DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY
Office for Scholarly Communication

Development and Application of a High-Throughput RNAi Screen to Reveal Novel Components of the DNA Sensing Pathway

Citation

Roy, Matthew Stephen. 2013. Development and Application of a High-Throughput RNAi Screen to Reveal Novel Components of the DNA Sensing Pathway. Doctoral dissertation, Harvard University.

Permanent link

http://nrs.harvard.edu/urn-3:HUL.InstRepos:11124826

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story .

Accessibility

© 2013 – Matthew Stephen Roy

All rights reserved.

Dissertation Advisor: Nir Hacohen Matthew Stephen Roy

Development and application of a high-throughput RNAi screen to reveal novel components of the DNA sensing pathway

Abstract

The mammalian immune system has evolved a complex and diverse set of mechanisms to detect and respond to pathogens by recognizing conserved molecular structures and inducing protective immune responses. While many of these mechanisms are capable of sensing diverse molecular structures, a large fraction of pathogen sensors recognize nucleic acids. Pathogen-derived nucleic acids trigger nucleic acid sensors that typically induce anti-viral or anti-microbial immunity, however host-derived nucleic acids may also activate these sensors and lead to increased risk of inflammatory or autoimmune disease. Animal models and humans lacking key DNA nucleases, such as Trex1/Dnase3, accumulate intracellular DNA and develop progressive autoimmunity marked by increased Type-I Interferon (IFN) expression and inflammatory signatures.

Double-stranded DNA (dsDNA) is a potent inducer of the Type-I IFN response. Many of the sensors and signaling components that drive the IFN signature following simulation with transfected dsDNA (also called 'Interferon Stimulatory DNA' or 'ISD') remain unknown. We set out to identify novel components of the ISD pathway by developing a large-scale loss-of-function genetic perturbation screen of 1003 candidate genes. We interrogated multiple human and murine primary and immortalized cells, tested several Type-I IFN reporters, and considered multiple loss-of-function strategies before proceeding with an RNAi screen whereby mouse embryonic fibroblasts were stimulated with ISD and Type-IFN pathway activation was assessed by measuring Cxcl10 protein by ELISA.

Candidate genes for testing in the RNAi screen were curated from quantitative proteomic screens, IFN-beta and ISD stimulated mRNA expression profiles, and a selection of domain-based proteins including helicases, cytoplasmically located DNA-binding proteins and a set of potential negative regulators including phosphatases, deubiquitinases and known signaling proteins.

We identified a number of novel ISD pathway components including Abcf1, Ptpn1 and Hells. We validated hits through siRNA-resistant cDNA rescue, chemical inhibition or targeted knockout. Additionally, we evaluated protein-protein interactions of our strongest validated hits to develop a network model of the ISD pathway. In addition to the identification of novel ISD pathway components, our enriched screening data set may provide a useful resource of candidate genes involved in the response to cytosolic DNA.

Table of Contents

Abstract iii

List of Figures and Tables vii

List of Abbreviations xi

Acknowledgements xiii

Attributions xv

Chapter 1:
Introduction: The Role of Nucleic Acids In Eliciting an Immune Response

1.1 – Introduction 2
1.2 – Nucleic acid recognition by Toll-like receptors 3
1.3 – Detection of RNA by cytosolic sensors 6
1.4 – Detection of DNA by cytosolic sensors 8
1.5 – Inflammasome activation by cytosolic DNA 9
1.6 – Interferon activation by cytosolic DNA sensors 10
1.7 – Endogenous DNA ligands and autoimmunity 17
1.8 – Conclusion 21

Chapter 2:
Development of a High-Throughput Screening Method to Detect Nucleic Acid Responses

2.1 – Introduction: Identification of a cellular based system to detect nucleic acid responses 24
2.2 – Ligands and transfection reagents 25
2.3 – Identification of cells for use in genetic and biochemical screen 33
2.4 – Development of a quantitative assay to detect type-I interferon responses of nucleic acids 46
2.5 – Genetic perturbation using RNA interference 64
2.6 – Screening strategy 69
2.7 – Conclusion 82

Chapter 3:
Generation of a candidate gene set by curation and quantitative proteomics

3.1 – Introduction: candidate gene selection 84
3.2 – Candidate gene selection: interferon-regulated, DNA-stimulated gene from published arrays 84
3.3 – Candidate gene selection: DNA SILAC 90
3.4 – Candidate gene selection: helicases 93
3.5 – Candidate gene selection: cytoplasmic DNA-binding proteins 98
3.6 – Candidate gene selection: putative negative regulators and signaling molecules 99
3.7 – Conclusion 104

Chapter 4:
A High-throughput Loss-of-Function RNAi Screen for of the ISD-Sending Pathway Reveals Identifies Known Components and Novel Regulators

4.1 – Introduction 110
4.2 – Pilot screen: phosphatases and deubiquitinases 110
4.3 – High-throughput loss-of-function RNAi screen 115
4.4 – Database development 119
4.5 – Secondary screening 133
4.6 – Conclusion 143

Chapter 5:
Validation and Characterization of Novel Regulators of the DNA Sensing Pathway

5.1 – Introduction 145
5.2 – Validation of putative DNA-sensors 145
5.3 – Abcf1 149
5.4 – Ifit1 157
5.5 – Reep4 167
5.6 – Putative ISD-sensing pathway signaling molecules 170
5.7 – Putative ISD-sending pathway candidates with no known ISD-interaction partners 171
5.8 – Conclusion 179

Chapter 6:
Concluding Remarks and Network Analysis
6.1 – Overview: Screening development, analysis and outcome 182
6.2 – Protein-protein interaction network analysis 184
6.3 – Predictions and concluding remarks 191

Chapter 7:
Material and Methods 198

References 204

List of Figures and Tables

Figure 1.1 Nucleic acid sensors activate the transcription of type-I IFN and other inflammatory genes. 4

Figure 2.1 Transfection of dsDNA induces a type I IFN response 27

Figure 2.2 Transfection is required to induce nucleic acid-specific type I IFN responses in B6 MEFs 28

Figure 2.3 Transfection reagents influence induction of Type-I IFN response 31

Figure 2.4 Haploid cell lines do not respond to viral stimulation 35

Figure 2.5 Transfection of double-stranded DNA induces a Type I IFN response in primary human bronchial epithelial cells (HBEC) but not in human cell lines 36

Figure 2.6 Human bronchial epithelial cells (HBECs) respond to DNA, are amenable to transfection and RNAi but require RIG-I and MAVS to respond to dsDNA 40

Figure 2.7. Multiple murine cell types respond to dsDNA transfection by inducing Type-I IFNs in a time dependent manner 43

Table 2.1 Developing a cell-autonomous assay for selecting bioactive RNAi treated cells. Generation of Type-I IFN-GFP reporters 48

Figure 2.8 Developing a cell-autonomous assay for selecting bioactive RNAi treated cells. Generation of Type-I IFN-GFP reporters 49

Figure 2.9 Developing a cell-autonomous assay for selecting bioactive RNAi treated cells. Generation of a CD-tagged protein library 52

Figure 2.10 Developing a cell-autonomous assay for selecting bioactive RNAi treated cells. Assessment of cell-surface markers following TLR stimulation in human and murine cells 55

Figure 2.11 Developing a cell-autonomous assay for selecting bioactive RNAi treated cells. Quantitative-RT-PCR dual-reporter 57

Figure 2.12 Developing a cell-autonomous assay for selecting bioactive RNAi treated cells. Luciferase interferon stimulated response element (ISRE) reporter development 58

Figure 2.13 Developing a cell-autonomous assay for selecting bioactive RNAi treated cells. ELISA-based protein detection of Cxcl10..... 63

Figure 2.14 Genetic perturbation using RNAi interference 65

Figure 2.15 Optimization of siRNA conditions in $p53^{-/-}$ MEFs: siRNA transfection conditions 70

Figure 2.16 Optimization of siRNA conditions in $p53^{-/-}$ MEFs: Cell culture conditions effect knockdown efficiency. Control gene panel validation 74

Figure 2.17 ISD stimulation following siRNA knockdown 77

Figure 2.18 Assay workflow 78

Figure 2.19 Z-factor test of control siRNAs confirms a robust screening method 81

Figure 3.1 Generation of a candidate gene set by curation and quantitative proteomics 85

Table 3.1 Candidate Gene List: Published Array Curation 87

Figure 3.2 Quantitative mass spectrometry identifies known components of the ISD sensing pathway 91

