exp Med* 214, 973-989.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* 37, 1-13.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4, 44-57.

Huang, D.W., Sherman, B.T., Tan, Q., Collins, J.R., Alvord, W.G., Roayaei, J., Stephens, R., Baseler, M.W., Lane, H.C., and Lempicki, R.A. (2007). The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biol* 8, R183.

Imamura, K., Imamachi, N., Akizuki, G., Kumakura, M., Kawaguchi, A., Nagata, K., Kato, A., Kawaguchi, Y., Sato, H., Yoneda, M., *et al.* (2014). Long noncoding RNA NEAT1-dependent SFPQ relocation from promoter region to paraspeckle mediates IL8 expression upon immune stimuli. *Mol Cell* 53, 393-406.

Improta, T., Schindler, C., Horvath, C.M., Kerr, I.M., Stark, G.R., and Darnell, J.E., Jr. (1994). Transcription factor ISGF-3 formation requires phosphorylated Stat91 protein, but

Stat113 protein is phosphorylated independently of Stat91 protein. *Proc Natl Acad Sci U S A* 91, 4776-4780.

Ishikawa, H., Ma, Z., and Barber, G.N. (2009). STING regulates intracellular DNA-mediated, type I interferon-dependent innate immunity. *Nature* 461, 788-792.

Janeway, C.A., Jr. (1992). The immune system evolved to discriminate infectious nonself from noninfectious self. *Immunol Today* 13, 11-16.

Jang, M.A., Kim, E.K., Now, H., Nguyen, N.T., Kim, W.J., Yoo, J.Y., Lee, J., Jeong, Y.M., Kim, C.H., Kim, O.H., *et al.* (2015). Mutations in DDX58, which encodes RIG-I, cause atypical Singleton-Merten syndrome. *Am J Hum Genet* 96, 266-274.

Johnston, M.D. (1981). The characteristics required for a Sendai virus preparation to induce high levels of interferon in human lymphoblastoid cells. *J Gen Virol* 56, 175-184.

Kambara, H., Gunawardane, L., Zebrowski, E., Kostadinova, L., Jobava, R., Krokowski, D., Hatzoglou, M., Anthony, D.D., and Valadkhan, S. (2014a). Regulation of Interferon-Stimulated Gene BST2 by a lncRNA Transcribed from a Shared Bidirectional Promoter. *Front Immunol* 5, 676.

Kambara, H., Niazi, F., Kostadinova, L., Moonka, D.K., Siegel, C.T., Post, A.B., Carnero, E., Barriocanal, M., Fortes, P., Anthony, D.D., *et al.* (2014b). Negative regulation of the interferon response by an interferon-induced long non-coding RNA. *Nucleic Acids Res* 42, 10668-10680.

Karolchik, D., Barber, G.P., Casper, J., Clawson, H., Cline, M.S., Diekhans, M., Dreszer, T.R., Fujita, P.A., Guruvadoo, L., Haeussler, M., *et al.* (2014). The UCSC Genome Browser database: 2014 update. *Nucleic Acids Res* 42, D764-770.

Kato, H., Oh, S.W., and Fujita, T. (2017). RIG-I-Like Receptors and Type I Interferonopathies. *J Interferon Cytokine Res* 37, 207-213.

Kawai, T., and Akira, S. (2006a). Innate immune recognition of viral infection. *Nat Immunol* 7, 131-137.

Kawai, T., and Akira, S. (2006b). TLR signaling. *Cell Death Differ* 13, 816-825.

Kawasaki, T., and Kawai, T. (2014). Toll-like receptor signaling pathways. *Front Immunol* 5, 461.

Kay, M.A., Glorioso, J.C., and Naldini, L. (2001). Viral vectors for gene therapy: the art of turning infectious agents into vehicles of therapeutics. *Nat Med* 7, 33-40.

- Kessler, D.S., Veals, S.A., Fu, X.Y., and Levy, D.E. (1990). Interferon-alpha regulates nuclear translocation and DNA-binding affinity of ISGF3, a multimeric transcriptional activator. *Genes Dev* 4, 1753-1765.
- Kiefer, K., Oropallo, M.A., Cancro, M.P., and Marshak-Rothstein, A. (2012). Role of type I interferons in the activation of autoreactive B cells. *Immunol Cell Biol* 90, 498-504.
- Kim, T.K., and Maniatis, T. (1997). The mechanism of transcriptional synergy of an in vitro assembled interferon-beta enhanceosome. *Mol Cell* 1, 119-129.
- Kobayashi, K.S., and Flavell, R.A. (2004). Shielding the double-edged sword: negative regulation of the innate immune system. *J Leukoc Biol* 75, 428-433.
- Kosiol, C., Vinar, T., da Fonseca, R.R., Hubisz, M.J., Bustamante, C.D., Nielsen, R., and Siepel, A. (2008). Patterns of positive selection in six Mammalian genomes. *PLoS Genet* 4, e1000144.
- Kretschmer, S., Wolf, C., Konig, N., Staroske, W., Guck, J., Hausler, M., Luksch, H., Nguyen, L.A., Kim, B., Alexopoulou, D., *et al.* (2015). SAMHD1 prevents autoimmunity by maintaining genome stability. *Ann Rheum Dis* 74, e17.
- Kvansakul, M., Caria, S., and Hinds, M.G. (2017). The Bcl-2 Family in Host-Virus Interactions. *Viruses* 9.
- Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., *et al.* (2001). Initial sequencing and analysis of the human genome. *Nature* 409, 860-921.
- Larner, A.C., Chaudhuri, A., and Darnell, J.E., Jr. (1986). Transcriptional induction by interferon. New protein(s) determine the extent and length of the induction. *J Biol Chem* 261, 453-459.
- Larner, A.C., Jonak, G., Cheng, Y.S., Korant, B., Knight, E., and Darnell, J.E., Jr. (1984). Transcriptional induction of two genes in human cells by beta interferon. *Proc Natl Acad Sci U S A* 81, 6733-6737.
- Laska, M.J., Trolborg, A., Hansen, B., Stengaard-Pedersen, K., Junker, P., Nexø, B.A., and Voss, A. (2014). Polymorphisms within Toll-like receptors are associated with systemic lupus erythematosus in a cohort of Danish females. *Rheumatology (Oxford)* 53, 48-55.
- Levy, D.E., Kessler, D.S., Pine, R., Reich, N., and Darnell, J.E., Jr. (1988). Interferon-induced nuclear factors that bind a shared promoter element correlate with positive and negative transcriptional control. *Genes Dev* 2, 383-393.

Li, A.P., Kaminski, D.L., and Rasmussen, A. (1995). Substrates of human hepatic cytochrome P450 3A4. *Toxicology* 104, 1-8.

Li, P., Du, J., Goodier, J.L., Hou, J., Kang, J., Kazazian, H.H., Jr., Zhao, K., and Yu, X.F. (2017). Aicardi-Goutieres syndrome protein TREX1 suppresses L1 and maintains genome integrity through exonuclease-independent ORF1p depletion. *Nucleic Acids Res* 45, 4619-4631.

Li, X.D., Sun, L., Seth, R.B., Pineda, G., and Chen, Z.J. (2005). Hepatitis C virus protease NS3/4A cleaves mitochondrial antiviral signaling protein off the mitochondria to evade innate immunity. *Proc Natl Acad Sci U S A* 102, 17717-17722.

Liddicoat, B.J., Chalk, A.M., and Walkley, C.R. (2016). ADAR1, inosine and the immune sensing system: distinguishing self from non-self. *Wiley Interdiscip Rev RNA* 7, 157-172.

Lin, M.F., Jungreis, I., and Kellis, M. (2011). PhyloCSF: a comparative genomics method to distinguish protein coding and non-coding regions. *Bioinformatics* 27, i275-282.

Lin, R., Heylbroeck, C., Genin, P., Pitha, P.M., and Hiscott, J. (1999). Essential role of interferon regulatory factor 3 in direct activation of RANTES chemokine transcription. *Mol Cell Biol* 19, 959-966.

Liu, B., Sun, L., Liu, Q., Gong, C., Yao, Y., Lv, X., Lin, L., Yao, H., Su, F., Li, D., *et al.* (2015). A cytoplasmic NF-kappaB interacting long noncoding RNA blocks IkappaB phosphorylation and suppresses breast cancer metastasis. *Cancer Cell* 27, 370-381.

Lomvardas, S., and Thanos, D. (2001). Nucleosome sliding via TBP DNA binding in vivo. *Cell* 106, 685-696.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15, 550.

Ma, Z., Jacobs, S.R., West, J.A., Stopford, C., Zhang, Z., Davis, Z., Barber, G.N., Glaunsinger, B.A., Dittmer, D.P., and Damania, B. (2015). Modulation of the cGAS-STING DNA sensing pathway by gammaherpesviruses. *Proc Natl Acad Sci U S A* 112, E4306-4315.

Mackenzie, K.J., Carroll, P., Lettice, L., Tarnauskaite, Z., Reddy, K., Dix, F., Revuelta, A., Abbondati, E., Rigby, R.E., Rabe, B., *et al.* (2016). Ribonuclease H2 mutations induce a cGAS/STING-dependent innate immune response. *EMBO J* 35, 831-844.

Martinez Valle, F., Balada, E., Ordi-Ros, J., and Vilardell-Tarres, M. (2008). DNase 1 and systemic lupus erythematosus. *Autoimmun Rev* 7, 359-363.

Mavragani, C.P., Sagalovskiy, I., Guo, Q., Nezos, A., Kapsogeorgou, E.K., Lu, P., Liang Zhou, J., Kirou, K.A., Seshan, S.V., Moutsopoulos, H.M., *et al.* (2016). Expression of Long

Interspersed Nuclear Element 1 Retroelements and Induction of Type I Interferon in Patients With Systemic Autoimmune Disease. *Arthritis Rheumatol* 68, 2686-2696.

McNab, F., Mayer-Barber, K., Sher, A., Wack, A., and O'Garra, A. (2015). Type I interferons in infectious disease. *Nat Rev Immunol* 15, 87-103.

Medzhitov, R., and Janeway, C.A., Jr. (1997). Innate immunity: the virtues of a nonclonal system of recognition. *Cell* 91, 295-298.

Meyerson, N.R., and Sawyer, S.L. (2011). Two-stepping through time: mammals and viruses. *Trends Microbiol* 19, 286-294.

Mogensen, T.H. (2009). Pathogen recognition and inflammatory signaling in innate immune defenses. *Clin Microbiol Rev* 22, 240-273, Table of Contents.

Nakamura, H., Tanimoto, K., Hiyama, K., Yunokawa, M., Kawamoto, T., Kato, Y., Yoshiga, K., Poellinger, L., Hiyama, E., and Nishiyama, M. (2008). Human mismatch repair gene, MLH1, is transcriptionally repressed by the hypoxia-inducible transcription factors, DEC1 and DEC2. *Oncogene* 27, 4200-4209.

Nishikura, K. (2010). Functions and regulation of RNA editing by ADAR deaminases. *Annu Rev Biochem* 79, 321-349.

Novick, D., Cohen, B., and Rubinstein, M. (1994). The human interferon alpha/beta receptor: characterization and molecular cloning. *Cell* 77, 391-400.

Oda, H., Nakagawa, K., Abe, J., Awaya, T., Funabiki, M., Hijikata, A., Nishikomori, R., Funatsuka, M., Ohshima, Y., Sugawara, Y., *et al.* (2014). Aicardi-Goutieres syndrome is caused by IFIH1 mutations. *Am J Hum Genet* 95, 121-125.

Oeckinghaus, A., and Ghosh, S. (2009). The NF-kappaB family of transcription factors and its regulation. *Cold Spring Harb Perspect Biol* 1, a000034.

Orecchini, E., Doria, M., Antonioni, A., Galardi, S., Ciafre, S.A., Frassinelli, L., Mancone, C., Montaldo, C., Tripodi, M., and Michienzi, A. (2017). ADAR1 restricts LINE-1 retrotransposition. *Nucleic Acids Res* 45, 155-168.

Ouyang, J., Zhu, X., Chen, Y., Wei, H., Chen, Q., Chi, X., Qi, B., Zhang, L., Zhao, Y., Gao, G.F., *et al.* (2014). NRAV, a long noncoding RNA, modulates antiviral responses through suppression of interferon-stimulated gene transcription. *Cell Host Microbe* 16, 616-626.

Parisien, J.P., Bamming, D., Komuro, A., Ramachandran, A., Rodriguez, J.J., Barber, G., Wojahn, R.D., and Horvath, C.M. (2009). A shared interface mediates paramyxovirus interference with antiviral RNA helicases MDA5 and LGP2. *J Virol* 83, 7252-7260.

Parisien, J.P., Lenoir, J.J., Mandhana, R., Rodriguez, K.R., Qian, K., Bruns, A.M., and Horvath, C.M. (2018). RNA sensor LGP2 inhibits TRAF ubiquitin ligase to negatively regulate innate immune signaling. *EMBO Rep* 19.

Paun, A., and Pitha, P.M. (2007). The IRF family, revisited. *Biochimie* 89, 744-753.

Peisley, A., Jo, M.H., Lin, C., Wu, B., Orme-Johnson, M., Walz, T., Hohng, S., and Hur, S. (2012). Kinetic mechanism for viral dsRNA length discrimination by MDA5 filaments. *Proc Natl Acad Sci U S A* 109, E3340-3349.

Peisley, A., Lin, C., Wu, B., Orme-Johnson, M., Liu, M., Walz, T., and Hur, S. (2011). Cooperative assembly and dynamic disassembly of MDA5 filaments for viral dsRNA recognition. *Proc Natl Acad Sci U S A* 108, 21010-21015.

Pichlmair, A., Lassnig, C., Eberle, C.A., Gorna, M.W., Baumann, C.L., Burkard, T.R., Burckstummer, T., Stefanovic, A., Krieger, S., Bennett, K.L., *et al.* (2011). IFIT1 is an antiviral protein that recognizes 5'-triphosphate RNA. *Nat Immunol* 12, 624-630.

Piganis, R.A., De Weerd, N.A., Gould, J.A., Schindler, C.W., Mansell, A., Nicholson, S.E., and Hertzog, P.J. (2011). Suppressor of cytokine signaling (SOCS) 1 inhibits type I interferon (IFN) signaling via the interferon alpha receptor (IFNAR1)-associated tyrosine kinase Tyk2. *J Biol Chem* 286, 33811-33818.

Pindel, A., and Sadler, A. (2011). The role of protein kinase R in the interferon response. *J Interferon Cytokine Res* 31, 59-70.

Porritt, R.A., and Hertzog, P.J. (2015). Dynamic control of type I IFN signalling by an integrated network of negative regulators. *Trends Immunol* 36, 150-160.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841-842.

Ramachandran, A., and Horvath, C.M. (2009). Paramyxovirus disruption of interferon signal transduction: STATus report. *J Interferon Cytokine Res* 29, 531-537.

Ramachandran, A., and Horvath, C.M. (2010). Dissociation of paramyxovirus interferon evasion activities: universal and virus-specific requirements for conserved V protein amino acids in MDA5 interference. *J Virol* 84, 11152-11163.

Ramachandran, A., Parisien, J.P., and Horvath, C.M. (2008). STAT2 is a primary target for measles virus V protein-mediated alpha/beta interferon signaling inhibition. *J Virol* 82, 8330-8338.

Reich, N., Evans, B., Levy, D., Fahey, D., Knight, E., Jr., and Darnell, J.E., Jr. (1987). Interferon-induced transcription of a gene encoding a 15-kDa protein depends on an upstream enhancer element. *Proc Natl Acad Sci U S A* 84, 6394-6398.

Rice, G.I., Bond, J., Asipu, A., Brunette, R.L., Manfield, I.W., Carr, I.M., Fuller, J.C., Jackson, R.M., Lamb, T., Briggs, T.A., *et al.* (2009). Mutations involved in Aicardi-Goutieres syndrome implicate SAMHD1 as regulator of the innate immune response. *Nat Genet* 41, 829-832.

Rice, G.I., Del Toro Duany, Y., Jenkinson, E.M., Forte, G.M., Anderson, B.H., Ariaudo, G., Bader-Meunier, B., Baildam, E.M., Battini, R., Beresford, M.W., *et al.* (2014). Gain-of-function mutations in IFIH1 cause a spectrum of human disease phenotypes associated with upregulated type I interferon signaling. *Nat Genet* 46, 503-509.

Rice, G.I., Forte, G.M., Szykiewicz, M., Chase, D.S., Aeby, A., Abdel-Hamid, M.S., Ackroyd, S., Allcock, R., Bailey, K.M., Balottin, U., *et al.* (2013). Assessment of interferon-related biomarkers in Aicardi-Goutieres syndrome associated with mutations in TREX1, RNASEH2A, RNASEH2B, RNASEH2C, SAMHD1, and ADAR: a case-control study. *Lancet Neurol* 12, 1159-1169.

Rice, G.I., Kasher, P.R., Forte, G.M., Mannion, N.M., Greenwood, S.M., Szykiewicz, M., Dickerson, J.E., Bhaskar, S.S., Zampini, M., Briggs, T.A., *et al.* (2012). Mutations in ADAR1 cause Aicardi-Goutieres syndrome associated with a type I interferon signature. *Nat Genet* 44, 1243-1248.

Rinn, J.L., and Chang, H.Y. (2012). Genome regulation by long noncoding RNAs. *Annu Rev Biochem* 81, 145-166.

Rodero, M.P., Tesser, A., Bartok, E., Rice, G.I., Della Mina, E., Depp, M., Beitz, B., Bondet, V., Cagnard, N., Duffy, D., *et al.* (2017). Type I interferon-mediated autoinflammation due to DNase II deficiency. *Nat Commun* 8, 2176.

Rodriguez, J.J., Parisien, J.P., and Horvath, C.M. (2002). Nipah virus V protein evades alpha and gamma interferons by preventing STAT1 and STAT2 activation and nuclear accumulation. *J Virol* 76, 11476-11483.

Rodriguez, J.J., Wang, L.F., and Horvath, C.M. (2003). Hendra virus V protein inhibits interferon signaling by preventing STAT1 and STAT2 nuclear accumulation. *J Virol* 77, 11842-11845.

Rodriguez, K.R., and Horvath, C.M. (2013). Amino acid requirements for MDA5 and LGP2 recognition by paramyxovirus V proteins: a single arginine distinguishes MDA5 from RIG-I. *J Virol* 87, 2974-2978.

Rodriguez, K.R., and Horvath, C.M. (2014). Paramyxovirus V protein interaction with the antiviral sensor LGP2 disrupts MDA5 signaling enhancement but is not relevant to LGP2-mediated RLR signaling inhibition. *J Virol* 88, 8180-8188.

Rutsch, F., MacDougall, M., Lu, C., Buers, I., Mamaeva, O., Nitschke, Y., Rice, G.I., Erlandsen, H., Kehl, H.G., Thiele, H., *et al.* (2015). A specific IFIH1 gain-of-function mutation causes Singleton-Merten syndrome. *Am J Hum Genet* 96, 275-282.

Ryals, J., Dierks, P., Ragg, H., and Weissmann, C. (1985). A 46-nucleotide promoter segment from an IFN-alpha gene renders an unrelated promoter inducible by virus. *Cell* 41, 497-507.

Sanchez, E.L., and Lagunoff, M. (2015). Viral activation of cellular metabolism. *Virology* 479-480, 609-618.

Schlee, M., and Hartmann, G. (2016). Discriminating self from non-self in nucleic acid sensing. *Nat Rev Immunol* 16, 566-580.

Schmittgen, T.D., and Livak, K.J. (2008). Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc* 3, 1101-1108.

Schoggins, J.W., Wilson, S.J., Panis, M., Murphy, M.Y., Jones, C.T., Bieniasz, P., and Rice, C.M. (2011). A diverse range of gene products are effectors of the type I interferon antiviral response. *Nature* 472, 481-485.

Shuai, K., Ziemiecki, A., Wilks, A.F., Harpur, A.G., Sadowski, H.B., Gilman, M.Z., and Darnell, J.E. (1993). Polypeptide signalling to the nucleus through tyrosine phosphorylation of Jak and Stat proteins. *Nature* 366, 580-583.

Siepel, A., Bejerano, G., Pedersen, J.S., Hinrichs, A.S., Hou, M., Rosenbloom, K., Clawson, H., Spieth, J., Hillier, L.W., Richards, S., *et al.* (2005). Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res* 15, 1034-1050.