Table 3.2 Candidate Gene List: DNA SILAC 94

Table 3.3 Candidate Gene List: Helicases 96

Table 3.4 Cytoplasmic protein prediction matrix 100

Table 3.5 Candidate Gene List: Cytoplasmic proteins 101

Table 3.6 Candidate Gene List: MiniScreen / Phosphatases / Deubiquitinases . 105

Table 3.7 Candidate Gene List: Signaling Molecules 107

Figure 4.1 Pilot Screen: Phosphatases and Deubiquitinases 111

Table 4.1 Negative Regulator Screen: Phosphatases 113

Table 4.2 Negative Regulator Screen: Deubiquitinases 114

Figure 4.2 High-throughput Loss-of-Function RNAi Screen: Control siRNA and replicate assessment indicates a robust screen..... 116

Figure 4.3 High-throughput Loss-of-Function RNAi Screen: Full screening results 120

Table 4.3 A High-throughput Loss-of-Function RNAi Screen: Microarray 125

Table 4.3 B High-throughput Loss-of-Function RNAi Screen: DNA SILAC 127

Table 4.3 C High-throughput Loss-of-Function RNAi Screen: Helicases 128

Table 4.3 D High-throughput Loss-of-Function RNAi Screen: Cytoplasmic DNA-binding proteins 129

Table 4.3 E High-throughput Loss-of-Function RNAi Screen: Signaling Molecules . 130

Table 4.3 F High-throughput Loss-of-Function RNAi Screen: Selected phosphatases and deubiquitinases 130

Figure 4.4 High-throughput Loss-of-Function RNAi Screen: Database development 131

Figure 4.5 Deconvolution of toxic siRNA pools reveals additional ISD pathway candidates 134

Figure 4.6 Secondary screening of top 200 candidates identifies DNA-specific response 137

Figure 4.7 Secondary screening of top 40 candidates identifies DNA-specific responses 140

Figure 5.1 Hmgb2 is a putative regulator of the ISD pathway 146

Figure 5.2 Putative DNA sensors Abcf1, Ifit1 and Reep4 147

Figure 5.3 Abcf1 is a putative regulator of the ISD pathway 151

Figure 5.4 Validation of *Ifit1* and its homologues as putative regulators of the ISD pathway 158

Figure 5.5 Reep4 is a putative regulator of the ISD pathway 168

Figure 5.6 Ptpn1 is a putative negative regulator of the ISD pathway..... 172

Figure 5.7 Sp110 is a putative regulator of the ISD pathway 173

Figure 5.8 Hells is a putative regulator of the ISD pathway 175

Table 6.1 Summary of putative ISD-sensing pathway candidates 185

Figure 6.1 Functional association predictions of SILAC candidates Abcf1, Ifit1 and Reep4..... 187

Figure 6.2 Functional association predictions of Sp110 and putative pro-apoptotic binding partners 188

Figure 6.3 Functional association predictions of Hells 189

Figure 6.4 Functional association predictions of putative negative regulators Ppp6c and Ptpn1 190

Figure 6.5 Putative roles of candidate genes as sensors, negative regulators and chromatin remodelers of the ISD pathway 192

Table 7.1 DNA and RNA ligands 200

Table 7.2 Quantitative RT-PCR Primers 201

Table 7.3 siRNA sequences 203

List of Abbreviations

3P-RNA	In vitro transcribed 3-prime triphosphate RNA
62.ISD	62bp dsDNA ligand derived from a viral RNA segment of influenza PR8 containing a T7 promoter consensus sequence
AdV	Adenovirus
AGS	Aicardi-Goutières syndrome
AIM2	Absent in melanoma 2
ANAs	Anti-nuclear antibodies
AP-1	Activation protein-1
APC	Antigen presenting cell
ASC	Apoptotic speck protein containing a CARD
BMDC	Bone-marrow derived dendritic cells
bp	Base pair
cDCs	Conventional dendritic cells
CMV	Cytomegalovirus
CT-DNA	Calf thymus DNA
DAI	DNA-dependent activator of interferon
DEAH	Aspartate-glutamate-alanine-histidine box
dsDNA	Double-stranded DNA
dsRNA	Double-stranded RNA
DUBA	Deubiquitinating enzyme A
ELISA	Enzyme-linked immunoblot staining assay
EMCV	Encephalomyocarditis virus
ES Cells	Embryonic Stem Cells
GFP	Green fluorescent protein
HCMV	Human cytomegalovirus
HSV	Herpes simplex virus
IFN	Interferon
IFN-β	Interferon beta
IL-6	Interleukin 6
ISD	Interferon stimulatory DNA, 45bp dsDNA ligand
LRRFIP1	Leucine-rich repeat in flightless-I interacting protein 1
MAMPS	Microbe-associated molecular patterns
MAPK	Mitogen-activated protein kinase
MAVS	Mitochondrial antiviral signaling protein
MCMV	Mouse cytomegalovirus
MDA5	Melanoma differentiation-associated gene 5
MDC	Myeloid dendritic cell
MDP	Muramyldipeptide
MEFs	Mouse embryonic fibroblasts
MoDCs	Monocyte derived dendritic cells
MYD88	Myeloid differentiation primary response gene 88
NA	Nucleic acids
NF-κB	Nuclear factor kappa b
pDCs	Plasmacytoid dendritic cells
Poly (dA:dT)	Poly(deoxyadenylic-deoxythymidylic) acid, double-stranded DNA sequence of poly(dA-dT)•poly(dT-dA).
Poly I:C	Polyinosinic-polycytidylic acid

List of Abbreviations, continued:

PR8-RNA	In vitro transcribed RNA derived from the 3' end of Influenza A virus (A/Puerto Rico/8/34(H1N1)) segment 8.
PRRs	Pattern recognition receptors
RIG-I	Retinoic acid-induced gene I
RLUs	Relative luminescence units
RLRs	RIG-I-like receptors
RNAi	RNA-interference
shISD	24bp dsDNA ligand
shRNA	Short-hairpin RNA
siRNA	Small-interfering RNA
SLE	Systemic Lupus Erythematosus
ssDNA	Single-stranded DNA
ssRNA	Single-stranded RNA
TAR	Trans-activation response
TBK1	Traf family member-associated NF-κB activator (TANK)-binding kinase 1
TIR	Toll-IL receptor
TLRs	Toll-like receptors
TNF-α	Tumor necrosis factor alpha
TNFR	Tumor necrosis factor receptor
TRC	The RNAi Consortium
TREX1	3'-repair exonuclease 1
TRIM	Interferon-inducible tripartite-motif
VACV	Vaccinia Virus

Acknowledgements

I have always said that I will never be the best at any one thing. But rather, I will be pretty damn good at a whole bunch of things. It was this sentiment that led me into the world of immunology in the first place. In an online job application, I described myself in the following way: “Generalist/Factotum: A P.I.’s dream candidate.” It was Christophe Benoist who was willing to take the risk of taking me on as a research study coordinator *cum* lab technician *cum* graduate student, and for that, I am extremely grateful. To be honest, I think he wanted me in the lab because I knew a lot about cycling and I had studied jazz in college. Regardless, it was in the Benoist/Mathis lab that I first discovered the beautiful and frightfully complicated world of immunology. It was the particular passion of a postdoc in the lab, Reinhard Obst, that pushed me to pursue my studies further. Thank you, Reinhard.

I would be remiss in saying that the road to this point was an easy one. To my advisor Nir, thank you for your patience, thoughtful advice and kind encouragement throughout this process. To the entire Hacohen lab, I extend to you my sincere thanks, especially to Karolina, my lab-sister. Your tireless optimism and infectious enthusiasm kept me going through my darkest days. It was a true pleasure to work with my collaborator Jayita Sen. Our coffee breaks resulted in some my most vibrant periods of productivity in the lab. I’m looking forward to the next one already.

I would like to extend my thanks to my Dissertation Advisory Committee members, Drs Shannon Turley, Sean Whelan and Kai Wucherpfennig. Thank you for pushing me through to the end. Thank you also to Dissertation Exam Committee, Drs Shannon Turley, Jon Kagan, Christophe Benoist and Igor Kramnik for kindly agreeing to participate.

To my parents, you have always supported me, no matter what, without judgment and always with encouragement. Thank you from the bottom of my heart.

And finally, to my wife, Mo, I have been in graduate school since the day we were married and you have been by my side, my rock, my greatest champion, my sounding board and my best friend for every second of every day between now and then. I have everything I have ever wanted in life... and it's all thanks to you. Now, Mo, the fun is really going to start.

Attributions

Portions of the work discussed in this dissertation have been published in the following manuscript.

"Identification of regulators of the innate immune response to cytosolic DNA and retroviral infection by an integrative approach."

Lee MN, Roy M, Ong SE, Mertins P, Villani AC, Li W, Dotiwala F, Sen J, Doench JG, Orzalli MH, Kramnik I, Knipe DM, Lieberman J, Carr SA, Hacohen N.