Skon, C.N., Lee, J.Y., Anderson, K.G., Masopust, D., Hogquist, K.A., and Jameson, S.C. (2013). Transcriptional downregulation of *S1pr1* is required for the establishment of resident memory CD8⁺ T cells. *Nat Immunol* 14, 1285-1293.

Smith, W.L., Garavito, R.M., and DeWitt, D.L. (1996). Prostaglandin endoperoxide H synthases (cyclooxygenases)-1 and -2. *J Biol Chem* 271, 33157-33160.

Stark, G.R., Kerr, I.M., Williams, B.R., Silverman, R.H., and Schreiber, R.D. (1998). How cells respond to interferons. *Annu Rev Biochem* 67, 227-264.

Sultana, H., Neelakanta, G., Foellmer, H.G., Montgomery, R.R., Anderson, J.F., Koski, R.A., Medzhitov, R.M., and Fikrig, E. (2012). Semaphorin 7A contributes to West Nile virus pathogenesis through TGF-beta1/Smad6 signaling. *J Immunol* 189, 3150-3158.

Sun, L., Wu, J., Du, F., Chen, X., and Chen, Z.J. (2013). Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway. *Science* 339, 786-791.

Takeuchi, O., and Akira, S. (2009). Innate immunity to virus infection. *Immunol Rev* 227, 75-86.

Tamura, T., Yanai, H., Savitsky, D., and Taniguchi, T. (2008). The IRF family transcription factors in immunity and oncogenesis. *Annu Rev Immunol* 26, 535-584.

Taniguchi, T., Ogasawara, K., Takaoka, A., and Tanaka, N. (2001). IRF family of transcription factors as regulators of host defense. *Annu Rev Immunol* 19, 623-655.

Teijaro, J.R., Studer, S., Leaf, N., Kiosses, W.B., Nguyen, N., Matsuki, K., Negishi, H., Taniguchi, T., Oldstone, M.B., and Rosen, H. (2016). S1PR1-mediated IFNAR1 degradation modulates plasmacytoid dendritic cell interferon-alpha autoamplification. *Proc Natl Acad Sci U S A* 113, 1351-1356.

ten Hoeve, J., de Jesus Ibarra-Sanchez, M., Fu, Y., Zhu, W., Tremblay, M., David, M., and Shuai, K. (2002). Identification of a nuclear Stat1 protein tyrosine phosphatase. *Mol Cell Biol* 22, 5662-5668.

Teng, Y., Pi, W., Wang, Y., and Cowell, J.K. (2016). WASF3 provides the conduit to facilitate invasion and metastasis in breast cancer cells through HER2/HER3 signaling. *Oncogene* 35, 4633-4640.

Tough, D.F. (2012). Modulation of T-cell function by type I interferon. *Immunol Cell Biol* 90, 492-497.

Tough, D.F., Borrow, P., and Sprent, J. (1996). Induction of bystander T cell proliferation by viruses and type I interferon in vivo. *Science* 272, 1947-1950.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25, 1105-1111.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* 28, 511-515.

Trinchieri, G. (2010). Type I interferon: friend or foe? *J Exp Med* 207, 2053-2063.

Ulane, C.M., Kentsis, A., Cruz, C.D., Parisien, J.P., Schneider, K.L., and Horvath, C.M. (2005). Composition and assembly of STAT-targeting ubiquitin ligase complexes: paramyxovirus V protein carboxyl terminus is an oligomerization domain. *J Virol* 79, 10180-10189.

Ungureanu, D., Vanhatupa, S., Gronholm, J., Palvimo, J.J., and Silvennoinen, O. (2005). SUMO-1 conjugation selectively modulates STAT1-mediated gene responses. *Blood* 106, 224-226.

Uze, G., Lutfalla, G., and Gresser, I. (1990). Genetic transfer of a functional human interferon alpha receptor into mouse cells: cloning and expression of its cDNA. *Cell* 60, 225-234.

Valadkhan, S., and Gunawardane, L.S. (2016). lncRNA-mediated regulation of the interferon response. *Virus Res* 212, 127-136.

van der Pouw Kraan, T.C., Wijbrandts, C.A., van Baarsen, L.G., Voskuyl, A.E., Rustenburg, F., Baggen, J.M., Ibrahim, S.M., Fero, M., Dijkmans, B.A., Tak, P.P., *et al.* (2007). Rheumatoid arthritis subtypes identified by genomic profiling of peripheral blood cells: assignment of a type I interferon signature in a subpopulation of patients. *Ann Rheum Dis* 66, 1008-1014.

Van Eyck, L., De Somer, L., Pombal, D., Bornschein, S., Frans, G., Humblet-Baron, S., Moens, L., de Zegher, F., Bossuyt, X., Wouters, C., *et al.* (2015). Brief Report: IFIH1 Mutation Causes Systemic Lupus Erythematosus With Selective IgA Deficiency. *Arthritis Rheumatol* 67, 1592-1597.

van Herwaarden, A.E., Wagenaar, E., van der Kruijssen, C.M., van Waterschoot, R.A., Smit, J.W., Song, J.Y., van der Valk, M.A., van Tellingen, O., van der Hoorn, J.W., Rosing, H., *et al.* (2007). Knockout of cytochrome P450 3A yields new mouse models for understanding xenobiotic metabolism. *J Clin Invest* 117, 3583-3592.

Vander Heiden, M.G. (2011). Targeting cancer metabolism: a therapeutic window opens. *Nat Rev Drug Discov* 10, 671-684.

Vaughan, J.C., Brandenburg, B., Hogle, J.M., and Zhuang, X. (2009). Rapid actin-dependent viral motility in live cells. *Biophys J* 97, 1647-1656.

Versteeg, G.A., and Garcia-Sastre, A. (2010). Viral tricks to grid-lock the type I interferon system. *Curr Opin Microbiol* 13, 508-516.

Wang, C.M., Chang, S.W., Wu, Y.J., Lin, J.C., Ho, H.H., Chou, T.C., Yang, B., Wu, J., and Chen, J.Y. (2014). Genetic variations in Toll-like receptors (TLRs 3/7/8) are associated with systemic lupus erythematosus in a Taiwanese population. *Sci Rep* 4, 3792.

Wang, P., Xu, J., Wang, Y., and Cao, X. (2017). An interferon-independent lncRNA promotes viral replication by modulating cellular metabolism. *Science* 358, 1051-1055.

Wang, Y., Lian, Q., Yang, B., Yan, S., Zhou, H., He, L., Lin, G., Lian, Z., Jiang, Z., and Sun, B. (2015). TRIM30alpha Is a Negative-Feedback Regulator of the Intracellular DNA and DNA Virus-Triggered Response by Targeting STING. *PLoS Pathog* 11, e1005012.

- Weber, F., Wagner, V., Rasmussen, S.B., Hartmann, R., and Paludan, S.R. (2006). Double-stranded RNA is produced by positive-strand RNA viruses and DNA viruses but not in detectable amounts by negative-strand RNA viruses. *J Virol* *80*, 5059-5064.
- Whittemore, L.A., and Maniatis, T. (1990). Postinduction turnoff of beta-interferon gene expression. *Mol Cell Biol* *10*, 1329-1337.
- Wies, E., Wang, M.K., Maharaj, N.P., Chen, K., Zhou, S., Finberg, R.W., and Gack, M.U. (2013). Dephosphorylation of the RNA sensors RIG-I and MDA5 by the phosphatase PP1 is essential for innate immune signaling. *Immunity* *38*, 437-449.
- Wong, J.J., Pung, Y.F., Sze, N.S., and Chin, K.C. (2006). HERC5 is an IFN-induced HECT-type E3 protein ligase that mediates type I IFN-induced ISGylation of protein targets. *Proc Natl Acad Sci U S A* *103*, 10735-10740.
- Wu, B., Peisley, A., Richards, C., Yao, H., Zeng, X., Lin, C., Chu, F., Walz, T., and Hur, S. (2013a). Structural basis for dsRNA recognition, filament formation, and antiviral signal activation by MDA5. *Cell* *152*, 276-289.
- Wu, J., Sun, L., Chen, X., Du, F., Shi, H., Chen, C., and Chen, Z.J. (2013b). Cyclic GMP-AMP is an endogenous second messenger in innate immune signaling by cytosolic DNA. *Science* *339*, 826-830.
- Wu, J.J., Li, W., Shao, Y., Avey, D., Fu, B., Gillen, J., Hand, T., Ma, S., Liu, X., Miley, W., *et al.* (2015a). Inhibition of cGAS DNA Sensing by a Herpesvirus Virion Protein. *Cell Host Microbe* *18*, 333-344.
- Wu, Y.W., Tang, W., and Zuo, J.P. (2015b). Toll-like receptors: potential targets for lupus treatment. *Acta Pharmacol Sin* *36*, 1395-1407.
- Yan, H., Krishnan, K., Greenlund, A.C., Gupta, S., Lim, J.T., Schreiber, R.D., Schindler, C.W., and Krolewski, J.J. (1996). Phosphorylated interferon-alpha receptor 1 subunit (IFNAR1) acts as a docking site for the latent form of the 113 kDa STAT2 protein. *EMBO J* *15*, 1064-1074.
- Yang, Y.G., Lindahl, T., and Barnes, D.E. (2007). Trex1 exonuclease degrades ssDNA to prevent chronic checkpoint activation and autoimmune disease. *Cell* *131*, 873-886.
- Yoon, J.H., Abdelmohsen, K., and Gorospe, M. (2013). Posttranscriptional gene regulation by long noncoding RNA. *J Mol Biol* *425*, 3723-3730.
- Zhang, D., and Zhang, D.E. (2011). Interferon-stimulated gene 15 and the protein ISGylation system. *J Interferon Cytokine Res* *31*, 119-130.

Zhang, J.H., Chung, T.D., and Oldenburg, K.R. (1999). A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *J Biomol Screen* 4, 67-73.

Zhang, X., Shi, H., Wu, J., Zhang, X., Sun, L., Chen, C., and Chen, Z.J. (2013). Cyclic GMP-AMP containing mixed phosphodiester linkages is an endogenous high-affinity ligand for STING. *Mol Cell* 51, 226-235.

Zhao, K., Du, J., Han, X., Goodier, J.L., Li, P., Zhou, X., Wei, W., Evans, S.L., Li, L., Zhang, W., *et al.* (2013). Modulation of LINE-1 and Alu/SVA retrotransposition by Aicardi-Goutieres syndrome-related SAMHD1. *Cell Rep* 4, 1108-1115.

Zhao, K., Du, J., Peng, Y., Li, P., Wang, S., Wang, Y., Hou, J., Kang, J., Zheng, W., Hua, S., *et al.* (2018). LINE1 contributes to autoimmunity through both RIG-I- and MDA5-mediated RNA sensing pathways. *J Autoimmun.*

Zheng, H., Qian, J., Varghese, B., Baker, D.P., and Fuchs, S. (2011). Ligand-stimulated downregulation of the alpha interferon receptor: role of protein kinase D2. *Mol Cell Biol* 31, 710-720.

Zhong, B., Zhang, L., Lei, C., Li, Y., Mao, A.P., Yang, Y., Wang, Y.Y., Zhang, X.L., and Shu, H.B. (2009). The ubiquitin ligase RNF5 regulates antiviral responses by mediating degradation of the adaptor protein MITA. *Immunity* 30, 397-407.

Zhu, J., Ghosh, A., and Sarkar, S.N. (2015). OASL-a new player in controlling antiviral innate immunity. *Curr Opin Virol* 12, 15-19.

Ziehr, B., Vincent, H.A., and Moorman, N.J. (2016). Human Cytomegalovirus pTRS1 and pIRS1 Antagonize Protein Kinase R To Facilitate Virus Replication. *J Virol* 90, 3839-3848.

Zinn, K., DiMaio, D., and Maniatis, T. (1983). Identification of two distinct regulatory regions adjacent to the human beta-interferon gene. *Cell* 34, 865-879.

Zust, R., Cervantes-Barragan, L., Habjan, M., Maier, R., Neuman, B.W., Ziebuhr, J., Szretter, K.J., Baker, S.C., Barchet, W., Diamond, M.S., *et al.* (2011). Ribose 2'-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5. *Nat Immunol* 12, 137-143.

APPENDIX A: MATERIALS AND METHODS

GENERAL METHODS

Cell culture

Namalwa B cells (ATCC CRL-1432) and THP-1 cells (ATCC TIB-202) were cultured in RPMI 1640 Medium (Life Technologies, Thermo Fisher Scientific). HEK293T (ATCC CRL-11270), 2fTGH cells (ECACC 12021508) and A549-DualTM cells (InvivoGen) were cultured in Dulbecco's Modified Eagle Medium (Life Technologies, Thermo Fisher Scientific).

For all cell lines except THP-1, medium was supplemented with 10% Cosmic calf serum (HyClone, GE Healthcare Life Sciences) and 1% Penicillin-Streptomycin (Gibco, Thermo Fisher Scientific). For A549-DualTM cells, media was also supplemented with Zeocin (100 µg/mL), blasticidin (10 µg/mL) and Normocin (100 µg/mL). For THP-1 cells, medium was supplemented with 10% heat-inactivated fetal bovine serum, 1% Penicillin-Streptomycin (Gibco, Thermo Fisher Scientific), 1 mM sodium pyruvate, 10 mM Hepes, 0.63% glucose solution and 0.00017% βME.

Virus infection

Sendai virus (Cantell strain) was grown in embryonated chicken eggs, and titers were determined on Vero cells. The A/Udorn/72 strain of influenza virus (gift of R. A. Lamb, Northwestern University) was propagated and titered on MDCK cells. HSV-1 (F strain) (gift of G. A. Smith, Northwestern University) was propagated and titered on Vero cells. Vesicular Stomatitis Virus (VSV, Indiana strain) was propagated and titered on U3A cells.

Cells were infected at a multiplicity of infection (MOI) of 5 for indicated times. Virus infections were performed in serum-free medium (SFM), with 1% bovine serum albumin (BSA) supplemented for influenza virus. At 1 hour post-infection, the inoculation medium was replaced with medium containing 2% Cosmic calf serum for the remainder of the infection.

Cytokine treatment and poly(I:C) transfection

For direct treatment with IFN α , cells were treated with 1000 U/mL recombinant IFN α (Roche) for indicated times. For poly(I:C) transfection, 2.5 μ g/mL of low and high molecular weight poly(I:C) (Invivogen) was transfected for indicated times using the Lipofectamine 2000 (Invitrogen, Thermo Fisher Scientific) transfection reagent.

RNA isolation and gene expression analysis by RT-qPCR

Total RNA was extracted using the TRIzol Reagent (Invitrogen, Thermo Fisher Scientific) and treated with DNase I (Invitrogen, Thermo Fisher Scientific). The RNA quantity was obtained using the NanoDrop 2000 instrument (Thermo Fisher Scientific). Total RNA (1 μ g) was random primed and reverse transcribed using SuperScript III Reverse Transcriptase (Invitrogen, Thermo Fisher Scientific).

Gene expression was measured by quantitative real-time PCR (qPCR) and normalized to glyceraldehyde 3-phosphate dehydrogenase (GAPDH) to determine relative abundance by the $2^{-\Delta\text{CT}}$ method or fold change over mock by the $2^{-\Delta\Delta\text{CT}}$ method

(Schmittgen and Livak, 2008). All primers are listed in Appendix B. Data are representative of ≥ 2 replicate experiments and plotted as average value of technical replicates ($n=3$) with error bars representing standard deviation in technical replicates. Statistical analysis was done using a two-tailed Student's t-test.

Immunoblotting

Cell extracts for immunoblotting were prepared by lysing samples in whole cell extract buffer (WCEB) (50 mM Tris, pH 8.0, 280 mM NaCl, 0.5% IGEPAL, 0.2 mM EDTA, 2 mM EGTA, 10% glycerol, 1 mM DTT, 0.1 mM NaVO₄) supplemented with protease inhibitor cocktail set III (MilliporeSigma). Proteins were separated by SDS-PAGE, transferred to nitrocellulose, probed with antibodies and visualized by chemiluminescence (PerkinElmer, Inc.). The antibodies used were anti-FLAG (MilliporeSigma, cat #F3165), anti-HA (MilliporeSigma, cat #H3663), anti-GAPDH (Santa Cruz Biotechnology, cat # sc-47724), anti-MDA5 (monoclonal antibody generated at Northwestern University (Bamming and Horvath, 2009)), anti-phospho-IRF3 (Ser396, Cell Signaling, cat #4947), and anti-IRF3 (Santa Cruz Biotechnology, cat #sc-9082).

METHODS SPECIFIC TO CHAPTER II

RNA-sequencing and transcript assembly

Duplicate samples of mock-infected and Sendai virus-infected cells for 6 hours were used to prepare RNA libraries for sequencing on Illumina HiSeq2000 platform (Illumina) to generate 100 bp paired-end sequencing reads. Raw data was filtered to remove adapter

sequences and low-quality reads. The remaining rRNA reads were removed by mapping to known human rRNA sequences. The clean, high-quality data was mapped to the human reference genome (GRCh37/hg19) (Lander et al., 2001) using TopHat (Trapnell et al., 2009) (2.0.10). The mapped reads for each sample were independently assembled into annotated and novel transcripts using the Cufflinks (Trapnell et al., 2010) (2.1.1) suite of programs.

Bioinformatic analyses

To determine differential expression between mock-infected and Sendai virus-infected samples, read counts for each RNA were generated by HTSeq (Anders et al., 2015) program (0.6.0) for each sample and analyzed by the DESeq2 (Love et al., 2014) program (1.2.10). Significance was calculated using the Wald test and a Benjamini-Hochberg False Discovery Rate (FDR) cut-off of 5% was used to assess statistically significant differential expression. The lowest quartile of RNAs based on expression were excluded from further analysis.

RefSeq genomic feature distribution information for coding exons, introns, 3' and 5' untranslated regions (UTRs), promoters (-1kb), and transcription termination sites (+1 kb) was downloaded from the UCSC genome browser (<http://genome.ucsc.edu/>) (Karolchik et al., 2014) and analyzed using the BEDtools program (Quinlan and Hall, 2010).

The online functional annotation tool, DAVID (Huang et al., 2009a, b) (6.8), was used to conduct gene enrichment analysis. A list of gene symbols corresponding to differentially expressed RNAs was mapped to DAVID gene IDs to determine which Gene Ontology biological processes and KEGG pathways were enriched. A Benjamini-Hochberg FDR cut-off of 5% was used to assess statistical significance. Clustering analysis of enriched GO biological processes was done using DAVID's heuristic fuzzy multiple-linkage partitioning method (Huang et al., 2007). An enrichment score cut-off of 3 was used to determine significance.

Gene lists from 11 SABiosciences PCR arrays for various metabolic processes were compiled into a single list of 597 unique genes. Expression changes of these genes were analyzed from RNA-sequencing data of Sendai virus-infected Namalwa cells. The PCR array lists used were: Mitochondria, Mitochondrial Energy Metabolism, Mitochondrial Energy Metabolism Plus, Oxidative Stress, Oxidative Stress Plus, Drug Metabolism, Drug Metabolism: Phase I Enzymes, Drug Metabolism: Phase II Enzymes, Amino Acid Metabolism I, and Amino Acid Metabolism II.

Base-by-base PhastCons (Siepel et al., 2005) conservation scores across 100 vertebrates were downloaded from the UCSC Genome Browser. For each RNA, a mean score was calculated if there was a score available for at least 80% of the bases in the sequence.

The PhyloCSF (Lin et al., 2011) software was locally installed and used to determine the coding potential of the longest start-to-stop open reading frame in each RNA. The multiple-species alignments needed for this analysis were prepared on Galaxy web platform at usegalaxy.org (Afgan et al., 2016).

Combining RNA-sequencing and ChIP-sequencing data

Sendai virus-regulated transcription factor occupancy of IRF3, p65, and total RNA Pol II in Namalwa cells was determined by ChIP-sequencing (Freaney et al., 2013). The genomic loci of transcription factor binding peaks that had at least a 2-fold enrichment following virus infection were compared to the RNA-encoding virus-inducible loci identified in the RNA-sequencing analysis. A transcription factor was bound nearby if it was within 5 kb of the transcript locus.

Lentivirus growth and transduction

Lentiviruses expressing non-silencing shRNA and AFF1-AS1-targeting shRNA from the pGIPz plasmid were harvested from 293T cells. Briefly, 293T cells were transfected with pCMV Δ R8.91, pUC-MDG, and pGIPz plasmids using polyethylenimine (PEI) in serum-free media. After 15 hours, the media was replaced with media supplemented with 10% Cosmic calf serum and 1% penicillin-streptomycin. The lentivirus-containing supernatant was harvested every 24 hours for the next 3 days. 2fTGH cells were transduced with this supernatant and puromycin-selected (2.5 μ g/mL) for 3 days to isolate shRNA-expressing cells.

siRNA transfection

Non-targeting and ZBED5-AS1-targeting SMARTpool siRNA (Dharmacon) were reverse transfected into 2fTGH cells at 25 nM in 6-well tissue culture plates using the Lipofectamine 2000 reagent (Invitrogen, Thermo Fisher Scientific). Cells were incubated for 48 hours before additional treatment or harvest.

Data availability

RNA-sequencing data is available in the GEO database under study GSE115266.

METHODS SPECIFIC TO CHAPTER III**Plasmids**

Wild type MDA5 was cloned into the p3xFLAG-CMV-10 vector as previously described to provide an N-terminal FLAG epitope tag (Rodriguez and Horvath, 2014). Point mutations in this plasmid were created using the QuikChange Lightning Site-Directed mutagenesis kit (Agilent Technologies). Paramyxovirus V proteins from PIV5, mumps virus, measles virus, Nipah virus and Hendra virus were cloned into the pEF-HA vector as previously described to provide an N-terminal HA epitope tag (Ramachandran et al., 2008; Rodriguez et al., 2002; Rodriguez et al., 2003; Ulane et al., 2005).

Luciferase Reporter Gene Assay

HEK293T cells were seeded in 24-well tissue culture plates and transfected using the Lipofectamine 2000 reagent (Invitrogen, Thermo Fisher Scientific). The -110 IFN

promoter-driven firefly luciferase reporter was co-transfected with a Renilla luciferase plasmid for signal normalization. The MDA5 plasmids were transfected at 25 ng/well and the V protein plasmids at 75 ng/well. Luciferase activity was measured 24 hours after transfection with the Dual Luciferase Reporter Assay (Promega Corporation). Data are plotted as average values ($n \geq 3$) with error bars representing the standard deviation. Statistical analysis was done using a two-tailed Student's t-test.

Co-immunoprecipitation

HEK293T cells were seeded in 10 cm tissue culture plates and transfected using standard calcium phosphate procedures. Each plate was transfected with 5 μ g of the MDA5 plasmids along with 5 μ g of each V protein. In some co-precipitation experiments, V proteins that could readily be distinguished by their SDS-PAGE electrophoretic mobilities were multiplexed, ie. measles virus V (33 kDa) and PIV5 virus V (25 kDa) or Hendra virus V (51 kDa) and mumps virus V (25 kDa). Lysates for immunoprecipitation were prepared in WCEB. After preclearing with Sepharose beads, 10% of the clarified lysate was analyzed directly and the rest was subjected to immunoprecipitation. Protein complexes were purified by overnight incubation with the FLAG M2 affinity resin (MilliporeSigma) and washed with WCEB. After elution with SDS, proteins were analyzed by immunoblotting.

Antiviral cytopathic effect assay

Cytopathic effect assays were carried out essentially as described (Bamming and Horvath, 2009). HEK293T cells were seeded in 10-cm tissue culture plates and

transfected using standard calcium phosphate procedures. Supernatants from HEK293T cells transfected with various MDA5 proteins alone and along with PIV5 V protein were harvested and used to treat 2fTGH cells. As a control, cells were also directly treated with recombinant IFN α (Roche). Following an 8-hour incubation, 2fTGH cells were infected with VSV at 6×10^3 plaque forming units (PFU)/well in serum-free media for 1 hour. Media was changed and cells incubated for 17 hours before fixing with 4% formaldehyde and staining with 0.3% crystal violet in 50% ethanol.

METHODS SPECIFIC TO APPENDIX D

Overall protocol summary

Day 1: A549 cells are plated in DMEM supplemented with 1% Cosmic calf serum in 384-well plate (3500 cells in 25 μ L per well)

Day 2: Cells are pre-incubated with compounds for 1 hour and then mock or Sendai virus-infected with 3000 PFU/well (MOI \sim 0.86)

Day 3: After 24-hour infection, supernatant media is harvested and analyzed for Lucia luciferase and SEAP.

Z-factor calculation

$$Z' = 1 - \frac{3\sigma_p + 3\sigma_n}{|\mu_p - \mu_n|}$$

APPENDIX B: PRIMERS USED FOR GENE EXPRESSION ANALYSIS

Table B.1 – Primers used for gene expression analysis

Gene	Forward Primer	Reverse Primer
GAPDH	5'-ACAGTCAGCCGCATCTTCTT-3'	5'-ACGACCAAATCCGTTGACTC-3'
AFF1-AS1	5'-GGCCTGTTTCATCTTCACAAC-3'	5'-TAGGAGATATTTTCGGCGCTTC-3'
ZBED5-AS1	5'-ACGCTTGTGCGAGAATGCAGAT-3'	5'-GAGAATGGGATGGGAGGACAC-3'
CCL5	5'-CTGCTTTGCCTACATTGCC-3'	5'-TCGGGTGACAAAGACGACTG-3'
CSAG3	5'-GCCATTGTCCAACAACCAC-3'	5'-TGGGAACCTTTGATAGGGCTTCT-3'
IFI6	CTGTGCCCATCTATCAGCAG	5'-GCAGGTAGCACAAAGAAAAGC
IFIT1	5'-CAGAACGGCTGCCTAATTT-3'	5'-GGCCTTTCAGGTGTTTCAC-3'
IFNβ (Chapter 2)	5'-CATTACCTGAAGGCCAAGGA-3'	5'-CAATTGTCCAGTCCCAGAGG-3'
IFNβ (Chapter 3)	5'-ACGCCGCATTGACCATCTAT-3'	5'-AGCCAGGAGGTTCTCAACAA-3'
ISG15	5'-GACCTGACGGTGAAGATGCT-3'	5'-CGATCTTCTGGGTGATCTGC-3'
USP18	5'-CCCCGGCAGATCTTGAAGAAA-3'	5'-CAACCAGGCCATGAGGGTAG-3'
PTPN7	5'-GCTTCCTGGAGCCTTCTCAG-3'	5'-CAGCCATGAGGTCTGCTGAA-3'
BCL7A	5'-CGCAGAGACGTCTGCAATCT-3'	5'-TCCAAATCCCCGACAACC-3'
DUSP2	5'-CTGCGAGGAGGCTTCGAC-3'	5'-GCTGGTTTTGTCCCCTGTTG-3'
RSAD2	5'-CCTGTCCGCTGGAAAGTGTT-3'	5'-GACACTTCTTTGTGGCGCTC-3'
XLOC_035959	5'-AGCTGGTAAGACATCGCTGG-3'	5'-AGCTGGTAAGACATCGCTGG-3'
XLOC_036708	5'-AGGGAAGAATGTGCAGTGCT-3'	5'-AGGGAAGAATGTGCAGTGCT-3'
XLOC_001887	5'-CAACACCAGCGGATGAAAGC-3'	5'-CAACACCAGCGGATGAAAGC-3'
XLOC_018179	5'-CCTGGGGAAACTGAGTGCTC-3'	5'-CCTGGGGAAACTGAGTGCTC-3'
XLOC_021911	5'-CTCGGACTGCGAGATGGAAT-3'	5'-CTCGGACTGCGAGATGGAAT-3'
XLOC_000681	5'-GTTGGTAAACGCTGTGTGGT-3'	5'-GTTGGTAAACGCTGTGTGGT-3'
XLOC_001629	5'-TGTGCAGCAACTCCTACTTGT-3'	5'-TGTGCAGCAACTCCTACTTGT-3'
XLOC_001875	5'-AAGATGCCGGGAAAAGGAGG-3'	5'-AAGATGCCGGGAAAAGGAGG-3'
XLOC_029768	5'-ACCACGGAACAGAAGAGCAG-3'	5'-ACCACGGAACAGAAGAGCAG-3'
XLOC_002741	5'-AGCCCATTACCATGTAAGCA-3'	5'-AGCCCATTACCATGTAAGCA-3'
XLOC_029636	5'-CCCACAGGCGTCTTGTAT-3'	5'-CCCACAGGCGTCTTGTAT-3'
XLOC_013418	5'-CCTGTTGAGTGTTGTTGCTGA-3'	5'-CCTGTTGAGTGTTGTTGCTGA-3'
XLOC_024204	5'-CTCCGTCAGTGTCTCCGTC-3'	5'-CTCCGTCAGTGTCTCCGTC-3'
XLOC_027202	5'-CTTCAAGGGCCTGCTACTGT-3'	5'-CTTCAAGGGCCTGCTACTGT-3'
XLOC_033596	5'- TCAAATTTGTTAAGATGTTGCCCA-3'	5'- TCAAATTTGTTAAGATGTTGCCCA-3'
XLOC_031219	5'-TGGCACAGAGGATTAAGCTCA- 3'	5'-TGGCACAGAGGATTAAGCTCA- 3'

XLOC_031822	5'- TGGGGAGATTAATAAGCCTCAGG-3'	5'- TGGGGAGATTAATAAGCCTCAGG-3'
XLOC_014649	5'-TTGGGTGGTAAAGATGCCCC-3'	5'-TTGGGTGGTAAAGATGCCCC-3'
XLOC_033830	5'-ACGCATTGCTGAGGACTGAC-3'	5'-ACGCATTGCTGAGGACTGAC-3'
XLOC_028329	5'-CTGCCTCAGTCAAAGCCTTT-3'	5'-CTGCCTCAGTCAAAGCCTTT-3'
XLOC_028240	5'-CTGGGGATTGTGTTGGCTCC-3'	5'-CTGGGGATTGTGTTGGCTCC-3'
XLOC_001127	5'-GCAAGAGGACAGAAGGATCCC-3'	5'-GCAAGAGGACAGAAGGATCCC-3'
XLOC_037108	5'-AAACAGACTGTCTACACCATGGT- 3'	5'-AAACAGACTGTCTACACCATGGT- 3'
XLOC_004672	5'-AGCCCATTTCTTCCTGCTC-3'	5'-AGCCCATTTCTTCCTGCTC-3'
XLOC_014660	5'-GGGAGGCAGAAGAGATGAGC-3'	5'-GGGAGGCAGAAGAGATGAGC-3'
XLOC_033807	5'-TGCAGTACAAACCTCCTGCA-3'	5'-TGCAGTACAAACCTCCTGCA-3'
XLOC_036689	5'-TGTCCTCAAATTTAAGGCAGGA- 3'	5'-TGTCCTCAAATTTAAGGCAGGA- 3'
XLOC_038802	5'-CGAGGAACAGCTAAGGTCCC-3'	5'-CGAGGAACAGCTAAGGTCCC-3'
XLOC_009436	5'-TCAGCATGAGACAGCCAACC-3'	5'-TCAGCATGAGACAGCCAACC-3'
XLOC_005868	5'-TGTCTGGTGCAGCAGAAGTC-3'	5'-TGTCTGGTGCAGCAGAAGTC-3'
XLOC_011151	5'-TGAGAAGAGTTCAGAGGAGGGA- 3'	5'-TGAGAAGAGTTCAGAGGAGGGA- 3'
XLOC_027917	5'-TGGTTCCAGGCATGATTCAGT-3'	5'-TGGTTCCAGGCATGATTCAGT-3'
XLOC_013194	5'-GTGCACTCTACCCAGGCAG-3'	5'-GTGCACTCTACCCAGGCAG-3'
XLOC_012443	5'-GGCACATGTTTACTCAGCTGT-3'	5'-GGCACATGTTTACTCAGCTGT-3'
XLOC_037958	5'-CCATGACGAGAGCTTGCAGA-3'	5'-CCATGACGAGAGCTTGCAGA-3'
XLOC_010526	5'-CAACACGGCTAATGGGTGC-3'	5'-CAACACGGCTAATGGGTGC-3'

**APPENDIX C: BIOLOGICAL REPLICATES OF AFF1-AS1 AND ZBED5-AS1
KNOCKDOWN EXPERIMENTS**

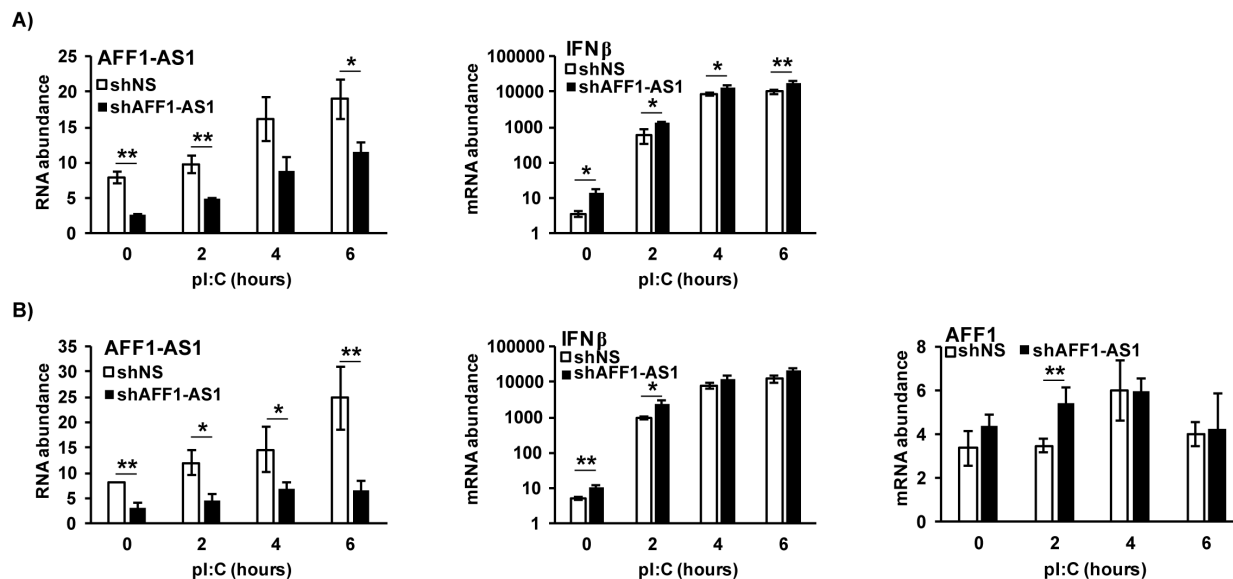


Figure C.1 – Biological replicates of AFF1-AS1 knockdown by shRNA in 2fTGH cells
 Cells transduced with non-silencing (shNS) or AFF1-AS1 targeting shRNA (shAFF1-AS1) were transfected with poly(I:C) for the indicated times in independent experiments shown in A) and B). Total RNA was harvested and expression of AFF1-AS1, IFN β and AFF1 was analyzed by RT-qPCR. Data are shown normalized to GAPDH expression. Bars indicate average values of technical replicates (n=3) with error bars representing standard deviation. Statistical analysis was done using a two-tailed Student's t-test (* p-value < 0.05, ** p-value < 0.005). (Related to Figure 2.16A-B)

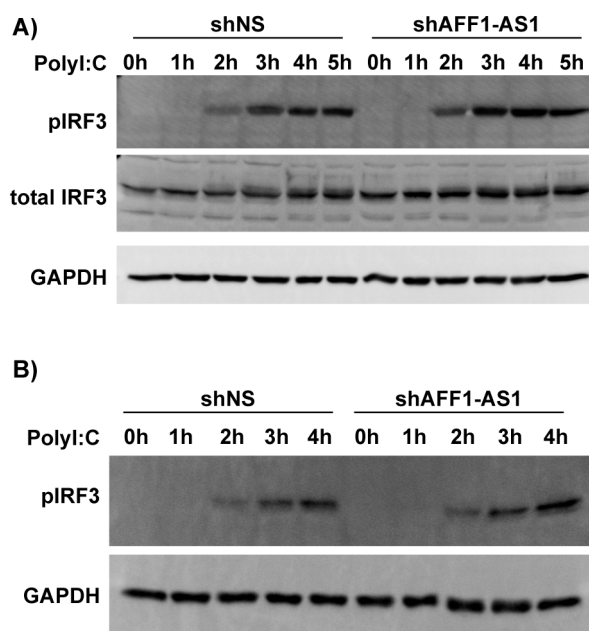


Figure C.2 - Biological replicates of AFF1-AS1 knockdown by shRNA in 2fTGH cells
Cells transduced with non-silencing (shNS) or AFF1-AS1 targeting shRNA (shAFF1-AS1) were transfected with poly(I:C) for the indicated times in independent experiments shown in A) and B). Whole cell lysates were collected at the indicated time points and analyzed by immunoblotting with antibodies against phospho-IRF3 (pIRF3), total IRF3 and GAPDH. (Related to Figure 2.16C)

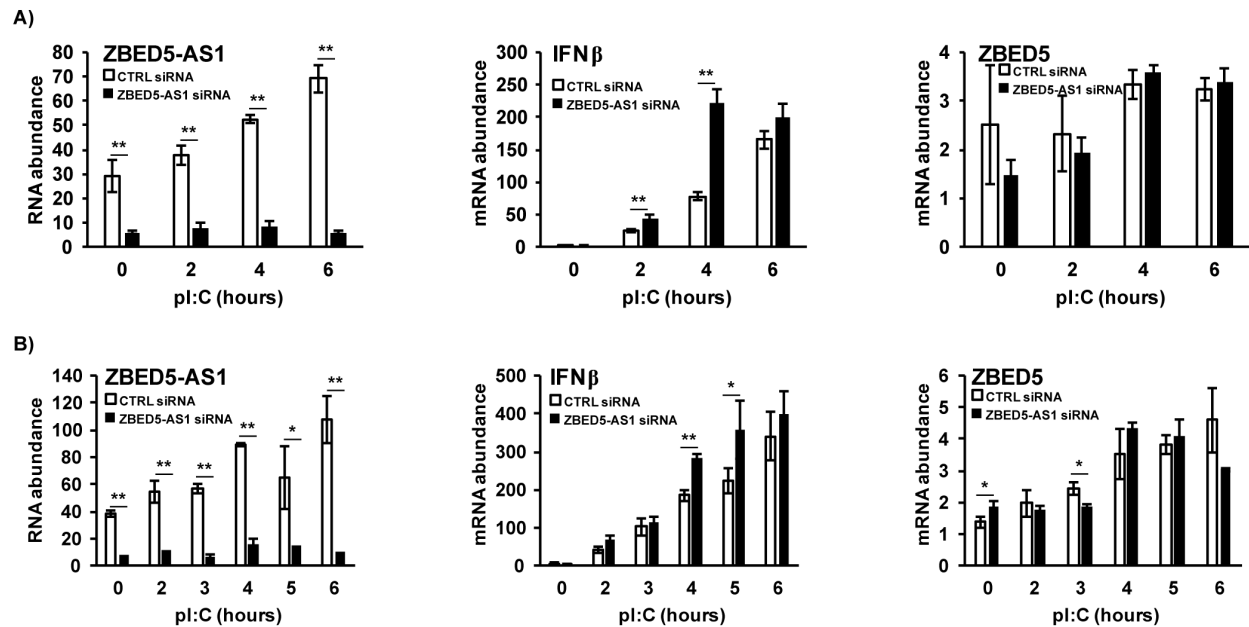


Figure C.3 – Biological replicates of ZBED5-AS1 knockdown by siRNA in 2fTGH cells

Cells transfected with control or ZBED5-AS1 targeting siRNA were transfected with poly(I:C) for the indicated times in independent experiments shown in A) and B). Total RNA was harvested and expression of ZBED5-AS1, IFN β and ZBED5 was analyzed by RT-qPCR. Data are shown normalized to GAPDH expression. Bars indicate average values of technical replicates (n=3) with error bars representing standard deviation. Statistical analysis was done using a two-tailed Student's t-test (* p-value < 0.05, ** p-value < 0.005). (Related to Figure 2.17)

**APPENDIX D: ASSAY DEVELOPMENT FOR IDENTIFICATION OF SMALL
MOLECULE REGULATORS OF CELLULAR INNATE IMMUNE SIGNALING**

INTRODUCTION

Type I interferon (IFN) is the primary antiviral cytokine expressed by the innate immune system following detection of non-self nucleic acids from viruses or genetic abnormalities. Excessive or inappropriate expression of IFN and its downstream effectors is linked to several diseases such as systemic lupus erythematosus (SLE) (Baechler et al., 2003; Bennett et al., 2003; Crow and Wohlgemuth, 2003), rheumatoid arthritis (van der Pouw Kraan et al., 2007), type I diabetes (Ferreira et al., 2014), and Aicardi-Goutieres Syndrome (AGS) (Crow et al., 2015; Rice et al., 2013). The connection between SLE and IFN has been established by detection of circulating IFN in patients as well as a chronic signature of upregulated interferon stimulated gene (ISG) expression (Baechler et al., 2003; Bennett et al., 2003; Crow and Wohlgemuth, 2003; Hooks et al., 1979). SLE and the other immune disorders that share the characteristics of increased IFN and ISG expression are collectively known as type I interferonopathies (Crow, 2011). Though these diseases are caused by a variety of genetic variants and monogenic mutations, targeting IFN could provide a common therapeutic for these heterogeneous disorders. Type I IFN expression is induced by the master innate immune transcription factors NF- κ B and IRF3. In chapter 3 of this thesis, a virus-encoded inhibitor of IFN production and signaling was studied as a potential therapeutic. In this section, we have developed a high-throughput screen (HTS) to identify small molecule compounds that could regulate NF- κ B and IRF3 signaling and provide leads for the development of therapeutics for type I interferonopathies.