Nat Immunol. 2013 Feb;14(2):179-85. doi: 10.1038/ni.2509. Epub 2012 Dec 23. PubMed PMID: 23263557.

Of note, M. Lee developed and conducted DNA SILAC experiments described in Chapter 3.2. RNAi pilot and full screens, including candidate list selections, were carried out and analyzed in cooperation with M. Lee described in Chapter 4. M. Lee performed immunoblots in Figures 5.3, 5.4 and 5.5, as well as cDNA rescue experiments, Trex1-/- experiments and small molecule studies described in Chapter 5. Furthermore, experiments connecting the ISD-sensing pathway to the SET complex were conceived and performed by M. Lee. Cxcl10 and Ifnb GFP-reporters MEFs were produced by W. Li, who also performed the initial experiments (Table 2.1). KBM7 viral infections were performed by J. Carette (Figure 2.4). T. Means assisted with RS4-11 and RAW264.7 stimulation experiments (Figure 2.10). HSVd109 experiments were carried out in collaboration with J. Sen. 293T ISRE reporters were developed in collaboration with B. Shum.

Chapter 1:

Introduction: The Role of Nucleic Acids In Eliciting an Immune Response

1.1 – *Introduction*

The mammalian immune system has evolved a complex and diverse set of mechanisms for recognizing and inducing protective immune responses to invading pathogens. Considered the first line of defense against invading pathogens, the innate immune system is characterized by germline-encoded pattern recognition receptors (PRRs) that recognize conserved motifs essential to the survival of invading pathogens[1]. Detection of pathogens by PRRs triggers signaling pathways that coordinate transcription of hundreds of inflammatory genes, the products of which directly control infection and drive the generation of T and B lymphocyte-mediated immune responses[2]. There are four major PRR families that include Toll-like receptors (TLRs), cytosolic retinoic acid-inducible gene I (RIG-I)-like receptors (RLRs), Nod-like receptors and C-type lectins and function as primary sensors that detect a wide range of microbe-associated molecular patterns (MAMPs)[3, 4]. While these receptors are capable of sensing diverse molecular structures, a large fraction of pathogen detectors recognize nucleic acids. While triggering of nucleic acid detectors by pathogen-derived nucleic acids typically induces anti-viral or anti-microbial immunity, host-derived nucleic acids may also activate these sensors and lead to increased risk of inflammatory or autoimmune disease. As such, the recognition of self nucleic acids by innate immune receptors has been linked to several autoimmune and autoinflammatory disorders.

In the past decade we have witnessed the rapid discovery of many innate sensors and with it, an appreciation for the importance of pathogen detection and responses in guiding the immune response[5, 6]. Among the recent discoveries is the recognition that double-stranded DNA (dsDNA) is a potent inducer of the antiviral, type I Interferon (IFN) response. However, many of the sensors and signaling components that drive the IFN signature following simulation with transfected dsDNA (also called

'Interferon Stimulatory DNA' or 'ISD') remain unknown. Summarized in Figure 1.1, the following discussion details our current understanding of nucleic acid sensing with a particular focus on the DNA sensing pathway members and their associated accessory proteins. Additionally, we discuss how regulation of these sensors may prevent the inappropriate recognition of self-nucleic acids.

1.2 – *Nucleic acid recognition by Toll-like receptors*

Pattern recognition receptors can be broadly characterized into two groups based on cellular localization. Members of the Toll-like receptor family sample pathogenic nucleic acids from the lumenal contents of endolysosomal compartments while a number of emerging receptors detect nucleic acids in the cytosol. Toll-like receptors are the best characterized among PRRs with 10 known TLRs in humans and 12 in mice that recognize conserved microbial motifs broadly shared by pathogens but distinguishable from host molecules. The TLR family can be generally categorized into those that are expressed on the cell surface and recognize microbial lipid and protein components (TLRs 2, 4, 5) and those that are expressed in endosomes and recognize nucleic acids (TLRs 3, 7, 8, 9)[5]. The TLRs implicated in the detection nucleic acids; double stranded RNA (dsRNA) (TLR3)[7], single stranded RNA (ssRNA) (TLR7 and TLR8)[8-10], and hypomethylated cytosine-guanosine (CpG) DNA (TLR9)[11, 12], share similarities in their architecture consisting of an extracellular domain with multiple leucine-rich repeats (LRRs) linked by a transmembrane domain to a conserved cytosolic signaling domain, the Toll/Il-1 receptor homology domain (TIR) domain[13]. TIR domain signaling, driven by adaptor molecules including MyD88 (Myeloid differentiation primary response gene 88) (TLR7, 8, 9) and/or TRIF (TIR-domain-containing adaptor inducing IFN-β) (TLR3) initiates a signal cascade culminating in the activation of nuclear factor kappa b (NF-κB),

Figure 1.1

Nucleic acid sensors activate the transcription of type-I IFN and other inflammatory genes. Localized to endosomal compartments, Toll Like Receptors (TLR) 3, 7 and 9 recognize double-stranded RNA, single-stranded RNA and non-methylated CpG DNA, respectively, leading to the production of type-I IFN and other inflammatory cytokines. RIG-I like RNA helicase receptors (RLRs) recognize viral nucleic acids. After recognition of viral RNA, retinoic acid-inducible gene-I (RIG-I) and Mda5 recruit mitochondria-associated viral sensor (MAVS, also known as IPS-1, Cardif, and VISA) and subsequently activates IRF3 and IRF7 in a TBK1- and IKKi-dependent manner. Cytosolic DNA of microbial or self-origin is a potent trigger of type-I IFN production via the STING–TBK1–IRF3 axis, as well as other proinflammatory cytokines (e.g., TNFa and IL-6), by engaging NF-kB signaling. Cytosolic DNA sensors DAI, RNA Polymerase III, IFI16, DHX36, DHX9, DDX41, LRRFIP1 and KU70 detect distinct DNA species and induce an antiviral response (discussed in detail in the text). The DNA-induced signaling pathway converges on the adaptor STING and the kinase TBK1, which phosphorylates IRF3 to mediate downstream signaling events leading to transcriptional induction of inflammatory genes. In addition to cytosolic DNA, bacterial small molecules c-di-AMP and c-di-GMP act as potent stimulators of the type I IFN response by engaging STING either as a direct sensor via (cGAS) or coactivator (discussed in text).

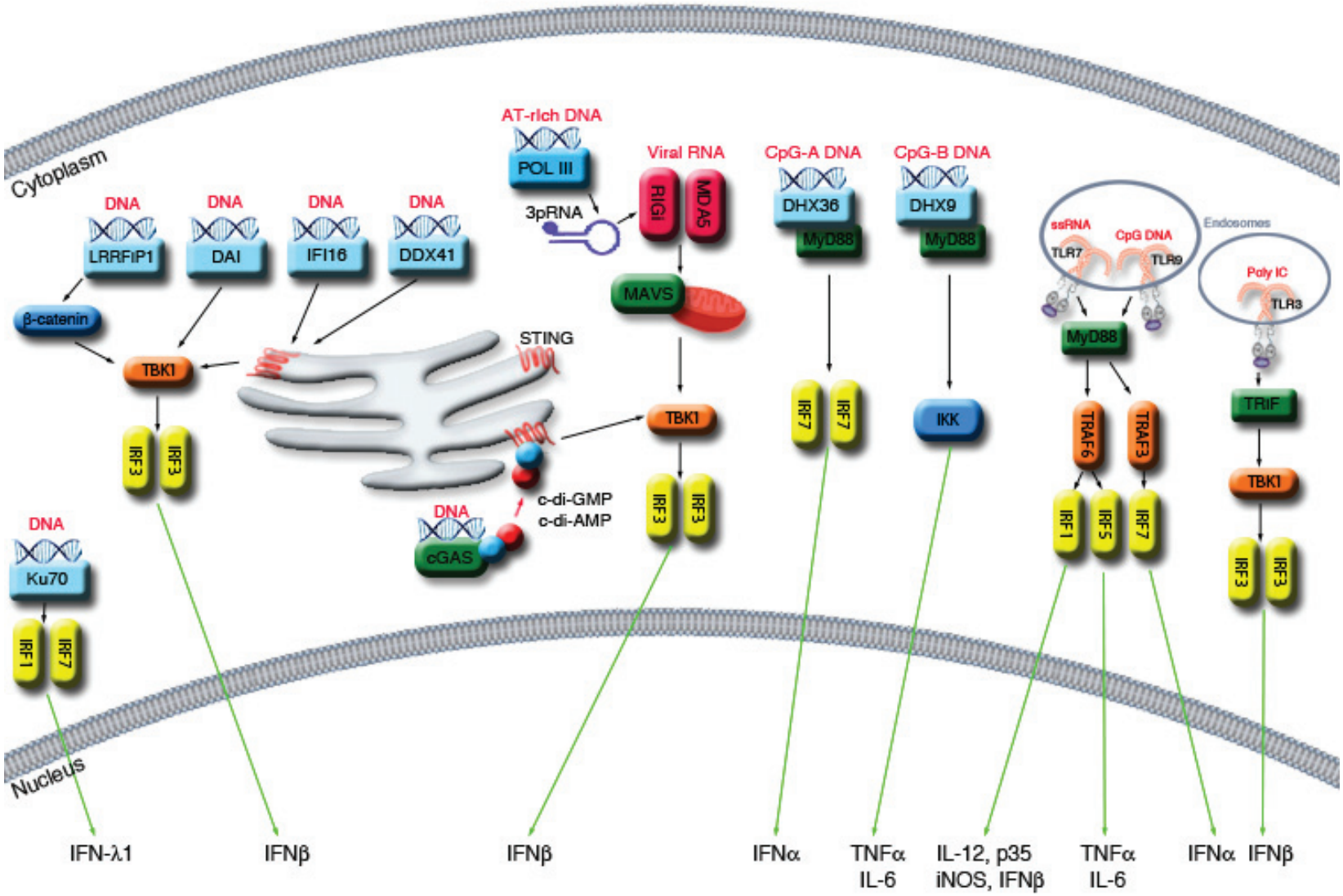


Figure 1.1 (continued)

mitogen-activated protein kinase (MAPK) and interferon regulatory factors 1, 3, 5 and 7 (IRF-1, -3, -5 and -7) [14]. Together these transcription factors drive the production of interferons, cytokines and chemokines as well as the induction of many additional genes important for the initiation of the immune response[15].

TLR9 was the first identified DNA receptor that recognizes nonmethylated CpG motifs, a hallmark of the bacterial genome[11, 12]. Largely restricted to B cells and plasmacytoid dendritic cells (pDCs), TLR9 has been shown to play an important role in detecting mouse and human cytomegalovirus (MCMV and HCMV, respectively), especially within pDCs[16-18], Herpes simplex virus (HSV)-1 and HSV-2[19, 20], and adenovirus[21]. Most of the evidence linking CpG-containing viruses to TLR9 has been generated through in vitro experiments, as TLR9 knockouts do not have a major defect in viral clearance. Together with TLR9 in vivo data, the restricted expression of TLR9, which does not account for DNA-induced immune responses in other cells types, suggests that additional sensing mechanisms must also exist and contribute to antiviral defenses.

1.3 – *Detection of RNA by cytosolic sensors*

While nucleic-acid sensing TLRs sample endolysosomal compartments for ligands degraded from pathogens, a different family of receptors is required to detect pathogen-derived nucleic acids that enter the cytosol. The last decade has revealed numerous classes of sensors for detecting nucleic acids including the RIG-I-like receptors (RLRs), inflammasome-activating sensors and type I IFN producing sensors (ISD sensors) which are the focus of this study.

A hallmark of RLR interaction with viral and synthetic RNA is the induction of type I IFNs. The detection of RNAs by RLRs and the subsequent response requires the

mitochondrial accessory protein MAVS[22-25]. Signaling via RIG-I but not MAVS may also require the membrane-bound protein STING, though the mechanism for this divergence remains unclear[26]. RLR-dependent MAVS signaling drives the recruitment of FADD, RIP1 and TRAF3, for the activation of mitogen-activated protein kinases (MAPKs), including TBK1 and IKK-I. MAPK activation subsequently enables transcription factors activation protein-1 (AP-1), IRF3/7 and NF-κB to translocate into the nucleus where the IFN-β promoter becomes activated[27].

Soon after the identification of TLRs and their cognate ligands, animal models targeting various nucleic acid sensing TLRs revealed reduced but not eliminated responses to their respective stimuli. The search for TLR3 independent sensing of synthetic dsRNA led to the discovery of an entire class of nucleic acid sensors, the DExD/H box RNA helicases known as RIG-I, melanoma differentiation-associated gene 5 (MDA5), and LGP2[28, 29]. RIG-I recognizes 5' triphosphate ssRNAs characteristic of uncapped viral RNAs including flavivirus and orthomyxovirus, and short, dsRNA polymers[30]. The specificity of RIG-I in detecting uncapped 5' triphosphate RNAs marks an important sensing criterion in distinguishing self from non-self RNAs[31]. In contrast, MDA5 may recognize longer RNA, such as synthetic dsRNA (poly I:C) and the genomes of picornaviruses such as encephalomyocarditis virus (EMCV)[32]. Lacking the caspase recruitment domains (CARDs) crucial for the activation of IRF3, the third member of the RLR family, LGP2 was originally described as a negative regulator of RIG-I and MDA5 function[33]. However, recent investigations with LGP2-deficient mice reveal a critical lack of protection against a variety of virus types such as EMCV and furthermore provides evidence that LGP2 acts as a coreceptor for some RIG-I and MDA5 ligands[34].

Evidence implicating NOD2 as an additional sensor adds further to complexity of cytosolic RNA sensing mechanisms. NOD2, a sensor of peptidoglycan derivative

muramyldipeptide (MDP), may also sense ssRNA[35, 36]. Though detection of MDP by NOD2 drives a classic pro-inflammatory signature, similar to its RLR-counterparts, NOD2 stimulation with ssRNA drives a type I IFN response.

1.4 – *Detection of DNA by cytosolic sensors*

The molecular basis of cytosolic DNA recognition is still being elucidated and many details remain unknown. Somewhat surprisingly, the first evidence of a cytosolic DNA response came in 1963 when two independent groups reported that DNA or RNA derived from pathogens or host cells was able to activate chicken or mouse fibroblast to produce interferon[37, 38]. That DNA could induce an immune response remained largely ignored until it was rediscovered decades later. Transfection of dsDNA but not ssDNA derived from pathogens or host cells was found to induce MHC gene expression, as well as various genes involved in antigen presentation and processing including proteasome proteins, invariant chain, HLA-DM, the costimulatory molecule B7.1, and a host of signaling molecules including Stat1, Stat3, MAP kinases and the transcription factor NF-κB[39]. Additionally, it was demonstrated that DNA released from dying cells and introduced into the cytosol of macrophages and bone-marrow derived dendritic cells (BMDCs) induced APCs to upregulate expression of MHC class I and II genes as well as various costimulatory molecules. Response was dependent on transfection of the ligand, regardless of the means to bypass the cell wall. Activation was sequence independent, induced by oligonucleotides (ODNs) as small as 25 bases in length and dose dependent. Additionally there was no response to unmethylated single-stranded CpG DNA demonstrating a divergence in immune responses induced by genomic DNA, independent of TLR9[40-42].

These findings led to two seminal studies demonstrating that the delivery of a synthetic long polymer of poly(dA-dT)·poly(dT-dA) (Poly (dA:dT)) DNA or a 45 base-pair immunostimulatory DNA (ISD) into the cytosol of mouse embryonic fibroblasts (MEFs), macrophages and dendritic cells (DCs) triggered the induction of type I IFN in a Tank-binding kinase 1 (TBK1)/IRF3-dependent manner[43, 44]. Although no receptor for this pathway was identified in these initial studies, the two reports demonstrated that a DNA sensor (or sensors) in the cytoplasm of cells could lead to the activation of the IRF3 pathway independent of TLR9. Since then, a number of groups have identified putative sensors of cytoplasmic DNA resulting in the emergence of two conceptually distinct signaling pathways. The first of these pathways leads to the activation of the inflammasome, characterized by the secretion of proinflammatory cytokines IL-1β and IL-18. The second pathway, the primary focus of this study, leads to the induction of type I IFNs.

1.5 – *Inflammasome activation by cytosolic DNA*

The inflammasome is a multiprotein complex whose activation results in processing of caspase-1, leading to the processing of cytokines such as IL-1β , a key pro-inflammatory mediator that induces pyroptosis and stimulates the recruitment of macrophages and DCs to sites of infection or injury[45, 46]. A role for cytosolic DNA sensing by the inflammasome was first described using various dsDNA ligands, demonstrating a required minimum ligand length (greater than 250bp long) to induce inflammasome activation[47]. Several groups identified AIM2 (absent in melanoma 2) as an essential factor in the initiation of cytosolic DNA-mediated inflammasome activation[48-51]. AIM2 recognition of cytosolic dsDNA requires two critical domains, the HIN200 DNA-binding domain and a pyrin domain that interacts with the inflammasome adapter protein

ASC (apoptotic speck protein containing a CARD). Following DNA transfection and subsequent binding, AIM2 associates with ASC via homotypic pyrin-pyrin domain interactions, which in turn recruit pro-caspase-1, essential for the activation of caspase-1 and proteolytic processing of IL-1β and IL-18.