DEVELOPMENT OF HIGH-THROUGHPUT SCREEN

Cell line and detection reagents

The A549-Dual cell line is a human epithelial cell-derived reporter cell line that stably expresses inducible reporters for the master immune transcription regulators NF- κ B and IRF3. The cells express secreted embryonic alkaline phosphatase (SEAP) under the control of a minimal promoter fused to NF- κ B binding sites and a secreted Lucia luciferase under the control of a minimal promoter fused to IRF3 binding sites. These reporters can be induced by RNA virus infection as shown in Figure D.1.

The SEAP is measured by a colorimetric enzyme assay and the Lucia luciferase is measured using the coelenterazine substrate in a luciferase reporter assay. While the Invivogen QUANTI-Blue reagent sold with the cell line for SEAP detection generates a stable signal and can be used effectively in HTS assays, the QUANTI-Luc reagent for Lucia luciferase detection could not be used as the half-life of the signal is only 5 minutes. Instead two alternate reagents were tested: Pierce Gaussia Luciferase Glow and Promega Renilla-Glo. Both reagents generate signals with a half-life greater than 1 hour, which is more suitable for a HTS assay.

Determining assay conditions

A preliminary experiment was done to determine optimal experimental conditions for cell density, amount of virus and length of infection to obtain the most robust signal from both reporters. The RNA virus Sendai virus, which is known to potently induce both NF- κ B and

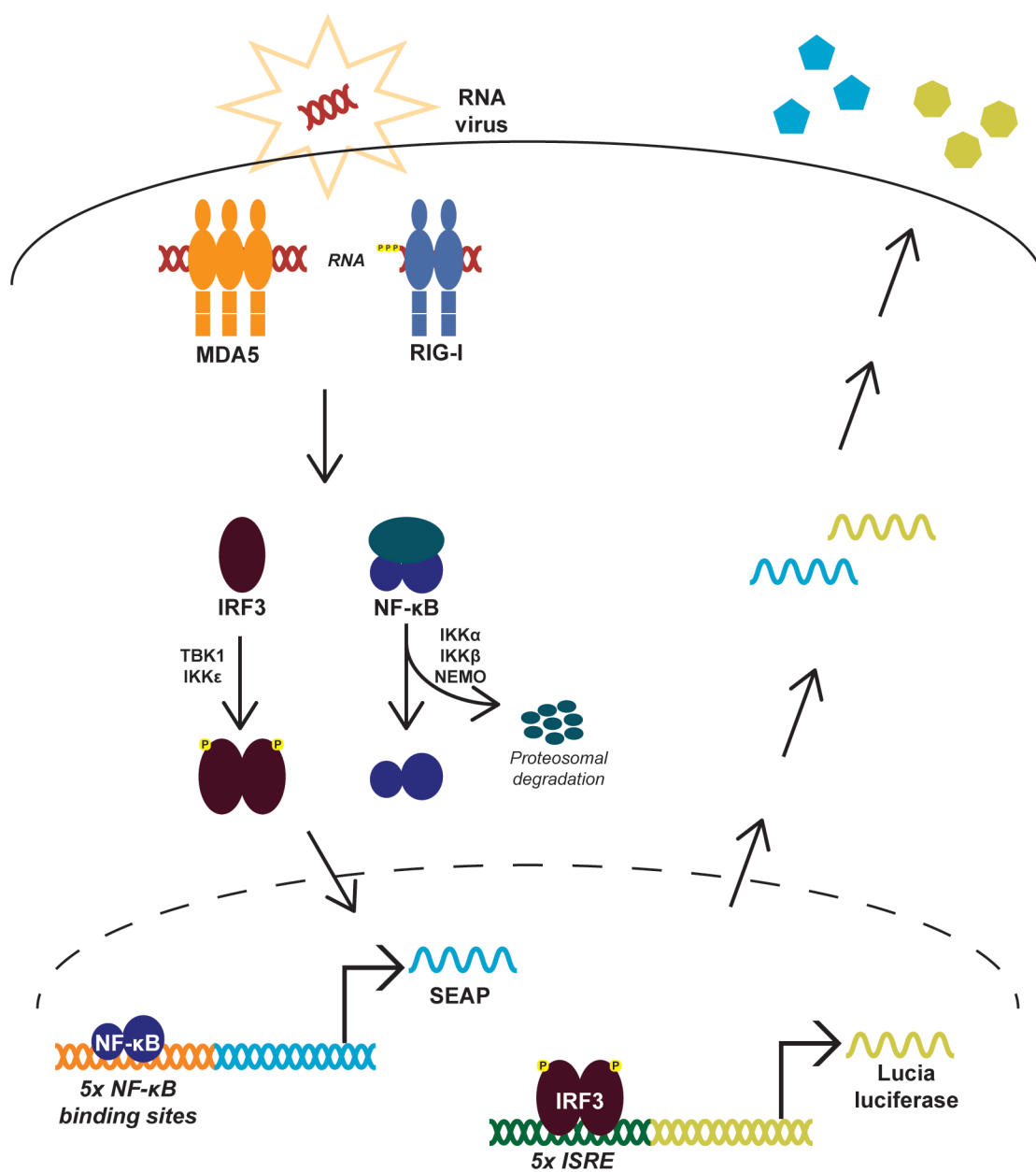


Figure D.1 – Illustration of A549-Dual cell reporter induction

A549-Dual cells stably express inducible NF-κB and IRF3 reporters. Both reporters can be simultaneously induced by RNA virus infection as shown in this diagram. The reporters are secreted by the cells into the supernatant media.

IRF3, was used to stimulate the expression of both reporters in A549-Dual cells. The original QUANTI-Blue and QUANTI-Luc reagents were used for this pilot experiment. A549-Dual cells were plated at various densities ranging from 1500 cells/well to 4500 cells/well in 384-well plates. Cells were infected with 0-18,000 PFU of Sendai virus per well for either 6 or 24 hours with 4 replicates for every condition. Though a 6-hour Sendai virus infection induces high IFN expression indicating NF- κ B and IRF3 activation, in this assay a 6-hour infection did not allow enough time for accumulation of either reporter at a detectable level (Fig. D.2). The 24-hour infection gave a much higher signal for both reporters across all conditions. From this pilot experiment, we concluded that a 24-hour infection of 3500 cells/well infected with 3000 PFU of Sendai virus/well would give an optimal signal for both reporters.

Determining assay quality

A second pilot experiment was conducted to determine the robustness of both reporters under the previously determined conditions. The Z-factor is a numeric quantification of the quality of a HTS assay based on the dynamic range of the signal measured as well as the signal variability (Zhang et al., 1999). A Z-factor below 0 indicates an unsuitable assay while a Z-factor greater than 0.5 indicates an excellent assay.

Supernatant media from mock- and Sendai virus-infected samples was used in the pilot experiment to quantify the quality of the HTS assay. The NF- κ B signal measured with the QUANTI-Blue reagent was very reproducible with little variability among the replicates

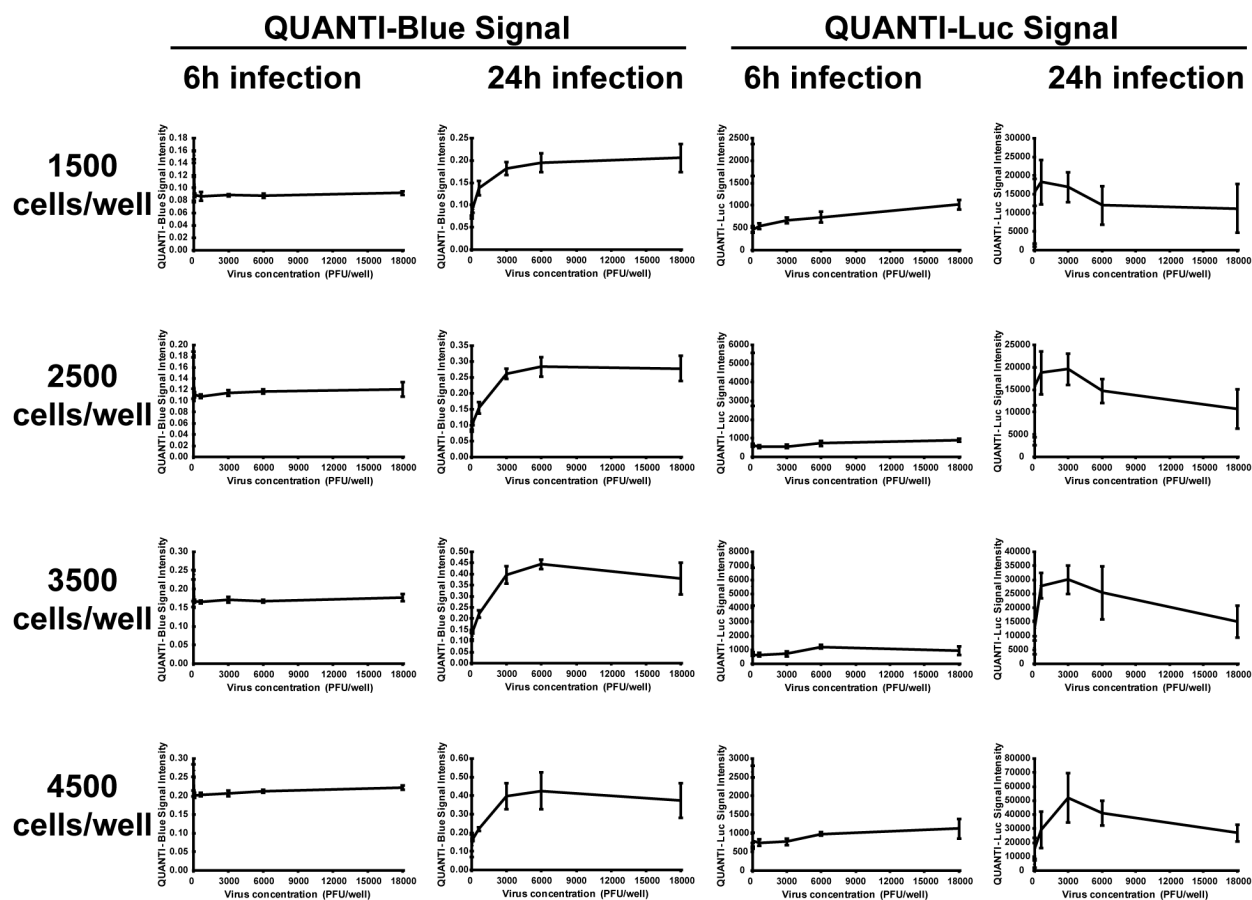


Figure D.2 – Preliminary experiment to determine assay conditions

A549-Dual cells were plated at various cell densities in 384-well plates and infected with 0-18,000 PFU/well Sendai virus for 6 or 24 hours. The supernatant from each well was collected and analyzed using the QUANTI-Blue and QUANTI-Luc reagents to quantify NF- κ B and IRF3 signaling, respectively. Error bars represent standard deviation between replicates (n=4).

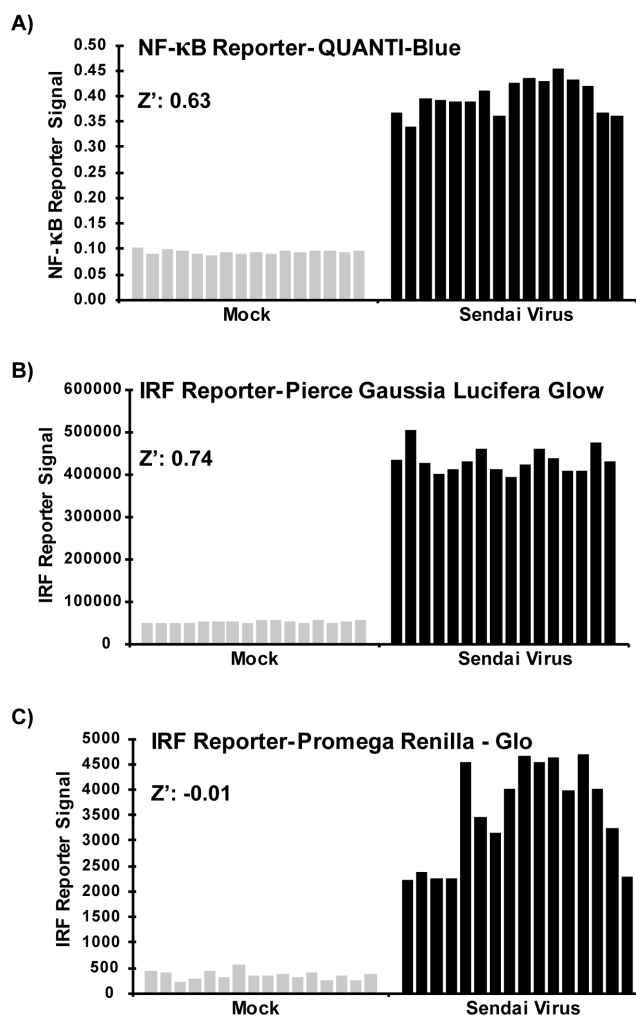


Figure D.3 – Determining HTS assay quality

Supernatant media containing secreted reporter proteins from mock- and Sendai virus-infected cells was assayed for A) NF- κ B signaling with the QUANTI-Blue reagent, and for IRF3 signaling with the B) Pierce Gaussia Luciferase Glow reagent and C) the Promega Renilla-Glo reagent. Z-factors were calculated for both reporters using this data.

resulting in a Z-factor of 0.63 (Fig. D.3A). The IRF3 signal was measured by two different reagents that generate a stable signal with a half-life greater than 1 hour: Pierce Gaussia Luciferase Glow and Promega Renilla-Glo Luciferase. The Gaussia Luciferase Glow reagent generated a consistent signal between all replicates and had a Z-factor of 0.74 (Fig. D.3B). However, the IRF3 signal measured with the Renilla-Glo reagent was variable, resulting in an unfavorable Z-factor of -0.01 (Fig. D.3C). The high Z-factors for the QUANTI-Blue and Gaussia Luciferase Glow reagents indicate that the assays for both pathways are extremely robust.

HIGH-THROUGHPUT SCREEN TO IDENTIFY COMPOUNDS OF INTEREST

Once the assay conditions were established and verified, the Spectrum Collection of known biologics was screened to identify potential regulators of innate immune signaling. After an initial screen of the entire Spectrum library, 480 compounds were selected for further analysis. These compounds were tested in duplicate to determine whether they altered NF- κ B or IRF3 signaling without cellular toxicity. Hoechst staining was used as a proxy for cell viability. The signals from the reporters as well as the measured nucleus size of each sample were normalized to cells infected with Sendai virus and mock-treated with a compound.

Compounds that did not affect cell viability resulted in nucleus sizes that were 80-120% of the control (Fig. D.4). However, several compounds were toxic to the cells, dramatically reducing the nucleus size. There was a lot of variability in the effect of the compounds on

reporter signaling, but most of them only had modest effects. Fifty compounds of interest were chosen for further analysis (Fig. D.4, dark blue and yellow dots; Table D.1). The compounds were chosen based on the following criteria: 1) no effect on cell viability; 2) reduction of the signal of either reporter below 50% or increase of the signal of either reporter above 115%; and 3) low deviation between duplicates. The fifty compounds of interest were tested in a counterscreen to verify that the compounds do not interfere with the assay or reagents used. Supernatant media from virus-infected cells was incubated with the compounds and then used in the assays, which demonstrated that the compounds do not directly affect the measurements. The selection of these fifty compounds for further analysis not only identifies leads for therapeutic potential, but importantly also establishes proof of concept for the HTS assay.

Titration analysis of compounds of interest

The fifty selected compounds were further analyzed in a dose-response assay. Compounds were not only analyzed in virus-infected cells but also in mock-infected cells to determine if they also affected basal signaling of the two reporters. The data are displayed as percent of control (Table D.2). None of the compounds dramatically affected basal signaling of either reporter, but most of them did alter virus-induced signaling.

Future directions

In order to determine a possible mechanism of action for each of these compounds, the assay will be repeated with stimuli other than Sendai virus. IL-1 β , TNF α , IFN α , and IFN γ

will be used as they should only activate 1 reporter each. This should enable clarification about where in the pathway each compound is interfering.

One caveat of these compounds is that almost all of them are flavonoid or steroid derivatives, which can be promiscuous molecules. However, now that the HTS assay conditions have been established, the assay can also be used to screen other libraries that have more promising drug candidates. A small subset of the DCL library, consisting of drug-like compounds with a wide variety of structures, was also tested with this assay as a proof of principle. Several compounds inhibited reporter signal indicating that this library could be promising.