In antigen presenting cells, AIM2 is indispensible for mounting a response to infection with MCMV and vaccinia Virus (VACV). AIM2-dependent IL-18 secretion and NK cell activation are essential for an early control of mouse CMV infection in vivo[52]. Furthermore, AIM2 deficient mice are more susceptible to infection with *Francisella tularensis*[50, 53] and AIM2-dependent inflammasome activation has been reported for other bacterial infections including *Listeria monocytogenes*[54, 55], and *Mycobacterium tuberculosis*[56]. While multiple lines of evidence point to AIM2 as a cytosolic DNA sensor essential for inflammasome formation and subsequent caspase-1 activation it is completely dispensable for type I IFN production.

1.6 – *Interferon activation by cytosolic DNA sensors*

Following the discovery of the TBK1-dependent, TLR9-independent cytosolic DNA sensing ISD pathway, a growing number of putative sensors have been identified. The growing number of sensors correlates with a diversity of cell-type and ligand-specific responses. Similar to the cytosolic RLR pathways, cytosolic DNA recognition leads to the production of type I IFNs following TBK1 and IRF3 activation. The transmembrane protein STING is essential for facilitating ISD-pathway gene induction, including type I IFN production in fibroblasts, macrophages, and DCs[26, 57]. Additionally, STING is required for both cytoplasmic dsDNA- and HSV-1-activated type I IFN production, demonstrated by STING deficient animals that succumb to lethal HSV-1 infection due to a lack of type I IFN production[57]. A precise mechanism of STING signaling remains

unclear, including whether or not STING itself recognizes intracellular DNA. Furthermore, the adaptors or receptors acting upstream of STING remain largely unknown. A growing number of putative DNA sensors have been identified and may provide further insights in the role of STING-mediated type I IFN production.

DNA-dependent activator of IRFs / DAI

DNA-dependent activator of IFN-regulatory factors (DAI, also known as ZBP-1 or DLM-1) was among the first of the cytosolic DNA sensors to be discovered. DAI was identified as a candidate in a screen for genes induced following IFN-β stimulation[58, 59]. Overexpression of DAI enhances type I IFN production following stimulation with cytosolic DNA. Small interfering RNAs (siRNAs) directed against DAI demonstrated reduced interferon production and resulted in enhanced replication of DNA but not RNA viruses. Furthermore, DAI binding to DNA was demonstrated in vitro and in cells using FRET. It was also demonstrated that DAI initiates DNA-dependent physical interactions with TBK-1 and IRF3. However, subsequent studies in DAI -/- mice revealed normal responses to synthetic and viral dsDNA[60]. The intact ISD sensing pathway indicates that DAI is either dispensable or a there are redundant sensors of cytosolic DNA.

RNA polymerase III

Following the discovery of DAI, two independent studies linked RNA polymerase III to cytosolic DNA sensing, preferentially for AT-rich DNA[61, 62]. Present in the cytosol, RNA polymerase III transcribes AT-rich DNA into uncapped 5' triphosphate moieties, the ligand for RIG-I, which in turn signals via MAVs to induce the type I IFN expression[31]. By generating a ligand to engage the RIG-I pathway, RNA polymerase III is not a cytosolic DNA sensor in the direct sense. RNA polymerase III is present in both human and mouse cell and detects both synthetic Poly (dA:dT) and pathogens including AT-rich virus, like adenovirus and Epstein-Barr virus. Furthermore, it was demonstrated that

RNA polymerase III mediates type I IFN responses during *Legionella pneumophila* and HSV-1 infection[63, 64].

IFI16/Ifi204

The PYHIN family member IFI16 was identified in human monocytes by affinity purification of immune stimulatory 70mer bait derived from vaccinia virus[65]. Similar to AIM2 and other PYHIN family members, IFI16 has and N-terminal pyrin domain two C-terminal DNA-binding HIN200 domains. IFI16 is predominantly expressed within the nucleus, but in some cells types, including macrophages, IFI16 gains access to the cytoplasmic compartment upon stimulation with transfected or viral DNA where it was found to interact in a complex with STING and TBK1 to trigger IFN-β production. In other cell types, including fibroblasts, IFI16 binds viral DNA in the nucleus during productive infection[66]. IFI16 requires signaling from STING to induce type I IFN production, regardless of its cellular localization.

The murine PYHIN protein Ifi204 (p204), which shares a 37% amino-acid identity and a similar domain architecture, is proposed to function in an analogous manner to IFI16. Corroborating homologous function, knockdown of Ifi204 in macrophages and MEFs leads to compromised IFN-β gene induction following stimulation with transfected DNA or HSV-1 infection[65]. Recent studies implicate a functional role of IFI16/Ifi204 in resistance to HSV-1 infection in the corneal epithelium[67]. Furthermore, in a mechanism to evade IFI16-mediate detection, the HSV nuclear protein ICP0 targets IFI16 for degradation[66]. Taken together, these data implicate that both IFI16 and p204 play a role in cytosolic DNA sensing, driving type I IFN expression in a STING dependent manner.

DHX36 and DHX9

The helicases DHX36 and DHX9 have emerged as putative TLR9-independent sensors of cytosolic CpG-A and CpG-B DNA in pDCs, respectively[68]. The aspartate–

glutamate–any amino acid–aspartate/ histidine box–containing helicases, belonging to the DEAH/RHA family of DExD/H helicases, were identified following affinity purification of CpG-DNA-conjugated beads from human primary PDCs. Competition assays demonstrated that DHX9 and DHX36 bound specifically to CpG subtypes, CpG-B or CpG-A, respectively.

These distinct CpG oligodeoxynucleotides induce either type I IFN (CpG-A) or proinflammatory cytokines, such as TNF- α and IL-6 (CpG-B), in pDCs[69]. Similarly, DHX9 activation leads to IRF7-dependent IFN-α production, while activation of DHX36 leads to NF-κB upregulation and the subsequent production of IL-6 and TNFα. Consistent with previous studies that implicated the existence of MyD88-dependent, TLR9-independent DNA sensors in pDCs, knockdown of DHX9 and DHX36 inhibited cytokine production following infection with HSV-1, while response to the RNA virus influenza A was unaffected[20, 70]. Following stimulation, DHX36 and DHX9 interact directly with the Toll/IL-1R domain of MyD88, triggering downstream signaling to activate IRF7 and NF-κB p50, respectively.

Recent studies have further implicated DHX36 and DHX9 as putative sensors of dsRNA[71, 72]. Following stimulation with Poly I:C, DHX36 forms a complex with helicases DDX1 and DDX21 and the adapter TRIF to trigger type I IFN responses[71]. In myeloid DCs, DHX9 interacts with MAVS in response to dsRNA, inducing type I IFN responses[72]. It remains unclear whether DHX36 and DHX9 play a role in the recognition of RNA viruses.

The current results describe complex dual sensing mechanisms mediated by DHX36 and DHX9. The helicases respond to cytosolic CpG DNA oligonucleotides and synthetic dsRNA driving the production of type I IFNs and proinflammatory cytokines. Activation of downstream effectors is initiated by adapters MyD88 (in response to CpGs),

TRIF (through a DHX36-dependent signaling mechanism), and MAVS, following dsRNA activation of DHX9.

DDX41

Following the discovery of the nucleic acid sensing capacity of DHX36 and DHX9, a screen of the DExD/H helicase family identified DDX41 as a putative cytosolic DNA sensor in both mouse and human DCs[73]. Knockdown of DDX41 results in impaired type I IFN and proinflammatory cytokine expression following stimulation with dsDNA oligonucleotides and DNA viruses, including HSV-1 and adenovirus, but not following Poly I:C stimulation or influenza infection. Immunoprecipitation studies revealed that DDX41 interacts with STING in both resting conditions and following stimulation with Poly (dA:dT). Furthermore, DDX41 and was shown to function upstream of IRF3, NF-κB and MAPK.

The current data suggest that DDX41 represents a cytosolic DNA sensor that, similar to IFI16, signals via STING to induce type I IFN and proinflammatory cytokine responses. Unlike the IFN-inducible IFI16, however, DDX41 is constitutively expressed at high levels in immune cells, indicative of a potential role in early immune surveillance of pathogenic dsDNA.