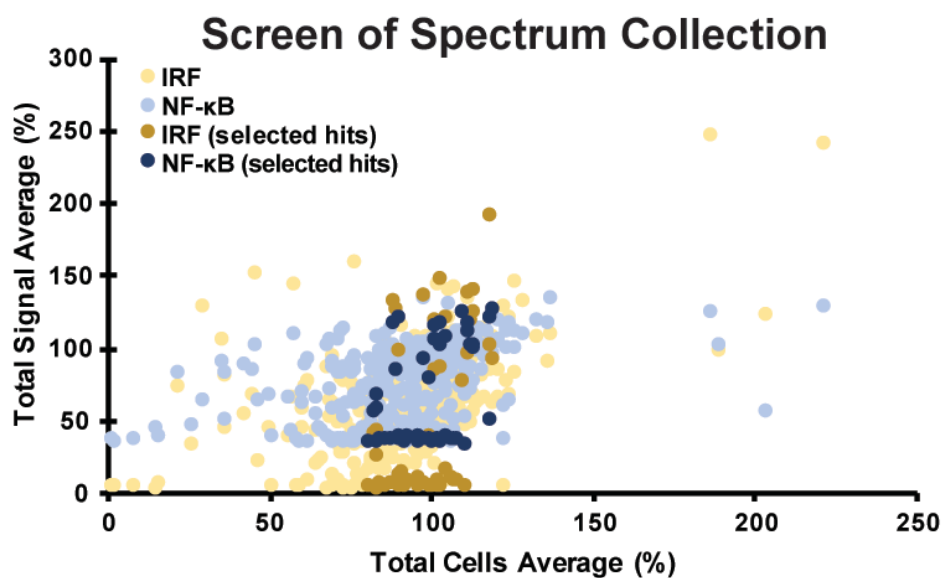


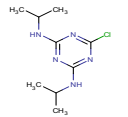
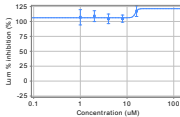
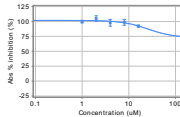
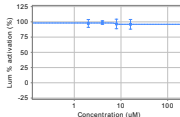
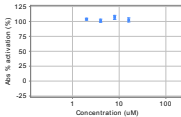
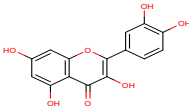
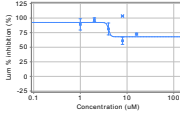
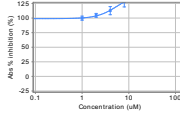
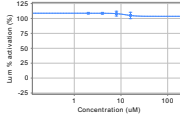
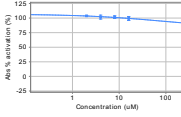
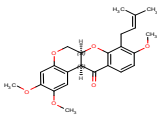
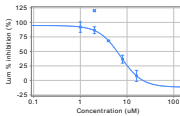
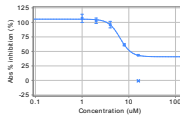
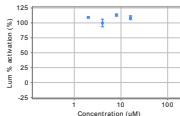
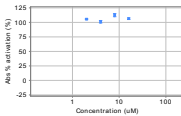
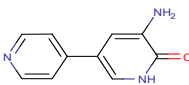
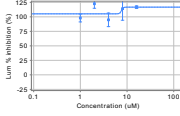
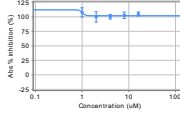
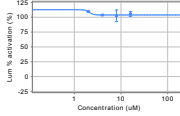
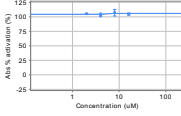
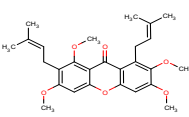
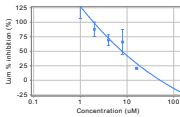
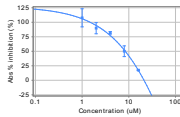
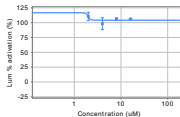
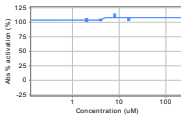
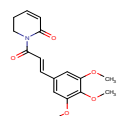
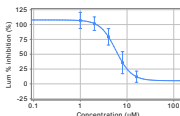
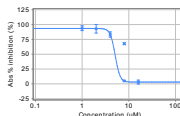
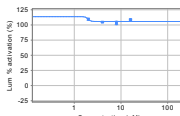
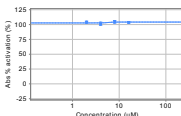
Figure D.4 – Screen of Spectrum Collection

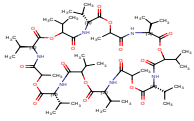
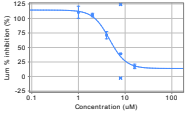
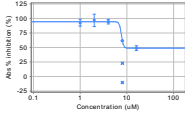
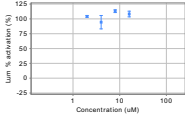
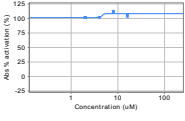
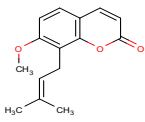
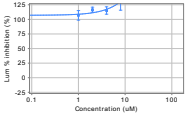
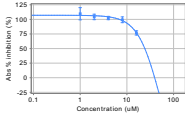
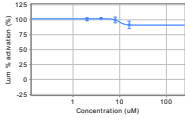
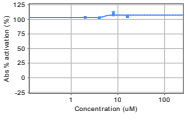
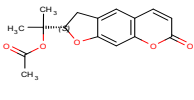
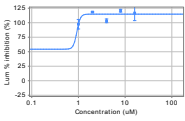
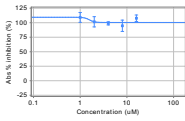
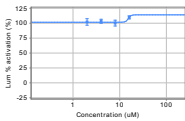
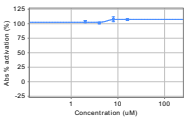
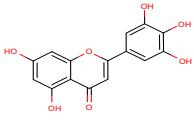
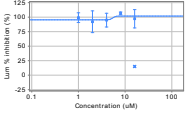
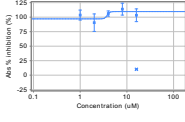
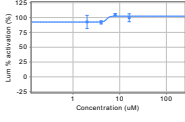
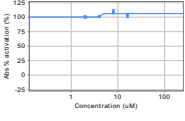
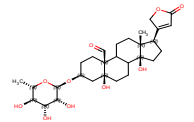
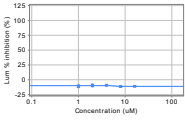
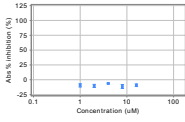
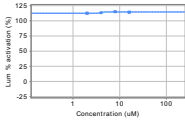
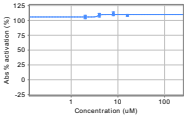
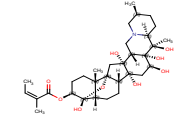
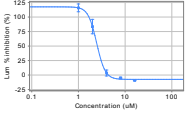
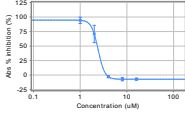
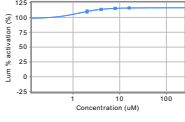
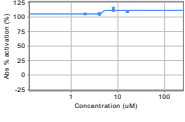
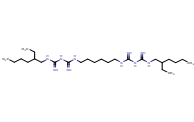
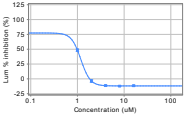
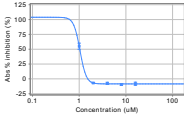
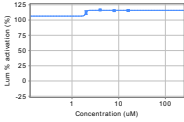
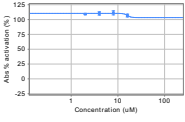
Compounds from the Spectrum Collection were screened in A549-Dual cells to identify compounds that affect NF-κB (blue dots) or IRF3 signaling (yellow dots). For each sample, cell viability was assessed by Hoescht staining. The average signal for each sample is plotted and the deviation between duplicates is not shown for clarity.

Table D.1 – Results of compounds of interest

Compound Name	IRF3 Signal (%)	NF-κB Signal (%)	Cell Size (%)
NUCC-0075261	4.16 ± 0.37	35.56 ± 0.80	83.46 ± 13.83
NUCC-0075634	4.45 ± 0.24	35.00 ± 0.80	92.03 ± 2.75
NUCC-0075843	4.71 ± 0.01	37.82 ± 1.60	85.98 ± 25.51
NUCC-0076096	4.73 ± 0.29	36.13 ± 0.00	83.56 ± 0.82
NUCC-0075505	4.82 ± 0.19	38.10 ± 1.20	101.55 ± 5.50
NUCC-0076123	5.05 ± 0.57	36.73 ± 2.05	90.06 ± 25.60
NUCC-0076558	5.17 ± 0.20	36.15 ± 0.41	96.21 ± 10.13
NUCC-0075708	5.29 ± 0.26	37.26 ± 3.19	90.16 ± 4.35
NUCC-0075378	5.39 ± 0.12	35.84 ± 0.40	99.92 ± 12.12
NUCC-0051540	5.67 ± 0.29	34.72 ± 0.40	110.65 ± 7.51
NUCC-0075972	5.68 ± 1.17	36.13 ± 0.00	80.86 ± 4.50
NUCC-0076594	5.70 ± 0.01	36.15 ± 0.41	102.46 ± 0.87
NUCC-0076498	5.88 ± 0.42	36.44 ± 0.82	91.94 ± 8.84
NUCC-0076516	5.96 ± 0.35	36.73 ± 0.41	99.02 ± 5.66
NUCC-0076137	5.98 ± 0.64	36.73 ± 2.05	85.52 ± 17.22
NUCC-0076191	6.46 ± 1.04	37.02 ± 0.82	87.57 ± 19.35
NUCC-0076176	6.50 ± 0.73	36.69 ± 0.80	102.65 ± 6.84
NUCC-0076269	7.60 ± 1.29	36.73 ± 2.86	83.86 ± 8.87
NUCC-0075565	7.67 ± 1.65	36.97 ± 2.00	97.26 ± 13.65
NUCC-0051653	8.21 ± 0.49	39.33 ± 1.64	92.97 ± 1.78
NUCC-0076525	9.25 ± 0.19	38.47 ± 0.41	106.68 ± 9.92
NUCC-0076579	9.65 ± 1.01	37.60 ± 0.82	107.74 ± 1.78
NUCC-0076325	9.69 ± 0.18	37.02 ± 0.82	101.59 ± 6.01
NUCC-0075379	11.03 ± 1.09	39.51 ± 0.80	96.40 ± 6.17
NUCC-0076485	11.30 ± 2.82	38.47 ± 0.41	105.89 ± 9.64
NUCC-0075798	12.37 ± 2.54	39.51 ± 0.80	90.37 ± 5.02
NUCC-0075550	15.22 ± 0.77	38.38 ± 0.80	91.06 ± 0.19
NUCC-0076454	15.86 ± 2.93	40.20 ± 0.41	104.85 ± 7.40
NUCC-0076689	26.61 ± 5.46	68.26 ± 3.27	82.90 ± 23.92
NUCC-0076840	39.99 ± 9.49	80.12 ± 8.59	99.02 ± 25.35
NUCC-0076763	41.04 ± 0.77	57.27 ± 1.64	82.50 ± 21.48
NUCC-0076691	42.55 ± 8.38	58.42 ± 6.54	83.00 ± 24.34
NUCC-0075646	77.81 ± 1.11	124.75 ± 3.19	109.94 ± 5.69
NUCC-0075384	85.26 ± 1.31	116.00 ± 0.40	101.34 ± 2.38
NUCC-0075367	87.81 ± 12.81	116.85 ± 7.18	102.42 ± 10.82
NUCC-0076902	91.93 ± 9.62	126.10 ± 20.45	118.83 ± 19.28
NUCC-0075252	96.26 ± 13.61	118.26 ± 5.19	111.38 ± 20.01
NUCC-0075594	97.95 ± 8.02	122.21 ± 4.39	89.82 ± 26.18
NUCC-0076625	103.03 ± 15.75	121.19 ± 5.32	117.99 ± 1.05
NUCC-0076927	118.48 ± 1.67	106.49 ± 4.33	101.03 ± 7.16
NUCC-0076793	119.16 ± 10.59	102.68 ± 11.04	112.51 ± 10.20
NUCC-0076643	121.09 ± 13.66	108.17 ± 4.91	104.53 ± 10.02
NUCC-0076644	125.41 ± 7.85	102.10 ± 2.05	112.93 ± 2.34
NUCC-0050450	127.32 ± 6.30	85.99 ± 6.49	89.05 ± 10.99
NUCC-0075599	133.64 ± 27.92	117.69 ± 5.99	88.69 ± 66.60
NUCC-0075250	136.22 ± 14.15	92.29 ± 5.19	97.89 ± 12.61
NUCC-0076436	138.06 ± 18.44	111.64 ± 5.73	111.37 ± 5.45
NUCC-0050357	139.55 ± 3.00	100.37 ± 5.19	112.85 ± 13.04
NUCC-0050413	148.83 ± 34.57	102.20 ± 22.50	102.90 ± 28.77
NUCC-0076449	191.22 ± 38.22	51.19 ± 6.14	118.19 ± 24.10

Table D.2 – Dose-response analysis of compounds of interest

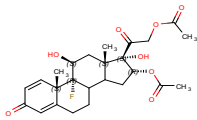
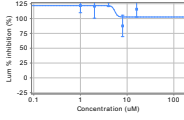
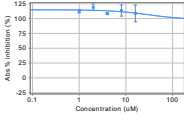
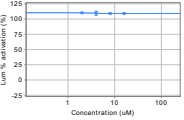
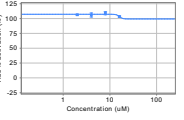
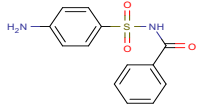
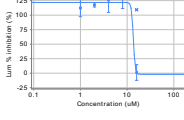
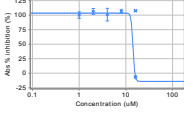
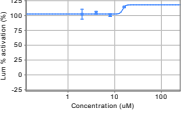
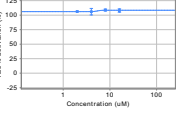
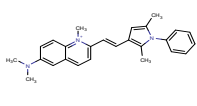
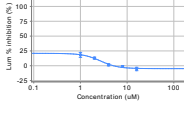
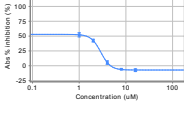
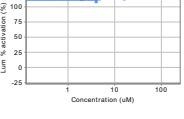
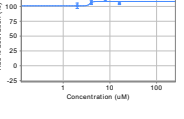
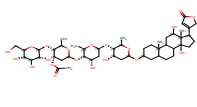
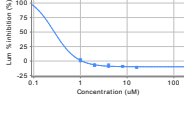
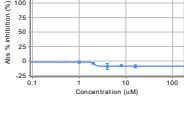
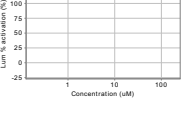
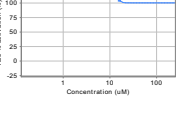
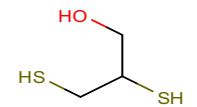
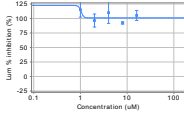
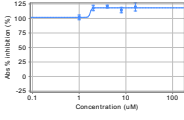
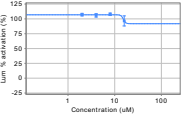
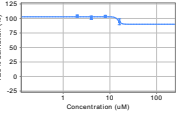
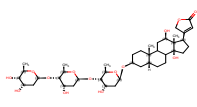
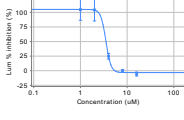
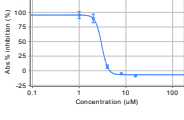
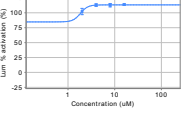
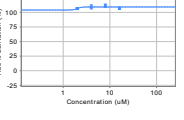
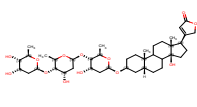
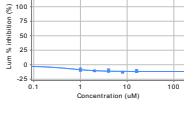
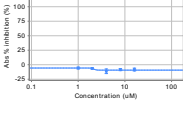
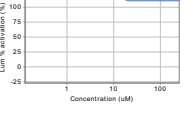
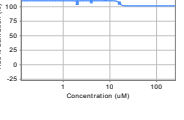
Compound Name	Structure	IRF3 Signaling (Virus-Infected)	NF-κB (Virus-infected)	IRF3 Signaling (Mock-Infected)	NF-κB (Mock-infected)
NUCC-0076927					
NUCC-0076902					
NUCC-0076840					
NUCC-0076793					
NUCC-0076763					
NUCC-0076691					

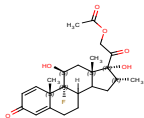
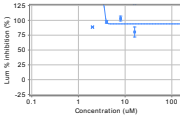
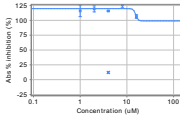
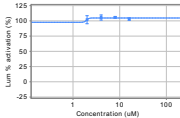
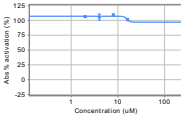
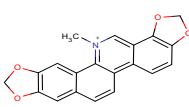
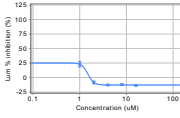
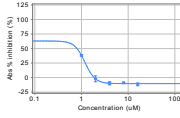
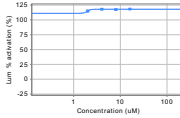
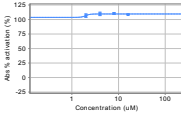
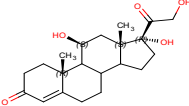
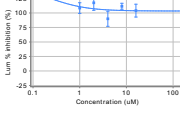
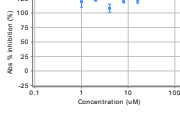
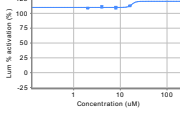
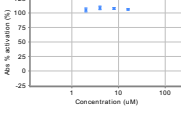
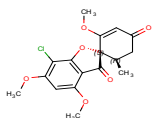
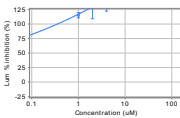
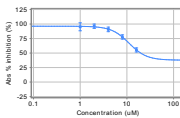
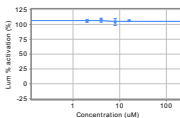
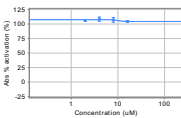
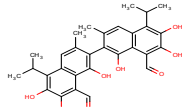
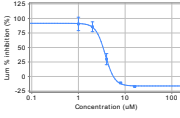
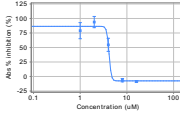
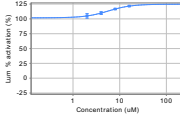
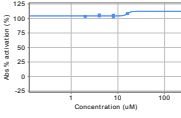
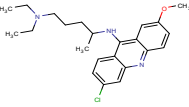
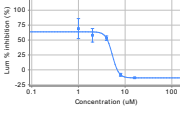
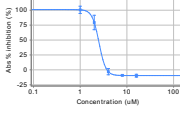
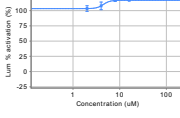
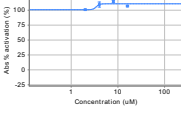
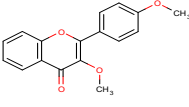
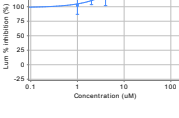
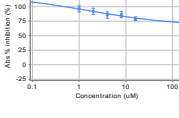
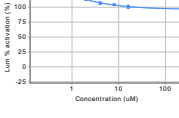
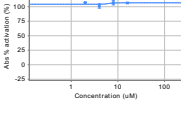
Compound Name	Structure	IRF3 Signaling (Virus-Infected)	NF-κB (Virus-infected)	IRF3 Signaling (Mock-Infected)	NF-κB (Mock-infected)
NUCC-0076689					
NUCC-0076644					
NUCC-0076643					
NUCC-0076625					
NUCC-0076594					
NUCC-0076579					
NUCC-0076558					

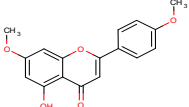
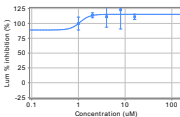
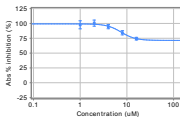
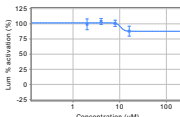
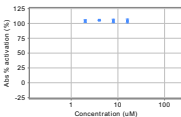
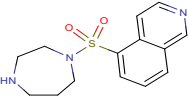
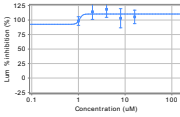
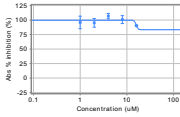
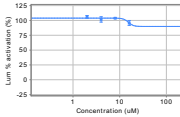
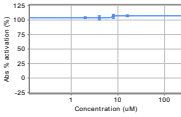
Compound Name	Structure	IRF3 Signaling (Virus-Infected)	NF-κB (Virus-infected)	IRF3 Signaling (Mock-Infected)	NF-κB (Mock-infected)
NUCC-0076525					
NUCC-0076516					
NUCC-0076498					
NUCC-0076485					
NUCC-0076454					
NUCC-0076449					
NUCC-0076436					

Compound Name	Structure	IRF3 Signaling (Virus-Infected)	NF-κB (Virus-infected)	IRF3 Signaling (Mock-Infected)	NF-κB (Mock-infected)
NUCC-0076325					
NUCC-0076269					
NUCC-0076191					
NUCC-0076176					
NUCC-0076137					
NUCC-0076123					
NUCC-0076096					

Compound Name	Structure	IRF3 Signaling (Virus-Infected)	NF-κB (Virus-infected)	IRF3 Signaling (Mock-Infected)	NF-κB (Mock-infected)
NUCC-0075972					
NUCC-0075843					
NUCC-0075798					
NUCC-0075708					
NUCC-0075646					
NUCC-0075634					
NUCC-0075599					

Compound Name	Structure	IRF3 Signaling (Virus-Infected)	NF-κB (Virus-infected)	IRF3 Signaling (Mock-Infected)	NF-κB (Mock-infected)
NUCC-0075594					
NUCC-0075565					
NUCC-0075550					
NUCC-0075505					
NUCC-0075384					
NUCC-0075379					
NUCC-0075378					

Compound Name	Structure	IRF3 Signaling (Virus-Infected)	NF-κB (Virus-infected)	IRF3 Signaling (Mock-Infected)	NF-κB (Mock-infected)
NUCC-0075367					
NUCC-0075261					
NUCC-0075252					
NUCC-0075250					
NUCC-0051653					
NUCC-0051540					
NUCC-0050450					

Compound Name	Structure	IRF3 Signaling (Virus-Infected)	NF-κB (Virus-infected)	IRF3 Signaling (Mock-Infected)	NF-κB (Mock-infected)
NUCC-0050413					
NUCC-0050357					

**APPENDIX E: ENRICHED GO BIOLOGICAL PROCESSES FOR VIRUS-INDUCED
GENES**

Table E.1 – Clusters of enriched GO biological processes terms for virus-induced genes

Annotation Cluster 1 - Enrichment Score: 22.771						
Term	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0034340: response to type I interferon	44	2.73	6.48E-27	6.96	5.17E-23	1.29E-23
GO:0071357: cellular response to type I interferon	42	2.61	7.71E-26	7.00	2.05E-22	1.54E-22
GO:0060337: type I interferon signaling pathway	42	2.61	7.71E-26	7.00	2.05E-22	1.54E-22
GO:0051607: defense response to virus	66	4.10	5.74E-20	3.55	7.63E-17	1.15E-16
GO:0009615: response to virus	77	4.78	6.35E-19	3.05	6.33E-16	1.27E-15
Annotation Cluster 2 - Enrichment Score: 15.765						
Term	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0034097: response to cytokine	158	9.81	1.55E-26	2.43	6.17E-23	3.09E-23
GO:0071345: cellular response to cytokine stimulus	138	8.57	7.57E-23	2.41	1.51E-19	1.51E-19
GO:0019221: cytokine-mediated signaling pathway	117	7.27	3.11E-22	2.60	4.96E-19	6.21E-19
GO:0070887: cellular response to chemical stimulus	324	20.12	8.00E-16	1.51	4.77E-13	1.55E-12
GO:0010033: response to organic substance	339	21.06	1.95E-15	1.48	1.06E-12	3.99E-12
GO:0071310: cellular response to organic substance	277	17.20	2.90E-15	1.56	1.44E-12	5.76E-12
GO:0007166: cell surface receptor signaling pathway	297	18.45	1.74E-09	1.37	2.21E-07	3.48E-06
GO:1901700: response to oxygen-containing compound	156	9.69	2.65E-04	1.31	6.68E-03	5.29E-01
Annotation Cluster 3 - Enrichment Score: 13.135						
Term	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0034097: response to cytokine	158	9.81	1.55E-26	2.43	6.17E-23	3.09E-23

GO:0071345: cellular response to cytokine stimulus	138	8.57	7.57E-23	2.41	1.51E-19	1.51E-19
GO:0019221: cytokine-mediated signaling pathway	117	7.27	3.11E-22	2.60	4.96E-19	6.21E-19
GO:0051607: defense response to virus	66	4.10	5.74E-20	3.55	7.63E-17	1.15E-16
GO:0006955: immune response	227	14.10	7.91E-20	1.81	9.02E-17	1.58E-16
GO:0009615: response to virus	77	4.78	6.35E-19	3.05	6.33E-16	1.27E-15
GO:0045087: innate immune response	141	8.76	2.11E-17	2.08	1.87E-14	4.22E-14
GO:0043207: response to external biotic stimulus	139	8.63	6.05E-17	2.07	8.86E-14	2.22E-13
GO:0051707: response to other organism	139	8.63	6.05E-17	2.07	8.86E-14	2.22E-13
GO:0009607: response to biotic stimulus	144	8.94	6.48E-17	2.03	8.05E-14	2.22E-13
GO:0002682: regulation of immune system process	197	12.24	7.45E-16	1.76	5.17E-13	1.55E-12
GO:0006952: defense response	211	13.11	1.31E-15	1.71	7.59E-13	2.66E-12
GO:0002252: immune effector process	123	7.64	1.80E-14	2.04	8.44E-12	3.59E-11
GO:0080134: regulation of response to stress	181	11.24	6.30E-14	1.74	2.65E-11	1.26E-10
GO:0001816: cytokine production	108	6.71	1.07E-13	2.11	4.07E-11	2.14E-10
GO:0001817: regulation of cytokine production	101	6.27	1.80E-13	2.16	5.74E-11	3.59E-10
GO:0050776: regulation of immune response	139	8.63	2.50E-13	1.88	7.38E-11	4.98E-10
GO:0031347: regulation of defense response	111	6.89	7.67E-13	2.03	2.04E-10	1.53E-09
GO:0098542: defense response to other organism	89	5.53	1.15E-12	2.22	2.77E-10	2.29E-09
GO:0009605: response to external stimulus	245	15.22	3.02E-10	1.46	4.73E-08	6.03E-07
GO:0001819: positive regulation of cytokine production	69	4.29	7.30E-10	2.20	1.04E-07	1.46E-06
GO:0043903: regulation of symbiosis, encompassing mutualism through parasitism	56	3.48	1.08E-09	2.42	1.46E-07	2.16E-06

GO:0050792: regulation of viral process	51	3.17	9.10E-09	2.40	9.43E-07	1.82E-05
GO:0044419: interspecies interaction between organisms	132	8.20	1.56E-08	1.63	1.57E-06	3.11E-05
GO:0044403: symbiosis, encompassing mutualism through parasitism	132	8.20	1.56E-08	1.63	1.57E-06	3.11E-05
GO:0044764: multi-organism cellular process	127	7.89	6.45E-08	1.61	5.20E-06	1.29E-04
GO:0016032: viral process	126	7.83	7.58E-08	1.61	5.87E-06	1.51E-04
GO:0002684: positive regulation of immune system process	123	7.64	2.95E-07	1.58	2.03E-05	5.90E-04
GO:0043900: regulation of multi-organism process	61	3.79	3.11E-07	2.00	2.12E-05	6.21E-04
GO:0050778: positive regulation of immune response	88	5.47	1.52E-05	1.59	6.32E-04	3.04E-02
Annotation Cluster 4 - Enrichment Score: 11.183						
Term	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0034341: response to interferon-gamma	44	2.73	1.34E-12	3.37	3.15E-10	2.68E-09
GO:0060333: interferon-gamma-mediated signaling pathway	29	1.80	9.81E-12	4.47	2.01E-09	1.96E-08
GO:0071346: cellular response to interferon-gamma	38	2.36	2.15E-11	3.47	3.89E-09	4.29E-08
Annotation Cluster 5 - Enrichment Score: 10.705						
Term	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0006986: response to unfolded protein	47	2.92	2.10E-13	3.37	6.44E-11	4.19E-10
GO:0035966: response to topologically incorrect protein	49	3.04	4.17E-13	3.21	1.19E-10	8.32E-10
GO:0034976: response to endoplasmic reticulum stress	60	3.73	1.10E-12	2.75	2.74E-10	2.19E-09
GO:0030968: endoplasmic reticulum unfolded protein response	38	2.36	1.04E-11	3.55	2.01E-09	2.07E-08

GO:0034620: cellular response to unfolded protein	38	2.36	2.15E-11	3.47	3.89E-09	4.29E-08
GO:0035967: cellular response to topologically incorrect protein	40	2.48	2.90E-11	3.31	4.92E-09	5.79E-08
GO:0036498: IRE1-mediated unfolded protein response	20	1.24	1.88E-06	3.52	1.06E-04	3.75E-03
Annotation Cluster 6 - Enrichment Score: 8.783						
Term	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0007249: I-kappaB kinase/NF-kappaB signaling	58	3.60	1.64E-12	2.78	3.73E-10	3.27E-09
GO:0043122: regulation of I-kappaB kinase/NF-kappaB signaling	51	3.17	2.95E-10	2.64	4.71E-08	5.90E-07
GO:0043123: positive regulation of I-kappaB kinase/NF-kappaB signaling	34	2.11	9.25E-06	2.30	4.22E-04	1.85E-02
Annotation Cluster 7 - Enrichment Score: 7.198						
Term	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0048525: negative regulation of viral process	34	2.11	1.03E-11	3.88	2.05E-09	2.05E-08
GO:1903901: negative regulation of viral life cycle	29	1.80	1.76E-10	4.02	2.86E-08	3.51E-07
GO:0043901: negative regulation of multi-organism process	39	2.42	6.14E-10	3.06	9.24E-08	1.22E-06
GO:0043903: regulation of symbiosis, encompassing mutualism through parasitism	56	3.48	1.08E-09	2.42	1.46E-07	2.16E-06
GO:0050792: regulation of viral process	51	3.17	9.10E-09	2.40	9.43E-07	1.82E-05
GO:0044419: interspecies interaction between organisms	132	8.20	1.56E-08	1.63	1.57E-06	3.11E-05
GO:0044403: symbiosis, encompassing mutualism through parasitism	132	8.20	1.56E-08	1.63	1.57E-06	3.11E-05

GO:0045071: negative regulation of viral genome replication	19	1.18	4.25E-08	4.59	3.65E-06	8.49E-05
GO:0044764: multi-organism cellular process	127	7.89	6.45E-08	1.61	5.20E-06	1.29E-04
GO:0016032: viral process	126	7.83	7.58E-08	1.61	5.87E-06	1.51E-04
GO:0043900: regulation of multi-organism process	61	3.79	3.11E-07	2.00	2.12E-05	6.21E-04
GO:0045069: regulation of viral genome replication	22	1.37	1.25E-06	3.35	7.53E-05	2.49E-03
GO:1903900: regulation of viral life cycle	34	2.11	5.00E-06	2.37	2.46E-04	9.97E-03
GO:0019079: viral genome replication	24	1.49	1.05E-05	2.79	4.71E-04	2.10E-02
GO:0019058: viral life cycle	46	2.86	0.094022738	1.25	5.94E-01	8.61E+01
Annotation Cluster 8 - Enrichment Score: 6.760						
Term	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0006955: immune response	227	14.10	7.91E-20	1.81	9.02E-17	1.58E-16
GO:0002682: regulation of immune system process	197	12.24	7.45E-16	1.76	5.17E-13	1.55E-12
GO:0001816: cytokine production	108	6.71	1.07E-13	2.11	4.07E-11	2.14E-10
GO:0001817: regulation of cytokine production	101	6.27	1.80E-13	2.16	5.74E-11	3.59E-10
GO:0050776: regulation of immune response	139	8.63	2.50E-13	1.88	7.38E-11	4.98E-10
GO:0045088: regulation of innate immune response	73	4.53	4.61E-13	2.50	1.27E-10	9.20E-10
GO:0031347: regulation of defense response	111	6.89	7.67E-13	2.03	2.04E-10	1.53E-09
GO:0001819: positive regulation of cytokine production	69	4.29	7.30E-10	2.20	1.04E-07	1.46E-06
GO:0002684: positive regulation of immune system process	123	7.64	2.95E-07	1.58	2.03E-05	5.90E-04
GO:0031349: positive regulation of defense response	61	3.79	2.73E-06	1.87	1.43E-04	5.45E-03
GO:0051240: positive regulation of multicellular organismal process	161	10.00	1.17E-05	1.39	5.14E-04	2.34E-02

GO:0050778: positive regulation of immune response	88	5.47	1.52E-05	1.59	6.32E-04	3.04E-02
GO:0002221: pattern recognition receptor signaling pathway	30	1.86	2.81E-05	2.33	1.08E-03	5.62E-02
GO:0002764: immune response-regulating signaling pathway	70	4.35	5.93E-05	1.63	2.01E-03	1.18E-01
GO:0045089: positive regulation of innate immune response	44	2.73	9.43E-05	1.86	2.90E-03	1.88E-01
GO:0002758: innate immune response-activating signal transduction	38	2.36	1.42E-04	1.93	4.14E-03	2.83E-01
GO:0002218: activation of innate immune response	38	2.36	2.54E-04	1.87	6.46E-03	5.06E-01
GO:0002757: immune response-activating signal transduction	64	3.98	2.61E-04	1.59	6.58E-03	5.19E-01
GO:0002224: toll-like receptor signaling pathway	22	1.37	5.42E-04	2.28	1.22E-02	1.08E+00
GO:0002253: activation of immune response	65	4.04	0.002032953	1.46	3.60E-02	3.98E+00
GO:0050852: T cell receptor signaling pathway	21	1.30	0.050290079	1.55	4.19E-01	6.43E+01
GO:0002768: immune response-regulating cell surface receptor signaling pathway	42	2.61	0.068146219	1.30	5.00E-01	7.56E+01
GO:0050851: antigen receptor-mediated signaling pathway	26	1.61	0.074805346	1.41	5.28E-01	7.88E+01
GO:0002429: immune response-activating cell surface receptor signaling pathway	37	2.30	0.132019799	1.24	6.97E-01	9.41E+01
Annotation Cluster 9 - Enrichment Score: 6.309						
Term:	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0060338: regulation of type I interferon-mediated signaling pathway	21	1.30	2.00E-12	6.64	4.44E-10	4.00E-09
GO:0043331: response to dsRNA	29	1.80	2.70E-11	4.31	4.69E-09	5.40E-08
GO:0043330: response to exogenous dsRNA	21	1.30	5.73E-11	5.76	9.53E-09	1.15E-07

GO:0033141: positive regulation of peptidyl-serine phosphorylation of STAT protein	14	0.87	6.73E-10	8.22	9.77E-08	1.34E-06
GO:0033139: regulation of peptidyl-serine phosphorylation of STAT protein	14	0.87	1.52E-09	7.85	1.99E-07	3.04E-06
GO:0060759: regulation of response to cytokine stimulus	37	2.30	1.82E-09	3.06	2.27E-07	3.63E-06
GO:0001959: regulation of cytokine-mediated signaling pathway	36	2.24	2.11E-09	3.10	2.47E-07	4.21E-06
GO:0002286: T cell activation involved in immune response	27	1.68	4.62E-09	3.74	5.27E-07	9.23E-06
GO:0002323: natural killer cell activation involved in immune response	15	0.93	2.07E-08	6.17	2.02E-06	4.14E-05
GO:0046651: lymphocyte proliferation	50	3.11	2.85E-08	2.34	2.62E-06	5.70E-05
GO:0032943: mononuclear cell proliferation	50	3.11	3.66E-08	2.33	3.21E-06	7.31E-05
GO:0042501: serine phosphorylation of STAT protein	14	0.87	4.18E-08	6.40	3.62E-06	8.34E-05
GO:0070661: leukocyte proliferation	51	3.17	8.43E-08	2.25	6.47E-06	1.68E-04
GO:0033138: positive regulation of peptidyl-serine phosphorylation	25	1.55	1.24E-07	3.43	9.34E-06	2.48E-04
GO:0030101: natural killer cell activation	22	1.37	5.00E-07	3.52	3.30E-05	9.98E-04
GO:0007259: JAK-STAT cascade	35	2.17	7.91E-07	2.52	4.97E-05	1.58E-03
GO:0097696: STAT cascade	35	2.17	1.58E-06	2.45	9.25E-05	3.15E-03
GO:0002285: lymphocyte activation involved in immune response	31	1.93	2.64E-06	2.57	1.42E-04	5.28E-03
GO:0042113: B cell activation	42	2.61	7.24E-06	2.11	3.46E-04	1.44E-02
GO:0033135: regulation of peptidyl-serine phosphorylation	26	1.61	9.43E-06	2.67	4.27E-04	1.88E-02
GO:0042100: B cell proliferation	21	1.30	1.32E-05	3.01	5.61E-04	2.64E-02

GO:0002263: cell activation involved in immune response	37	2.30	2.08E-05	2.13	8.34E-04	4.16E-02
GO:0030183: B cell differentiation	24	1.49	2.70E-05	2.64	1.05E-03	5.39E-02
GO:0018105: peptidyl-serine phosphorylation	40	2.48	3.86E-05	2.01	1.40E-03	7.70E-02
GO:0002366: leukocyte activation involved in immune response	36	2.24	4.04E-05	2.09	1.46E-03	8.06E-02
GO:0018209: peptidyl-serine modification	41	2.55	5.65E-05	1.95	1.94E-03	1.13E-01
GO:0006959: humoral immune response	33	2.05	0.005890338	1.64	8.71E-02	1.11E+01
GO:1904892: regulation of STAT cascade	22	1.37	0.008897507	1.82	1.20E-01	1.63E+01
GO:0046425: regulation of JAK-STAT cascade	22	1.37	0.008897507	1.82	1.20E-01	1.63E+01
GO:0018193: peptidyl-amino acid modification	115	7.14	0.029439599	1.20	2.93E-01	4.49E+01
Annotation Cluster 10 - Enrichment Score: 6.297						
Term	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0008219: cell death	233	14.47	5.13E-10	1.47	7.87E-08	1.02E-06
GO:0006915: apoptotic process	210	13.04	1.97E-09	1.48	2.39E-07	3.94E-06
GO:0012501: programmed cell death	220	13.66	2.10E-09	1.47	2.50E-07	4.19E-06
GO:0010941: regulation of cell death	182	11.30	2.56E-08	1.49	2.40E-06	5.10E-05
GO:0042981: regulation of apoptotic process	171	10.62	2.89E-08	1.51	2.62E-06	5.78E-05
GO:0043067: regulation of programmed cell death	172	10.68	3.34E-08	1.50	3.00E-06	6.68E-05
GO:0043065: positive regulation of apoptotic process	80	4.97	1.77E-06	1.72	1.02E-04	3.54E-03
GO:0043068: positive regulation of programmed cell death	80	4.97	2.51E-06	1.71	1.37E-04	5.01E-03
GO:0010942: positive regulation of cell death	82	5.09	5.10E-06	1.67	2.50E-04	1.02E-02
GO:0060548: negative regulation of cell death	102	6.34	4.02E-04	1.40	9.55E-03	7.99E-01
GO:0043066: negative regulation of apoptotic process	93	5.78	7.48E-04	1.40	1.55E-02	1.48E+00
GO:0043069: negative regulation of programmed cell death	94	5.84	7.56E-04	1.40	1.56E-02	1.50E+00

Annotation Cluster 11 - Enrichment Score: 5.871						
Term:	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0009966: regulation of signal transduction	299	18.57	7.33E-09	1.35	7.80E-07	1.46E-05
GO:0048585: negative regulation of response to stimulus	167	10.37	4.82E-08	1.50	4.05E-06	9.62E-05
GO:0010646: regulation of cell communication	320	19.88	6.61E-08	1.31	5.22E-06	1.32E-04
GO:0023051: regulation of signaling	322	20.00	1.59E-07	1.29	1.16E-05	3.17E-04
GO:1902532: negative regulation of intracellular signal transduction	67	4.16	2.23E-06	1.82	1.24E-04	4.45E-03
GO:0009968: negative regulation of signal transduction	124	7.70	8.35E-05	1.40	2.65E-03	1.67E-01
GO:0010648: negative regulation of cell communication	132	8.20	1.15E-04	1.38	3.50E-03	2.30E-01
GO:0023057: negative regulation of signaling	132	8.20	1.36E-04	1.37	4.05E-03	2.71E-01
Annotation Cluster 12 - Enrichment Score: 5.549						
Term:	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0070486: leukocyte aggregation	87	5.40	6.04E-14	2.36	2.68E-11	1.21E-10
GO:0071593: lymphocyte aggregation	86	5.34	6.84E-14	2.37	2.73E-11	1.37E-10
GO:0007159: leukocyte cell-cell adhesion	91	5.65	1.13E-13	2.29	3.92E-11	2.26E-10
GO:0042110: T cell activation	85	5.28	1.66E-13	2.35	5.53E-11	3.32E-10
GO:0070489: T cell aggregation	85	5.28	1.66E-13	2.35	5.53E-11	3.32E-10
GO:0001775: cell activation	137	8.51	9.61E-13	1.86	2.47E-10	1.92E-09
GO:0016337: single organismal cell-cell adhesion	114	7.08	6.47E-12	1.94	1.40E-09	1.29E-08
GO:0045321: leukocyte activation	116	7.20	9.35E-12	1.92	1.96E-09	1.87E-08
GO:0098602: single organism cell adhesion	119	7.39	1.44E-11	1.88	2.67E-09	2.88E-08
GO:0046649: lymphocyte activation	103	6.40	2.56E-11	1.98	4.54E-09	5.11E-08