LRRFIP1

Identified in a cDNA screen as an HIV-1 trans-activation response (TAR) RNA-interacting protein, Leucine-rich repeat in flightless-I interacting protein 1 (LRRFIP1) was shown to bind nucleic acids through its N-terminal domain[74]. Subsequently, a targeted siRNA screening focusing on leucine-rich repeat-containing or interacting proteins, found that LRRFIP1 inhibited type I IFN production following stimulation with *Listeria monocytogenes* or infection with VSV[75]. Furthermore, LRRFIP1 knockdown inhibited

IFN production in response to Poly I:C, Poly (dA:dT) and a Z-form DNA analog, Poly (dG:dC).

Interestingly, a novel β-catenin-dependent pathway regulates LRRFIP1-induced type I IFN production. β-catenin, an integral component of the Wnt signaling pathway, is phosphorylated following LRRFIP1-nucleic acid binding where it subsequently translocates to the nucleus. Following nuclear translocation, LRRFIP1 binds to IRF3, leading to enhanced recruitment of the histone acetyltransferase p300, enhancing transcription of the IFN-β gene[76].

Together, these data demonstrate a role for LRRFIP1 in cytosolic nucleic acid recognition leading to enhance IFN-β transcription via a novel β-catenin-dependent signaling pathway.

KU70

A key component of the DNA repair pathway, KU70 was recently identified as a novel regulator of the type III IFN pathway[77]. Similar to type I IFNs, type III IFNs can also exert broad antiviral activity, yet they use a distinct heterodimeric receptor complex (IFN-lR1/IL- 10R2)[78]. KU70 was identified as a putative regulator of cytosolic DNA sensing in cytosolic extracts of HEK293 cells utilizing DNA-conjugated beads. Knockdown of KU70 partially abrogated HSV-1-induced type III IFN induction. Furthermore, KU70 knockdown reduced to type III IFN expression in murine macrophages and DCs following transfection of plasmid DNA in an IRF1- and IRF7-dependent manner.

The role of KU70 in the sensing of cytoplasmic DNA remains to be fully elucidated. It is unclear whether KU70 recognizes bacterial infections and DNA viruses in primary immune cells. Further, the signal transduction pathway between Ku70 and

IRF1 and IRF7 remains elusive. Additionally, it is unclear how the KU70-mediated activation of type III IFN expression contributes to the overall immune response.

Cyclic dinucleotides

Recently, a number of studies have provided compelling evidence of a role for cyclic GMP-AMP (cGAMP) as a novel second messenger triggered by cytosolic DNA that leads to STING-dependent expression of type I IFNs[78-82]. First, it was demonstrated that STING was required to mediate the intracellular response to bacterial derived cGAMP[79]. STING binds these small molecules directly through its C-terminal domain (CTD), leading to TBK1-dependent expression of type I IFNs. The resolution of crystal structures of human STING-CTD bound to cGAMP further demonstrated the putative dual role of STING as an adaptor of DNA sensing and as a direct sensor of cGAMP and potentially of other second messenger molecules[6, 83-85]. An in vitro complementation assay further demonstrated that STING directly binds cGAMP, triggering IRF3 activation and the induction of IFN-β in response to transfected DNA or DNA viruses[82].

It was subsequently demonstrated that chemically synthesized cGAMP potently induced IRF3 phosphorylation and type I IFN induction. Furthermore, cGAMP is induced following HSV-1 and vaccinia virus infection providing strong physiological evidence that cytosolic DNA stimulation leads to STING activation through a novel cGAMP-second messenger. Furthermore, recent evidence links DDX41 as a sensor for cGAMP potentially facilitating cGAMP signaling via STING[86].

Cellular extracts from active fractions of cGAMP-producing L929 cells revealed the interferon-inducible candidate E330016A19, prospectively called cGAS (cGAMP synthase)[87, 88]. Overexpression of cGAS in the presence of STING led to increased type I IFN production, while knockdown compromised cellular responses to transfected DNA or to DNA viruses including HSV-1 or VACV. Together, these studies provide compelling

evidence that cGAS-dependent DNA-sensing operates through a second messenger signaling system leading to the potent induction of type I IFNs.

1.7 – *Endogenous DNA ligands and autoimmunity*

As our knowledge of pattern recognition receptors and their role in microbial defense expands, so too does our understanding that these same receptors may be involved in the initiation of autoimmunity. While PRRs provide the first line of defense against infection, a growing body of evidence supports the notion that autoantigens, particularly nucleic acids, released from apoptotic bodies, necrotic or pyroptotic cells can be recognized by the same receptors. Recognition of self-DNA is largely avoided, as cellular DNA sensors are predominantly located in the cytoplasm while host DNA is typically limited to the mitochondria and nucleus, thus preventing inadvertent activation of proinflammatory cytokines pathways. While the precise series of events between activation of innate immunity and initiation of autoimmunity is not well understood, there are several striking examples of this connection that demonstrate missteps of regulatory mechanisms required to subvert endogenous DNA-induced immune responses. The cellular endonucleases DNase-I, DNase-II, and DNase-III (also known as Trex1) required for the clearance of extracellular, lysosomal, and cytosolic DNA, respectively, represent major regulatory checkpoints that, when altered or absent have deleterious consequences to the host, leading to the inappropriate activation of cytokines including type I IFN production. Next, we review the role of the endonucleases as critical regulatory checkpoints necessary for the prevention of self-DNA-induced activation of proinflammatory cytokines.

DNase-I

Crucial to reducing the potential for self-nucleic acid recognition is the degradation of extracellular nucleic acids. DNase-I is the major nuclease present in serum, urine and secreta and is responsible for the removal of DNA at sites of high cell turnover where it degrades extracellular dsDNA into tri- or tetraoligonucleotides. Mice deficient in DNase-I develop high titers of anti-nuclear antibodies (ANAs), with deposition of immune complexes in the glomeruli and full-blown glomerulonephritis[89, 90]. It was demonstrated that disease severity increased in a DNase-I dose-dependent manner, and correlated with increased levels of apoptosis. Similarly, a subset of patients with Systemic Lupus Erythematosus (SLE) have been found to have decreased levels of circulating DNase-I in the serum and increased apoptosis[91]. Furthermore, mutations in DNase-I are associated with SLE, and low DNase-I activity correlates with glomerulonephritis[92]. Together, these studies point to DNase-I as a crucial regulator of self-DNA, leading to the degradation and destruction of otherwise stimulatory circulating DNA.

DNase-II

DNase-II is expressed in lysosomes, where it degrades DNA from phagocytosed apoptotic and necrotic cells. *Drosophila* deficient in DNase-II demonstrate high levels of endogenous DNA that escaped degradation. The accumulation of endogenous DNA leads to the constitutive expression of the antibacterial genes for diptericin and attacin in a Toll-independent manner[93]. Mice deficient in DNase-II die in utero of anemia, coupled with a massive IFN-β release owing to a failure of definitive erythropoiesis in mouse fetal liver[94, 95]. The failure to degrade nuclei during erythropoiesis and subsequent accumulation of DNA in macrophages leads to a cytokine storm. DNase-II/IFNAR-double-deficient mice are rescued from the anemia phenotype but later develop

polyarthritis, which is dependent on TNF-α[95]. Furthermore, TLR3 and TLR9 deficiency had no effect on the lethality[96].

Adult mice with an inducible knockout form of DNase-II develop severe inflammatory joint disease similar to human Rheumatoid Arthritis, including synovitus with villus proliferation and pannus formation that filled the joint cavity, eroded cartilage, destroyed bones, and occasionally penetrated the bone marrow[97]. Macrophages carrying undigested DNA produced greater amounts of TNF-α and IFN-β when compared to age-matched littermate controls suggesting that macrophages could initiate arthritis. A recent study demonstrated the DNase-II-dependent embryonic lethality could be rescued by loss of STING function, completely preventing the arthritis phenotype, characteristic of DNase-II/IFNAR-double-deficient and inducible DNase-II deficient mice[98].

The sensors that detect DNase-II substrates are unclear, but it is likely that one or more of the aforementioned cytosolic DNA sensor(s) drive the interferon response in the presence of undigested DNA, further highlighting a central role for DNA-induced immune responses in autoimmune diseases.

DNase-III / TREX1

Recently, TREX1 (also known as DNase-III) was identified as a negative regulator of cytosolic DNA sensors. Mutations in 3'-repair exonuclease 1, or TREX1, can lead to autoimmunity as demonstrated in several cases of monogenic familial lupus in which heterozygous mutations in TREX1 gene were found in affected members of a family with chilblain lupus as well as patients with sporadic SLE[99]. Additional mutations have been described as the underlying cause of a congenital neurological disorder with striking similarities to SLE, Aicardi-Goutières syndrome (AGS)[100-104]. The pathology of AGS is indicative of an aberrant immune response; in fact, symptoms closely parallel

those of acquired *in utero* viral infection. Similar to SLE, patients with AGS have elevated levels of IFN-α, progressive autoantibody activation, including elevated levels of IgG and IgM, skin lesions that bear pathological similarities in both diseases and intracranial calcification, with preference for basal ganglia, which occurs in up to 30% of patients with cerebral lupus[105].