GO:0048534: hematopoietic or lymphoid organ development	113	7.02	1.09E-09	1.79	1.45E-07	2.17E-06
GO:0002520: immune system development	117	7.27	1.65E-09	1.76	2.13E-07	3.30E-06
GO:0002250: adaptive immune response	72	4.47	1.95E-09	2.11	2.39E-07	3.89E-06
GO:0030097: hemopoiesis	106	6.58	7.73E-09	1.77	8.11E-07	1.54E-05
GO:0030098: lymphocyte differentiation	56	3.48	2.41E-08	2.23	2.29E-06	4.81E-05
GO:0046651: lymphocyte proliferation	50	3.11	2.85E-08	2.34	2.62E-06	5.70E-05
GO:0032943: mononuclear cell proliferation	50	3.11	3.66E-08	2.33	3.21E-06	7.31E-05
GO:0030155: regulation of cell adhesion	93	5.78	4.79E-08	1.78	4.07E-06	9.57E-05
GO:0002521: leukocyte differentiation	74	4.60	6.50E-08	1.93	5.18E-06	1.30E-04
GO:0070661: leukocyte proliferation	51	3.17	8.43E-08	2.25	6.47E-06	1.68E-04
GO:0022407: regulation of cell-cell adhesion	63	3.91	1.42E-07	2.01	1.05E-05	2.83E-04
GO:0098609: cell-cell adhesion	144	8.94	2.09E-07	1.53	1.49E-05	4.18E-04
GO:0050865: regulation of cell activation	74	4.60	5.95E-07	1.82	3.83E-05	1.19E-03
GO:0007155: cell adhesion	192	11.93	9.99E-07	1.39	6.18E-05	2.00E-03
GO:0022610: biological adhesion	192	11.93	1.31E-06	1.39	7.87E-05	2.62E-03
GO:0002694: regulation of leukocyte activation	69	4.29	1.60E-06	1.82	9.31E-05	3.19E-03
GO:1903037: regulation of leukocyte cell-cell adhesion	51	3.17	2.63E-06	2.01	1.43E-04	5.25E-03
GO:0030217: T cell differentiation	38	2.36	2.70E-06	2.30	1.44E-04	5.39E-03
GO:0051249: regulation of lymphocyte activation	62	3.85	2.71E-06	1.86	1.43E-04	5.41E-03
GO:0050863: regulation of T cell activation	49	3.04	3.64E-06	2.02	1.88E-04	7.28E-03
GO:0045785: positive regulation of cell adhesion	57	3.54	1.26E-05	1.83	5.39E-04	2.52E-02
GO:0022409: positive regulation of cell-cell adhesion	40	2.48	4.64E-05	1.99	1.64E-03	9.26E-02

GO:0045619: regulation of lymphocyte differentiation	26	1.61	1.39E-04	2.29	4.11E-03	2.77E-01
GO:0042098: T cell proliferation	30	1.86	1.51E-04	2.13	4.35E-03	3.01E-01
GO:0046634: regulation of alpha-beta T cell activation	17	1.06	1.68E-04	2.91	4.65E-03	3.34E-01
GO:1902105: regulation of leukocyte differentiation	37	2.30	3.25E-04	1.87	7.92E-03	6.47E-01
GO:1903039: positive regulation of leukocyte cell-cell adhesion	34	2.11	3.54E-04	1.92	8.44E-03	7.04E-01
GO:0042129: regulation of T cell proliferation	26	1.61	4.63E-04	2.12	1.08E-02	9.21E-01
GO:1903706: regulation of hemopoiesis	45	2.80	6.33E-04	1.69	1.37E-02	1.26E+00
GO:0045580: regulation of T cell differentiation	21	1.30	7.00E-04	2.29	1.49E-02	1.39E+00
GO:0050867: positive regulation of cell activation	44	2.73	7.01E-04	1.70	1.48E-02	1.39E+00
GO:0050870: positive regulation of T cell activation	32	1.99	8.10E-04	1.88	1.65E-02	1.60E+00
GO:0050670: regulation of lymphocyte proliferation	30	1.86	0.001328156	1.87	2.52E-02	2.62E+00
GO:0002696: positive regulation of leukocyte activation	42	2.61	0.001336536	1.67	2.53E-02	2.64E+00
GO:0070663: regulation of leukocyte proliferation	31	1.93	0.001419108	1.84	2.67E-02	2.80E+00
GO:0032944: regulation of mononuclear cell proliferation	30	1.86	0.001437405	1.86	2.70E-02	2.83E+00
GO:1902107: positive regulation of leukocyte differentiation	21	1.30	0.004337362	1.98	6.78E-02	8.31E+00
GO:0050671: positive regulation of lymphocyte proliferation	21	1.30	0.004733075	1.96	7.26E-02	9.04E+00
Annotation Cluster 12 - Enrichment Score: 5.549 (continued)						
Term:	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0032946: positive regulation of mononuclear cell proliferation	21	1.30	0.005158131	1.95	7.83E-02	9.81E+00

GO:0042102: positive regulation of T cell proliferation	17	1.06	0.005271224	2.14	7.92E-02	1.00E+01
GO:0051251: positive regulation of lymphocyte activation	37	2.30	0.00574873	1.59	8.56E-02	1.09E+01
GO:1903708: positive regulation of hemopoiesis	24	1.49	0.006303844	1.82	9.11E-02	1.19E+01
GO:0046637: regulation of alpha-beta T cell differentiation	11	0.68	0.007020659	2.66	9.97E-02	1.31E+01
GO:0070665: positive regulation of leukocyte proliferation	21	1.30	0.008417206	1.86	1.15E-01	1.55E+01
GO:0045582: positive regulation of T cell differentiation	13	0.81	0.008914243	2.32	1.20E-01	1.64E+01
GO:0046635: positive regulation of alpha-beta T cell activation	11	0.68	0.009250623	2.56	1.24E-01	1.69E+01
GO:0045621: positive regulation of lymphocyte differentiation	14	0.87	0.012926971	2.13	1.61E-01	2.29E+01
GO:0030099: myeloid cell differentiation	40	2.48	0.019589544	1.44	2.18E-01	3.26E+01
GO:0046638: positive regulation of alpha-beta T cell differentiation	8	0.50	0.035136654	2.53	3.29E-01	5.10E+01
GO:0031295: T cell costimulation	12	0.75	0.06162725	1.83	4.73E-01	7.19E+01
GO:0031294: lymphocyte costimulation	12	0.75	0.066213196	1.80	4.92E-01	7.45E+01
GO:0045637: regulation of myeloid cell differentiation	20	1.24	0.206056222	1.29	8.14E-01	9.90E+01
Annotation Cluster 13 - Enrichment Score: 5.518						
Term:	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0032496: response to lipopolysaccharide	56	3.48	9.40E-09	2.29	9.61E-07	1.88E-05
GO:0002237: response to molecule of bacterial origin	57	3.54	1.88E-08	2.22	1.85E-06	3.74E-05
GO:0071216: cellular response to biotic stimulus	36	2.24	4.73E-07	2.54	3.17E-05	9.44E-04

GO:0071222: cellular response to lipopolysaccharide	32	1.99	8.65E-07	2.65	5.39E-05	1.73E-03
GO:0071219: cellular response to molecule of bacterial origin	32	1.99	2.10E-06	2.55	1.17E-04	4.19E-03
GO:0009617: response to bacterium	74	4.60	1.23E-05	1.68	5.35E-04	2.47E-02
GO:0033993: response to lipid	97	6.02	7.59E-04	1.39	1.56E-02	1.51E+00
GO:0071396: cellular response to lipid	59	3.66	0.00506604	1.43	7.71E-02	9.64E+00
Annotation Cluster 14 - Enrichment Score: 5.463						
Term:	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0051252: regulation of RNA metabolic process	397	24.66	9.96E-10	1.30	1.37E-07	1.99E-06
GO:2001141: regulation of RNA biosynthetic process	384	23.85	2.90E-09	1.30	3.35E-07	5.79E-06
GO:0006355: regulation of transcription, DNA-templated	379	23.54	4.74E-09	1.30	5.33E-07	9.46E-06
GO:1903506: regulation of nucleic acid-templated transcription	380	23.60	6.79E-09	1.29	7.32E-07	1.35E-05
GO:0097659: nucleic acid-templated transcription	389	24.16	2.83E-08	1.27	2.63E-06	5.65E-05
GO:0010604: positive regulation of macromolecule metabolic process	307	19.07	3.37E-08	1.33	2.99E-06	6.74E-05
GO:0010556: regulation of macromolecule biosynthetic process	417	25.90	4.97E-08	1.25	4.13E-06	9.93E-05
GO:0009893: positive regulation of metabolic process	323	20.06	6.15E-08	1.31	5.01E-06	1.23E-04
GO:0006351: transcription, DNA-templated	372	23.11	6.71E-08	1.27	5.25E-06	1.34E-04
GO:0032774: RNA biosynthetic process	396	24.60	1.60E-07	1.25	1.16E-05	3.19E-04
GO:0019219: regulation of nucleobase-containing compound metabolic process	413	25.65	1.71E-07	1.24	1.23E-05	3.41E-04
GO:0051171: regulation of nitrogen compound metabolic process	438	27.20	2.33E-07	1.23	1.64E-05	4.64E-04

GO:0010468: regulation of gene expression	433	26.89	2.38E-07	1.23	1.67E-05	4.76E-04
GO:2000112: regulation of cellular macromolecule biosynthetic process	402	24.97	2.50E-07	1.24	1.73E-05	4.99E-04
GO:0031325: positive regulation of cellular metabolic process	296	18.39	1.13E-06	1.29	6.87E-05	2.25E-03
GO:0016070: RNA metabolic process	452	28.07	2.01E-06	1.20	1.13E-04	4.01E-03
GO:0034645: cellular macromolecule biosynthetic process	475	29.50	1.91E-05	1.17	7.76E-04	3.80E-02
GO:0010628: positive regulation of gene expression	184	11.43	2.00E-05	1.34	8.07E-04	4.00E-02
GO:0051254: positive regulation of RNA metabolic process	157	9.75	3.21E-05	1.37	1.21E-03	6.41E-02
GO:0034654: nucleobase-containing compound biosynthetic process	416	25.84	4.28E-05	1.18	1.53E-03	8.54E-02
GO:0018130: heterocycle biosynthetic process	420	26.09	5.32E-05	1.17	1.85E-03	1.06E-01
GO:1903508: positive regulation of nucleic acid-templated transcription	149	9.25	6.42E-05	1.36	2.16E-03	1.28E-01
GO:0045893: positive regulation of transcription, DNA-templated	149	9.25	6.42E-05	1.36	2.16E-03	1.28E-01
GO:0019438: aromatic compound biosynthetic process	420	26.09	7.03E-05	1.17	2.35E-03	1.40E-01
GO:0051173: positive regulation of nitrogen compound metabolic process	187	11.61	8.07E-05	1.31	2.60E-03	1.61E-01
GO:0006357: regulation of transcription from RNA polymerase II promoter	193	11.99	8.58E-05	1.30	2.69E-03	1.71E-01
GO:1902680: positive regulation of RNA biosynthetic process	150	9.32	9.42E-05	1.35	2.92E-03	1.88E-01

GO:0010557: positive regulation of macromolecule biosynthetic process	169	10.50	1.92E-04	1.31	5.15E-03	3.83E-01
GO:0045935: positive regulation of nucleobase-containing compound metabolic process	173	10.75	3.11E-04	1.29	7.67E-03	6.19E-01
GO:0045944: positive regulation of transcription from RNA polymerase II promoter	115	7.14	5.03E-04	1.36	1.15E-02	9.99E-01
GO:0031328: positive regulation of cellular biosynthetic process	177	10.99	6.54E-04	1.26	1.41E-02	1.30E+00
GO:0009891: positive regulation of biosynthetic process	180	11.18	6.59E-04	1.26	1.41E-02	1.31E+00
GO:0010467: gene expression	484	30.06	7.40E-04	1.13	1.55E-02	1.47E+00
GO:0006366: transcription from RNA polymerase II promoter	181	11.24	0.002219147	1.23	3.87E-02	4.34E+00
Annotation Cluster 15 - Enrichment Score: 5.384						
Term:	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0051090: regulation of sequence-specific DNA binding transcription factor activity	65	4.04	6.78E-09	2.15	7.41E-07	1.35E-05
GO:0051091: positive regulation of sequence-specific DNA binding transcription factor activity	39	2.42	2.55E-05	2.06	9.91E-04	5.09E-02
GO:0051092: positive regulation of NF-kappaB transcription factor activity	24	1.49	4.08E-04	2.23	9.65E-03	8.12E-01
Annotation Cluster 16 - Enrichment Score: 4.944						
Term:	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0043903: regulation of symbiosis, encompassing mutualism through parasitism	56	3.48	1.08E-09	2.42	1.46E-07	2.16E-06
GO:0050792: regulation of viral process	51	3.17	9.10E-09	2.40	9.43E-07	1.82E-05

GO:0050688: regulation of defense response to virus	20	1.24	5.48E-05	2.84	1.89E-03	1.09E-01
GO:0002831: regulation of response to biotic stimulus	25	1.55	2.26E-04	2.27	5.85E-03	4.50E-01
GO:0050691: regulation of defense response to virus by host	9	0.56	0.002652123	3.58	4.51E-02	5.16E+00
GO:0002230: positive regulation of defense response to virus by host	7	0.43	0.00673177	3.92	9.60E-02	1.26E+01
Annotation Cluster 17 - Enrichment Score: 4.870						
Term:	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0042107: cytokine metabolic process	24	1.49	8.90E-06	2.82	4.13E-04	1.78E-02
GO:0042035: regulation of cytokine biosynthetic process	23	1.43	9.14E-06	2.89	4.21E-04	1.82E-02
GO:0042108: positive regulation of cytokine biosynthetic process	17	1.06	1.93E-05	3.44	7.82E-04	3.85E-02
GO:0042089: cytokine biosynthetic process	23	1.43	2.11E-05	2.75	8.37E-04	4.21E-02
Annotation Cluster 18 - Enrichment Score: 4.353						
Term:	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0065009: regulation of molecular function	304	18.88	1.33E-07	1.31	9.89E-06	2.65E-04
GO:0044093: positive regulation of molecular function	200	12.42	7.99E-06	1.34	3.77E-04	1.59E-02
GO:0050790: regulation of catalytic activity	245	15.22	1.57E-05	1.28	6.50E-04	3.14E-02
GO:0043547: positive regulation of GTPase activity	79	4.91	1.27E-04	1.54	3.79E-03	2.52E-01
GO:0043085: positive regulation of catalytic activity	166	10.31	1.55E-04	1.31	4.45E-03	3.10E-01
GO:0051345: positive regulation of hydrolase activity	106	6.58	1.67E-04	1.43	4.66E-03	3.33E-01
GO:0043087: regulation of GTPase activity	84	5.22	1.74E-04	1.50	4.80E-03	3.48E-01
GO:0051336: regulation of hydrolase activity	139	8.63	0.001553755	1.28	2.88E-02	3.06E+00

Annotation Cluster 19 - Enrichment Score: 4.247						
Term:	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0035556: intracellular signal transduction	316	19.63	1.07E-13	1.47	3.89E-11	2.14E-10
GO:1902531: regulation of intracellular signal transduction	216	13.42	1.38E-11	1.55	2.62E-09	2.75E-08
GO:0009966: regulation of signal transduction	299	18.57	7.33E-09	1.35	7.80E-07	1.46E-05
GO:0031399: regulation of protein modification process	195	12.11	6.00E-08	1.45	4.93E-06	1.20E-04
GO:0010646: regulation of cell communication	320	19.88	6.61E-08	1.31	5.22E-06	1.32E-04
GO:0001932: regulation of protein phosphorylation	159	9.88	9.16E-08	1.51	6.96E-06	1.83E-04
GO:0023051: regulation of signaling	322	20.00	1.59E-07	1.29	1.16E-05	3.17E-04
GO:0042325: regulation of phosphorylation	164	10.19	6.41E-07	1.45	4.09E-05	1.28E-03
GO:0051246: regulation of protein metabolic process	266	16.52	1.39E-06	1.31	8.31E-05	2.79E-03
GO:0071900: regulation of protein serine/threonine kinase activity	70	4.35	1.46E-06	1.81	8.64E-05	2.92E-03
GO:0006468: protein phosphorylation	204	12.67	2.65E-06	1.36	1.42E-04	5.29E-03
GO:0048584: positive regulation of response to stimulus	220	13.66	3.01E-06	1.34	1.57E-04	6.01E-03
GO:1902533: positive regulation of intracellular signal transduction	114	7.08	3.34E-06	1.54	1.73E-04	6.67E-03
GO:0019220: regulation of phosphate metabolic process	180	11.18	4.60E-06	1.38	2.31E-04	9.19E-03
GO:0051174: regulation of phosphorus metabolic process	180	11.18	4.96E-06	1.38	2.46E-04	9.91E-03
GO:0001934: positive regulation of protein phosphorylation	109	6.77	1.18E-05	1.51	5.13E-04	2.35E-02
GO:0023056: positive regulation of signaling	173	10.75	1.37E-05	1.37	5.77E-04	2.74E-02

GO:0010647: positive regulation of cell communication	172	10.68	1.46E-05	1.36	6.08E-04	2.91E-02
GO:0050790: regulation of catalytic activity	245	15.22	1.57E-05	1.28	6.50E-04	3.14E-02
GO:0009967: positive regulation of signal transduction	159	9.88	2.34E-05	1.37	9.21E-04	4.68E-02
GO:0051247: positive regulation of protein metabolic process	166	10.31	2.89E-05	1.36	1.11E-03	5.76E-02
GO:0032268: regulation of cellular protein metabolic process	242	15.03	3.10E-05	1.27	1.18E-03	6.20E-02
GO:0042327: positive regulation of phosphorylation	111	6.89	3.35E-05	1.47	1.26E-03	6.69E-02
GO:0045859: regulation of protein kinase activity	91	5.65	3.51E-05	1.54	1.31E-03	7.01E-02
GO:0080135: regulation of cellular response to stress	76	4.72	3.52E-05	1.61	1.31E-03	7.03E-02
GO:0036211: protein modification process	369	22.92	5.69E-05	1.19	1.95E-03	1.14E-01
GO:0006464: cellular protein modification process	369	22.92	5.69E-05	1.19	1.95E-03	1.14E-01
GO:0043405: regulation of MAP kinase activity	48	2.98	8.28E-05	1.81	2.64E-03	1.65E-01
GO:0043549: regulation of kinase activity	95	5.90	8.47E-05	1.49	2.67E-03	1.69E-01
GO:0006793: phosphorus metabolic process	304	18.88	9.03E-05	1.21	2.82E-03	1.80E-01
GO:0006796: phosphate-containing compound metabolic process	302	18.76	1.40E-04	1.21	4.12E-03	2.79E-01
GO:0045937: positive regulation of phosphate metabolic process	119	7.39	1.41E-04	1.40	4.11E-03	2.80E-01
GO:0010562: positive regulation of phosphorus metabolic process	119	7.39	1.41E-04	1.40	4.11E-03	2.80E-01
GO:0043085: positive regulation of catalytic activity	166	10.31	1.55E-04	1.31	4.45E-03	3.10E-01
GO:0031401: positive regulation of protein modification process	127	7.89	2.02E-04	1.37	5.35E-03	4.02E-01

GO:0016310: phosphorylation	225	13.98	2.44E-04	1.24	6.27E-03	4.87E-01
GO:0043408: regulation of MAPK cascade	82	5.09	2.45E-04	1.49	6.26E-03	4.89E-01
GO:0023014: signal transduction by protein phosphorylation	101	6.27	2.69E-04	1.42	6.74E-03	5.35E-01
GO:0051403: stress-activated MAPK cascade	37	2.30	2.76E-04	1.89	6.91E-03	5.51E-01
GO:0031098: stress-activated protein kinase signaling cascade	38	2.36	3.49E-04	1.85	8.37E-03	6.94E-01
GO:0000165: MAPK cascade	97	6.02	3.52E-04	1.42	8.42E-03	7.00E-01
GO:0045860: positive regulation of protein kinase activity	60	3.73	4.92E-04	1.57	1.13E-02	9.78E-01
GO:0032270: positive regulation of cellular protein metabolic process	150	9.32	5.03E-04	1.30	1.16E-02	9.99E-01
GO:0071902: positive regulation of protein serine/threonine kinase activity	42	2.61	6.57E-04	1.73	1.41E-02	1.30E+00
GO:0051338: regulation of transferase activity	106	6.58	7.29E-04	1.37	1.54E-02	1.45E+00
GO:0033674: positive regulation of kinase activity	62	3.85	0.00103661	1.51	2.06E-02	2.05E+00
GO:0046328: regulation of JNK cascade	26	1.61	0.00144515	1.97	2.71E-02	2.85E+00
GO:0007254: JNK cascade	27	1.68	0.003504773	1.82	5.69E-02	6.77E+00
GO:0043406: positive regulation of MAP kinase activity	30	1.86	0.005214207	1.71	7.90E-02	9.91E+00
GO:0043506: regulation of JUN kinase activity	15	0.93	0.005860228	2.26	8.69E-02	1.11E+01
GO:0032872: regulation of stress-activated MAPK cascade	28	1.74	0.005996233	1.73	8.85E-02	1.13E+01
GO:0070302: regulation of stress-activated protein kinase signaling cascade	28	1.74	0.006407862	1.72	9.22E-02	1.20E+01
GO:0046330: positive regulation of JNK cascade	19	1.18	0.007257919	1.97	1.02E-01	1.35E+01
GO:0032147: activation of protein kinase activity	37	2.30	0.010248089	1.53	1.34E-01	1.86E+01