TREX1 deficient mice have elevated cytokine production and auto-antibodies leading to lethal autoimmune non-infectious inflammatory myocarditis[104]. Increased cytokine production was linked to the accumulation of cytosolic DNA derived from endogenous retroelements, indicating that one of the functions of Trex1 is to degrade cytosolic DNA derived from reverse transcribed retroelements. Crossing TREX1-deficient mice with mice lacking IRF3 or IFNR rescued TREX1$^{-/-}$ animals from death, linking type I IFN production to the observed autoimmunity and further implicating TREX1 in the regulation DNA activators that would otherwise trigger innate immune signaling-dependent autoimmunity[104]. Recently, a role for TREX1 in HIV infection has described whereby HIV DNA uses TREX1 to rapidly digest viral HIV retroelements to avoid STING-dependent innate immune signaling that would otherwise lead to viral suppression. Together, these data suggest that TREX1 acts as negative regulator of the STING-dependent ISD-sensing pathway [106].

Clearance of apoptotic cells

The role of apoptotic cells, their recognition by components of the innate immune system and subsequent activation of adaptive immunity through what the body may perceive as a “danger signal” has been well established as a potential factor in the development of SLE[107]. Clearance deficiencies, accompanied by the accumulation of apoptotic debris, maturation of antigen presenting cells (APCs) and the subsequent formation of anti-nuclear antibodies has been demonstrated in patients with SLE or with

deficiencies in complement component C1q[108-111]. Apoptotic cells have also been implicated in the activation of T cells in a TLR-independent manner[112]. Exposure of immature dendritic cells to apoptotic cells resulted in production of type I interferons, the upregulation of both MHC class I and II, and the subsequent development of cytotoxic T cells.

Additionally, both nucleosome and dsDNA components of apoptotic cells have been shown to induce dendritic cell activation via toll-dependent and independent pathways[40, 113, 114]. Severe lymphoproliferation and a broad spectrum of autoimmune pathologies characterize mice lacking a TYRO3, AXL and MER (TAM) family of receptor tyrosine kinases, shown to be important for phagocytosis and clearance of apoptotic cells[115, 116]. TAM receptors have also been implicated in the negative regulation of TLR signaling, suggesting that the phenotype of TAM receptor-deficient mice may not be solely attributed to apoptotic clearance defects. Failure to clear apoptotic cells, however, is likely to lead to the release of self nucleic acid ligands and contributes to disease through mechanisms central to the ISD-sensing pathway.

1.8 – Conclusion

The past decade has seen rapid progress in understanding how cells recognize and respond to microbial threats via cytosolic DNA recognition. Multiple innate immune receptors have evolved to sample the lumenal contents of endolysosomal compartments as well as the cytosol to recognize diverse pathogens with a limited set of receptors. The same set of receptors may, in some contexts, detect endogenous nucleic acids resulting in the stimulation of an immune response and subsequent autoimmune pathologies.

Understanding how different cell types recognize and respond to cytoplasmic DNA has profound implications for the diagnosis, prevention, and treatment of a variety

of diseases and may lead to the development of new vaccine strategies. Identifying novel sensors, adaptors, signaling molecules and other components of the ISD-sensing pathway may provide new candidates that could be targeted for therapeutic intervention of infectious, as well as autoimmune, disease.

To this end, we set out to identify and validate novel components of the ISD-sensing pathway. While our primary goal is to identify novel DNA sensor candidates, we hope to also identify components that function as accessory proteins in the signaling cascade following ISD stimulation and leading to the activation of type I IFNs. In the chapters that follow, we describe the development, execution, analysis and validation of a large-scale loss-of-function genetic perturbation screen targeting more than 1,000 candidates genes designed to reveal new components in the ISD-sensing signaling network.

Chapter 2:

Development of a High-Throughput Screening Method to Detect Nucleic Acid Responses

2.1 – *Introduction: Identification of a cellular based system to detect nucleic acid responses*

The following chapter details the development of a reliable, repeatable and robust assay for the detection of the immune response to cytosolic nucleic acids for use in our high throughput loss-of-function RNA-interference (RNAi) screen. Our approach rests on the notion that quantitative measurement of type I IFNs induced by nucleic acids introduced to the cytosol will allow us to use RNAi to identify non-redundant factors required for the normal nucleic acid response in the ISD-sensing pathway.

Many of the genes and pathways that contribute to the immune responses induced by cytosolic nucleic acids remain unknown. To enable genome-wide RNAi screens to find these pathways, we have developed a cellular model system to detect innate immune responses to nucleic acids (NA) in the cytosol, described in the following three steps.

First, to stimulate NA-sensing pathways, we generated and tested several kinds of NA stimuli: (1) 48bp and 24bp dsDNA or immunostimulatory DNA (ISD), which are the major focus of our studies; (2) synthetic and in vitro transcribed ssRNA and dsRNA ligands specific to other (TLR-dependent and independent) sensors including TLR3, TLR7/8 and RIG-I; (3) negative controls ligands, including a TLR9 ligand, unmethylated CpG ssDNA, and an inactive 12bp dsDNA oligonucleotide. Second, we identified an optimal cell line that is able to withstand multiple passages, can be amplified rapidly for use in screens, is amenable to siRNA transfection and/or lentiviral infection and is not sensitive to nonspecific effects of RNAi. Multiple human and murine cell types show strong type I interferon responses to nucleic acid ligands. We identified $p53^{-/-}$ MEFS as an optimal screening cell line for their reproducibility, low sensitivity to transfection and strength of response. Third, to identify the most robust assays, we tested quantitative

assays to detect type I interferon responses, including CD-tagging, GFP and Luciferase reporters, quantitative PCR-based interferon detection and protein detection by ELISA.

We considered two general approaches to our genetic perturbation screens: genome-wide lentiviral shRNA screens and arrayed candidate siRNA screens. For both approaches, optimization of the genetic perturbation system required refinement of multiple conditions including transfection reagent selection, volume, and order of RNAi treatment. Additionally, we assessed siRNA concentration, time to optimal knockdown, cell seeding density and media conditions. Finally, dsDNA stimulation variables, including ligand to transfection ratio, complex incubation time, supernatant collection time and cell survival, were considered following siRNA knockdown.

The screening workflow described below is robust, reproducible and simple to implement. We assessed the validity of our system with the Z-factor scoring method, demonstrating that our screening performance was indicative of a robust screen. The subsequent screen of dsDNA sensing factors resulted in a valuable resource of candidates and novel regulators of the ISD-sensing pathway.

2.2 – *Ligands and transfection reagents*

First we recapitulated recent findings that support the evidence of a cytosolic nucleic acid sensor in mouse embryonic fibroblasts (MEFs) [40, 42-44]. MEFs provide a simple model for the study of innate immunity because they allow infection with various viruses and effectively express type I interferon genes[14]. Zero to eight hours following transfection of 1ug/ml of a 48bp dsDNA ligand or interferon stimulatory DNA (ISD), a 24bp dsDNA ligand or short ISD (shISD), single stranded 20-mer Type-B CpG or Polyinosinic-polycytidylic acid (Poly I:C), a synthetic analog of double-stranded RNA (dsRNA), total RNA was purified at different time points from C57/Bl6 MEF and lysates

were assessed by quantitative RT-PCR (qRT-PCR) for the expression of type I IFNs, including *Ifn-β*, *Cxcl10*, and *Mx1* (Figure 2.1). These cytokines are hallmarks of the first phase immune response, regulated by Irf3 and induced by cytosolic nucleic acids[14]. Experiments aimed at recapitulating the described studies demonstrated robust interferon expression in response to lipid-transfected ISD (described in [44]) that increased throughout an 8-hour time course (Figure 2.1).

Responses were dependent on transfection but were not induced by transfection reagents alone (Figure 2.2A). We confirmed that the responses are sequence-independent, length-dependent but not concentration dependent. Contrary to previous studies[42, 117], 24bp dsDNA ODNs transfected in equimolar amounts to 48bp DNA ODNs consistently induce the expression of *Cxcl10* an order of magnitude lower than the 48bp ODN. However, we demonstrated 12bp dsDNA does not induce *Ifn-β* (not shown) or *Cxcl10* expression, suggesting a minimum length DNA ODN is required to stimulate the IDS-sensing pathway. Synthetic dsDNA polymers (Poly (dA:dT), a repetitive synthetic double-stranded DNA sequence of poly(dA-dT)•poly(dT-dA) and a synthetic analog of B-DNA, induce *Ifn-β* or *Cxcl10*. To discriminate between specific responses to dsDNA and non-specific responses to other nucleic acids, we developed a series of nucleic acid ligands to function as positive and negative controls for use in our secondary screening and proteomics approaches. In addition to 12bp dsDNA that does not stimulate IFN; TLR9-dependent, single-stranded bacterial B-type CpG, a synthetic oligonucleotide that contains unmethylated CpG dinucleotides, did not stimulate type I IFN directly or with lipid-mediated transfection(Figure 2.2B).