GO:0051347: positive regulation of transferase activity	69	4.29	0.012444306	1.33	1.57E-01	2.21E+01
GO:0000187: activation of MAPK activity	21	1.30	0.014137245	1.77	1.71E-01	2.47E+01
GO:0043507: positive regulation of JUN kinase activity	12	0.75	0.014625682	2.28	1.75E-01	2.55E+01
GO:0043410: positive regulation of MAPK cascade	54	3.35	0.015968744	1.37	1.86E-01	2.75E+01
GO:0032874: positive regulation of stress-activated MAPK cascade	20	1.24	0.016829087	1.77	1.94E-01	2.87E+01
GO:0070304: positive regulation of stress-activated protein kinase signaling cascade	20	1.24	0.018043562	1.76	2.04E-01	3.05E+01
GO:0007257: activation of JUN kinase activity	7	0.43	0.083131745	2.27	5.57E-01	8.23E+01
Annotation Cluster 20 - Enrichment Score: 4.187						
Term:	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0010558: negative regulation of macromolecule biosynthetic process	162	10.06	4.45E-06	1.41	2.25E-04	8.88E-03
GO:0031324: negative regulation of cellular metabolic process	247	15.34	5.69E-06	1.30	2.75E-04	1.14E-02
GO:0031327: negative regulation of cellular biosynthetic process	167	10.37	6.88E-06	1.39	3.31E-04	1.37E-02
GO:0009890: negative regulation of biosynthetic process	168	10.43	1.03E-05	1.38	4.63E-04	2.05E-02
GO:2000113: negative regulation of cellular macromolecule biosynthetic process	151	9.38	1.13E-05	1.41	5.00E-04	2.26E-02
GO:0010605: negative regulation of macromolecule metabolic process	243	15.09	1.25E-05	1.29	5.37E-04	2.49E-02
GO:0009892: negative regulation of metabolic process	256	15.90	5.22E-05	1.25	1.82E-03	1.04E-01
GO:0051172: negative regulation of nitrogen compound metabolic process	164	10.19	5.77E-05	1.34	1.96E-03	1.15E-01

GO:0045892: negative regulation of transcription, DNA-templated	128	7.95	8.04E-05	1.40	2.60E-03	1.60E-01
GO:1902679: negative regulation of RNA biosynthetic process	134	8.32	8.16E-05	1.38	2.61E-03	1.63E-01
GO:0006357: regulation of transcription from RNA polymerase II promoter	193	11.99	8.58E-05	1.30	2.69E-03	1.71E-01
GO:0051253: negative regulation of RNA metabolic process	138	8.57	9.42E-05	1.37	2.91E-03	1.88E-01
GO:0045934: negative regulation of nucleobase-containing compound metabolic process	148	9.19	2.15E-04	1.33	5.60E-03	4.28E-01
GO:0010629: negative regulation of gene expression	158	9.81	2.45E-04	1.31	6.26E-03	4.88E-01
GO:1903507: negative regulation of nucleic acid-templated transcription	129	8.01	3.25E-04	1.35	7.95E-03	6.47E-01
GO:0006366: transcription from RNA polymerase II promoter	181	11.24	0.002219147	1.23	3.87E-02	4.34E+00
GO:0000122: negative regulation of transcription from RNA polymerase II promoter	82	5.09	0.004290765	1.35	6.72E-02	8.23E+00
Annotation Cluster 21 - Enrichment Score: 3.858						
Term:	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0097190: apoptotic signaling pathway	82	5.09	1.04E-06	1.73	6.38E-05	2.08E-03
GO:0097191: extrinsic apoptotic signaling pathway	40	2.48	4.61E-06	2.19	2.30E-04	9.20E-03
GO:2001236: regulation of extrinsic apoptotic signaling pathway	30	1.86	3.58E-05	2.30	1.32E-03	7.14E-02
GO:2001233: regulation of apoptotic signaling pathway	54	3.35	6.77E-05	1.75	2.27E-03	1.35E-01
GO:0008625: extrinsic apoptotic signaling pathway via death domain receptors	18	1.12	1.61E-04	2.81	4.57E-03	3.21E-01

GO:2001238: positive regulation of extrinsic apoptotic signaling pathway	14	0.87	2.92E-04	3.20	7.23E-03	5.81E-01
GO:1902041: regulation of extrinsic apoptotic signaling pathway via death domain receptors	13	0.81	0.001264897	2.92	2.43E-02	2.50E+00
GO:2001235: positive regulation of apoptotic signaling pathway	26	1.61	0.004476848	1.81	6.98E-02	8.57E+00
GO:2001234: negative regulation of apoptotic signaling pathway	29	1.80	0.006169293	1.70	8.98E-02	1.16E+01
Annotation Cluster 22 - Enrichment Score: 3.705						
Term:	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0002819: regulation of adaptive immune response	32	1.99	2.26E-08	3.08	2.18E-06	4.52E-05
GO:0002822: regulation of adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains	28	1.74	5.11E-07	2.95	3.34E-05	1.02E-03
GO:0002697: regulation of immune effector process	54	3.35	7.00E-07	2.05	4.43E-05	1.40E-03
GO:0002460: adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains	43	2.67	1.20E-04	1.85	3.64E-03	2.40E-01
GO:0002703: regulation of leukocyte mediated immunity	28	1.74	1.32E-04	2.21	3.94E-03	2.64E-01
GO:0002706: regulation of lymphocyte mediated immunity	23	1.43	1.39E-04	2.45	4.11E-03	2.78E-01
GO:0002456: T cell mediated immunity	18	1.12	1.61E-04	2.81	4.57E-03	3.21E-01
GO:0002440: production of molecular mediator of immune response	29	1.80	1.75E-04	2.14	4.80E-03	3.50E-01

GO:0002709: regulation of T cell mediated immunity	14	0.87	1.95E-04	3.32	5.20E-03	3.88E-01
GO:0002700: regulation of production of molecular mediator of immune response	21	1.30	3.36E-04	2.42	8.12E-03	6.69E-01
GO:0002821: positive regulation of adaptive immune response	17	1.06	5.13E-04	2.65	1.17E-02	1.02E+00
GO:0002824: positive regulation of adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains	16	0.99	8.70E-04	2.63	1.76E-02	1.72E+00
GO:0002443: leukocyte mediated immunity	47	2.92	0.001081732	1.63	2.13E-02	2.14E+00
GO:0002699: positive regulation of immune effector process	26	1.61	0.001577059	1.96	2.91E-02	3.10E+00
GO:0002449: lymphocyte mediated immunity	38	2.36	0.002910931	1.64	4.88E-02	5.65E+00
GO:0002705: positive regulation of leukocyte mediated immunity	15	0.93	0.012015881	2.08	1.53E-01	2.14E+01
GO:0002711: positive regulation of T cell mediated immunity	8	0.50	0.014866409	2.99	1.78E-01	2.58E+01
GO:0002708: positive regulation of lymphocyte mediated immunity	13	0.81	0.015309235	2.17	1.81E-01	2.65E+01
Annotation Cluster 23 - Enrichment Score: 3.434						
Term:	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0007596: blood coagulation	51	3.17	3.18E-05	1.84	1.21E-03	6.35E-02
GO:0042060: wound healing	71	4.41	6.39E-05	1.62	2.16E-03	1.28E-01
GO:0007599: hemostasis	51	3.17	1.04E-04	1.76	3.18E-03	2.07E-01
GO:0050817: coagulation	51	3.17	1.04E-04	1.76	3.18E-03	2.07E-01
GO:0009611: response to wounding	80	4.97	1.63E-04	1.52	4.60E-03	3.25E-01
GO:0050878: regulation of body fluid levels	60	3.73	0.002531393	1.47	4.35E-02	4.93E+00

GO:0030168: platelet activation	19	1.18	0.101598659	1.46	6.16E-01	8.82E+01
Annotation Cluster 24 - Enrichment Score: 3.392						
Term:	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0031349: positive regulation of defense response	61	3.79	2.73E-06	1.87	1.43E-04	5.45E-03
GO:0050729: positive regulation of inflammatory response	23	1.43	1.59E-04	2.42	4.51E-03	3.16E-01
GO:0032101: regulation of response to external stimulus	84	5.22	6.74E-04	1.44	1.44E-02	1.34E+00
GO:0032103: positive regulation of response to external stimulus	37	2.30	0.001727162	1.71	3.13E-02	3.39E+00
GO:0050727: regulation of inflammatory response	37	2.30	0.021609879	1.45	2.36E-01	3.54E+01
Annotation Cluster 25 - Enrichment Score: 3.233						
Term:	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0002683: negative regulation of immune system process	60	3.73	1.81E-06	1.91	1.03E-04	3.61E-03
GO:0050863: regulation of T cell activation	49	3.04	3.64E-06	2.02	1.88E-04	7.28E-03
GO:0051250: negative regulation of lymphocyte activation	25	1.55	3.82E-05	2.53	1.39E-03	7.63E-02
GO:0002698: negative regulation of immune effector process	22	1.37	4.21E-05	2.71	1.51E-03	8.40E-02
GO:0007162: negative regulation of cell adhesion	36	2.24	7.99E-05	2.03	2.61E-03	1.59E-01
GO:0050777: negative regulation of immune response	24	1.49	1.24E-04	2.41	3.73E-03	2.47E-01
GO:0022408: negative regulation of cell-cell adhesion	25	1.55	1.59E-04	2.32	4.53E-03	3.16E-01
GO:0002695: negative regulation of leukocyte activation	26	1.61	1.75E-04	2.26	4.81E-03	3.49E-01
GO:0050866: negative regulation of cell activation	28	1.74	1.83E-04	2.17	4.94E-03	3.65E-01

GO:0050868: negative regulation of T cell activation	19	1.18	2.39E-04	2.63	6.19E-03	4.77E-01
GO:1903038: negative regulation of leukocyte cell-cell adhesion	19	1.18	6.29E-04	2.44	1.37E-02	1.25E+00
GO:0002820: negative regulation of adaptive immune response	9	0.56	0.007088534	3.08	1.00E-01	1.32E+01
GO:0070664: negative regulation of leukocyte proliferation	10	0.62	0.103950678	1.79	6.23E-01	8.88E+01
GO:0042130: negative regulation of T cell proliferation	8	0.50	0.134430204	1.86	7.02E-01	9.44E+01
GO:0050672: negative regulation of lymphocyte proliferation	9	0.56	0.153609256	1.71	7.36E-01	9.64E+01
GO:0032945: negative regulation of mononuclear cell proliferation	9	0.56	0.153609256	1.71	7.36E-01	9.64E+01
Annotation Cluster 26 - Enrichment Score: 3.187						
Term:	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0015031: protein transport	223	13.85	6.71E-10	1.48	9.91E-08	1.34E-06
GO:0045184: establishment of protein localization	238	14.78	8.35E-10	1.45	1.17E-07	1.67E-06
GO:0008104: protein localization	274	17.02	5.34E-09	1.38	5.92E-07	1.07E-05
GO:0033036: macromolecule localization	305	18.94	1.79E-08	1.34	1.79E-06	3.57E-05
GO:0070727: cellular macromolecule localization	186	11.55	4.15E-07	1.42	2.80E-05	8.28E-04
GO:0043124: negative regulation of I-kappaB kinase/NF-kappaB signaling	18	1.12	4.90E-07	4.19	3.26E-05	9.79E-04
GO:0034613: cellular protein localization	184	11.43	5.81E-07	1.42	3.77E-05	1.16E-03
GO:1903828: negative regulation of cellular protein localization	29	1.80	5.58E-06	2.57	2.71E-04	1.11E-02
GO:0006886: intracellular protein transport	122	7.58	1.07E-05	1.48	4.77E-04	2.14E-02

GO:0090317: negative regulation of intracellular protein transport	21	1.30	1.32E-05	3.01	5.61E-04	2.64E-02
GO:0051641: cellular localization	262	16.27	1.37E-05	1.27	5.78E-04	2.74E-02
GO:1904950: negative regulation of establishment of protein localization	36	2.24	1.75E-05	2.18	7.20E-04	3.50E-02
GO:0032387: negative regulation of intracellular transport	23	1.43	5.29E-05	2.60	1.84E-03	1.06E-01
GO:0051224: negative regulation of protein transport	33	2.05	7.80E-05	2.11	2.57E-03	1.56E-01
GO:0051235: maintenance of location	44	2.73	7.99E-05	1.87	2.60E-03	1.59E-01
GO:0046907: intracellular transport	171	10.62	8.46E-05	1.32	2.68E-03	1.69E-01
GO:0051649: establishment of localization in cell	208	12.92	9.22E-05	1.28	2.87E-03	1.84E-01
GO:0042306: regulation of protein import into nucleus	30	1.86	1.23E-04	2.15	3.70E-03	2.44E-01
GO:0034504: protein localization to nucleus	51	3.17	1.48E-04	1.73	4.27E-03	2.94E-01
GO:1904589: regulation of protein import	30	1.86	1.67E-04	2.11	4.67E-03	3.33E-01
GO:0042308: negative regulation of protein import into nucleus	15	0.93	2.03E-04	3.14	5.38E-03	4.05E-01
GO:1904590: negative regulation of protein import	15	0.93	2.03E-04	3.14	5.38E-03	4.05E-01
GO:0060341: regulation of cellular localization	97	6.02	2.60E-04	1.43	6.59E-03	5.18E-01
GO:0032880: regulation of protein localization	109	6.77	2.83E-04	1.40	7.07E-03	5.64E-01
GO:0046823: negative regulation of nucleocytoplasmic transport	16	0.99	2.89E-04	2.90	7.19E-03	5.76E-01
GO:0033365: protein localization to organelle	103	6.40	3.22E-04	1.41	7.89E-03	6.40E-01
GO:0051223: regulation of protein transport	88	5.47	3.31E-04	1.45	8.04E-03	6.59E-01
GO:0070201: regulation of establishment of protein localization	95	5.90	3.44E-04	1.43	8.28E-03	6.84E-01

GO:1900181: negative regulation of protein localization to nucleus	16	0.99	4.04E-04	2.82	9.57E-03	8.03E-01
GO:0033157: regulation of intracellular protein transport	50	3.11	5.21E-04	1.65	1.18E-02	1.04E+00
GO:0046822: regulation of nucleocytoplasmic transport	33	2.05	5.37E-04	1.90	1.22E-02	1.07E+00
GO:0042992: negative regulation of transcription factor import into nucleus	11	0.68	5.39E-04	3.67	1.22E-02	1.07E+00
GO:0032386: regulation of intracellular transport	60	3.73	5.45E-04	1.57	1.23E-02	1.08E+00
GO:1903827: regulation of cellular protein localization	69	4.29	5.55E-04	1.51	1.24E-02	1.10E+00
GO:0006606: protein import into nucleus	40	2.48	6.07E-04	1.76	1.35E-02	1.21E+00
GO:0044744: protein targeting to nucleus	40	2.48	6.07E-04	1.76	1.35E-02	1.21E+00
GO:1902593: single-organism nuclear import	40	2.48	6.52E-04	1.76	1.41E-02	1.29E+00
GO:1900180: regulation of protein localization to nucleus	33	2.05	7.42E-04	1.87	1.55E-02	1.47E+00
GO:0050663: cytokine secretion	27	1.68	7.81E-04	2.02	1.60E-02	1.55E+00
GO:0042345: regulation of NF-kappaB import into nucleus	12	0.75	8.94E-04	3.22	1.80E-02	1.77E+00
GO:0042990: regulation of transcription factor import into nucleus	18	1.12	0.001043444	2.41	2.07E-02	2.06E+00
GO:0051170: nuclear import	41	2.55	0.001151953	1.69	2.25E-02	2.28E+00
GO:0042991: transcription factor import into nucleus	18	1.12	0.00118241	2.39	2.29E-02	2.33E+00
GO:0042347: negative regulation of NF-kappaB import into nucleus	7	0.43	0.001585276	5.08	2.91E-02	3.12E+00
GO:0042348: NF-kappaB import into nucleus	11	0.68	0.001901111	3.16	3.41E-02	3.73E+00
GO:0072594: establishment of protein localization to organelle	76	4.72	0.002168813	1.41	3.80E-02	4.24E+00

GO:0032507: maintenance of protein location in cell	17	1.06	0.00219675	2.33	3.84E-02	4.30E+00
GO:1903533: regulation of protein targeting	40	2.48	0.002227962	1.64	3.88E-02	4.36E+00
GO:0045185: maintenance of protein location	17	1.06	0.005271224	2.14	7.92E-02	1.00E+01
GO:0017038: protein import	41	2.55	0.006400891	1.54	9.23E-02	1.20E+01
GO:0051051: negative regulation of transport	52	3.23	0.009214832	1.43	1.24E-01	1.69E+01
GO:0050707: regulation of cytokine secretion	22	1.37	0.009580344	1.81	1.27E-01	1.75E+01
GO:0051220: cytoplasmic sequestering of protein	9	0.56	0.011617115	2.85	1.49E-01	2.08E+01
GO:0042994: cytoplasmic sequestering of transcription factor	6	0.37	0.012156562	4.11	1.54E-01	2.17E+01
GO:1902580: single-organism cellular localization	111	6.89	0.012186809	1.24	1.54E-01	2.17E+01
GO:0051651: maintenance of location in cell	17	1.06	0.013260922	1.94	1.64E-01	2.34E+01
GO:0050710: negative regulation of cytokine secretion	10	0.62	0.013763865	2.57	1.69E-01	2.42E+01
GO:0006605: protein targeting	74	4.60	0.015279006	1.30	1.81E-01	2.65E+01
GO:1904951: positive regulation of establishment of protein localization	57	3.54	0.015604882	1.36	1.83E-01	2.70E+01
GO:0031503: protein complex localization	17	1.06	0.02301841	1.82	2.45E-01	3.72E+01
GO:1902582: single-organism intracellular transport	71	4.41	0.033231847	1.26	3.17E-01	4.91E+01
GO:0072595: maintenance of protein localization in organelle	7	0.43	0.03566124	2.78	3.32E-01	5.16E+01
GO:1903531: negative regulation of secretion by cell	22	1.37	0.03635249	1.59	3.37E-01	5.23E+01
GO:0050709: negative regulation of protein secretion	15	0.93	0.04687626	1.75	4.00E-01	6.17E+01

GO:0046824: positive regulation of nucleocytoplasmic transport	16	0.99	0.050736835	1.69	4.21E-01	6.46E+01
GO:0051048: negative regulation of secretion	23	1.43	0.065041542	1.47	4.87E-01	7.39E+01
GO:0051169: nuclear transport	48	2.98	0.083495951	1.25	5.58E-01	8.25E+01
GO:0006913: nucleocytoplasmic transport	47	2.92	0.091121346	1.25	5.84E-01	8.52E+01
GO:0042307: positive regulation of protein import into nucleus	13	0.81	0.092299951	1.65	5.88E-01	8.55E+01
GO:1904591: positive regulation of protein import	13	0.81	0.103591997	1.62	6.22E-01	8.87E+01
GO:1903829: positive regulation of cellular protein localization	37	2.30	0.119619879	1.26	6.68E-01	9.21E+01
GO:0032388: positive regulation of intracellular transport	30	1.86	0.130987666	1.28	6.96E-01	9.39E+01
GO:0090316: positive regulation of intracellular protein transport	27	1.68	0.136959293	1.30	7.06E-01	9.47E+01
GO:1900182: positive regulation of protein localization to nucleus	15	0.93	0.154081421	1.45	7.37E-01	9.65E+01
GO:0042346: positive regulation of NF-kappaB import into nucleus	4	0.25	0.353229946	1.90	9.32E-01	1.00E+02
Annotation Cluster 27 - Enrichment Score: 3.139						
Term:	Count	%	P-value	Fold Enrichment	Benjamini	FDR
GO:0071356: cellular response to tumor necrosis factor	39	2.42	7.14E-05	1.97	2.37E-03	1.43E-01
GO:0034612: response to tumor necrosis factor	40	2.48	1.68E-04	1.88	4.65E-03	3.35E-01
GO:0033209: tumor necrosis factor-mediated signaling pathway	19	1.18	0.031894593	1.69	3.10E-01	4.76E+01