Viral and synthetic RNA ligands require RIG-I and MDA-5 to induce an immune response. Thus, we developed a number of ligands specific to the cytosolic RNA sensing pathway. In addition to synthetic double-stranded RNA polyinosine-polycytidylic acid

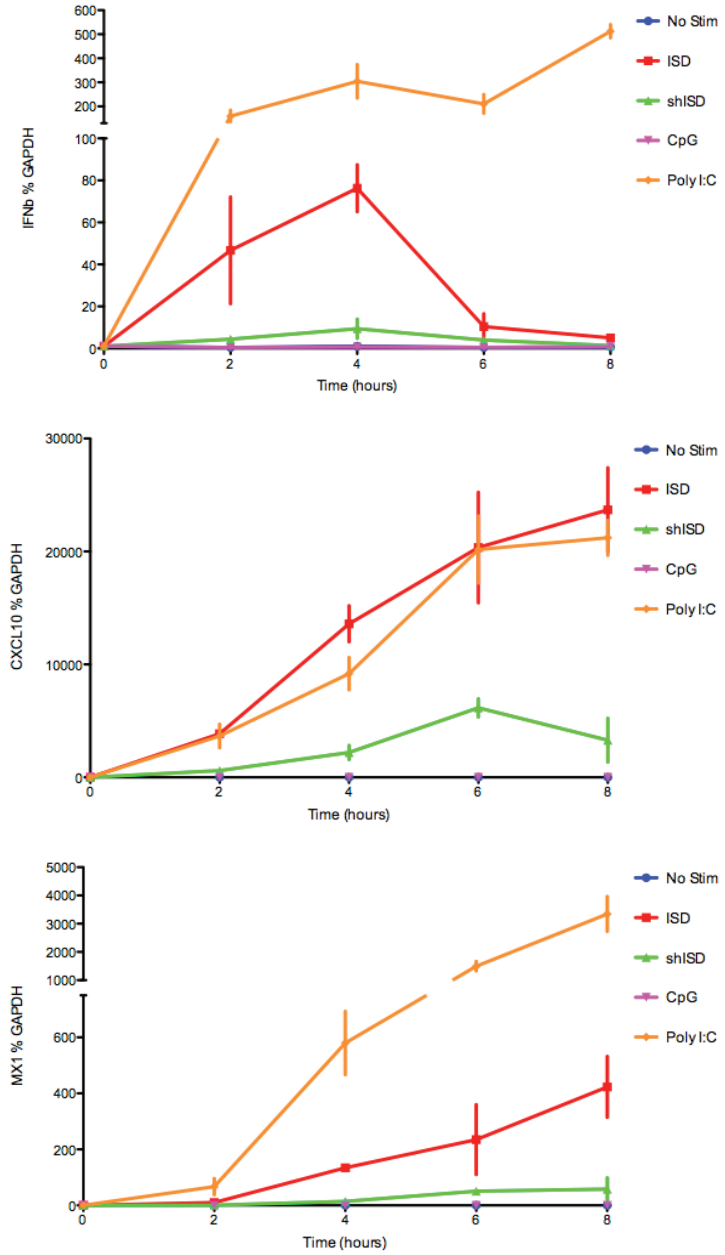


Figure 2.1

Transfection of dsDNA induces a type I IFN response. B6 MEFs were transfected with 45-base pair double-stranded DNA ligand (ISD), 24-base pair double-stranded DNA ligand (shISD), Type-B CpG or poly I:C for 2 to 8 hours. Lysates were collected and total RNA was isolated. Quantitative RT-PCR of Type-I interferon genes was performed from cDNAs prepared from RNA isolates. From top to bottom, *Ifnb*, *Cxcl10* and *Mx1* were measured relative to percent *Gapdh*.

Figure 2.2

Transfection is required to induce nucleic acid-specific type I IFN responses in B6 MEFs. A-C) B6 MEFs are transfected with 1ug/mL of the indicated DNA and RNA ligands for 0-8 hours. *Cxcl10* induction was measured by quantitative RT-PCR. D) B6 MEFs are transfected with ligands with the indicated concentrations for six hours. *Cxcl10* induction was measured by quantitative RT-PCR. E) Lipofectamine 2000 was titrated in combination with 1ug/mL of ligand at the indicated ratios in 96-well plates. *Cxcl10* induction was measured by quantitative RT-PCR.

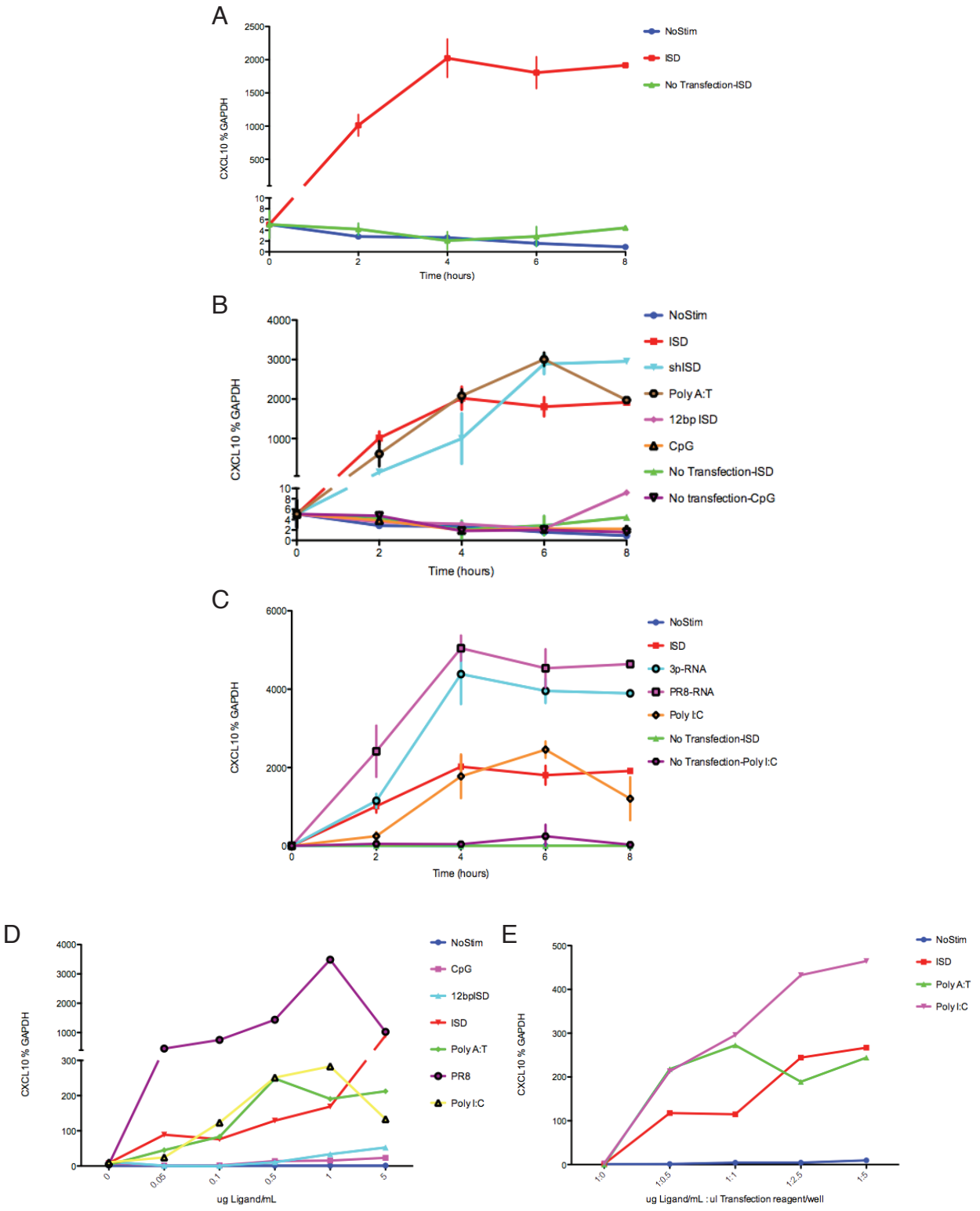


Figure 2.2 (continued)

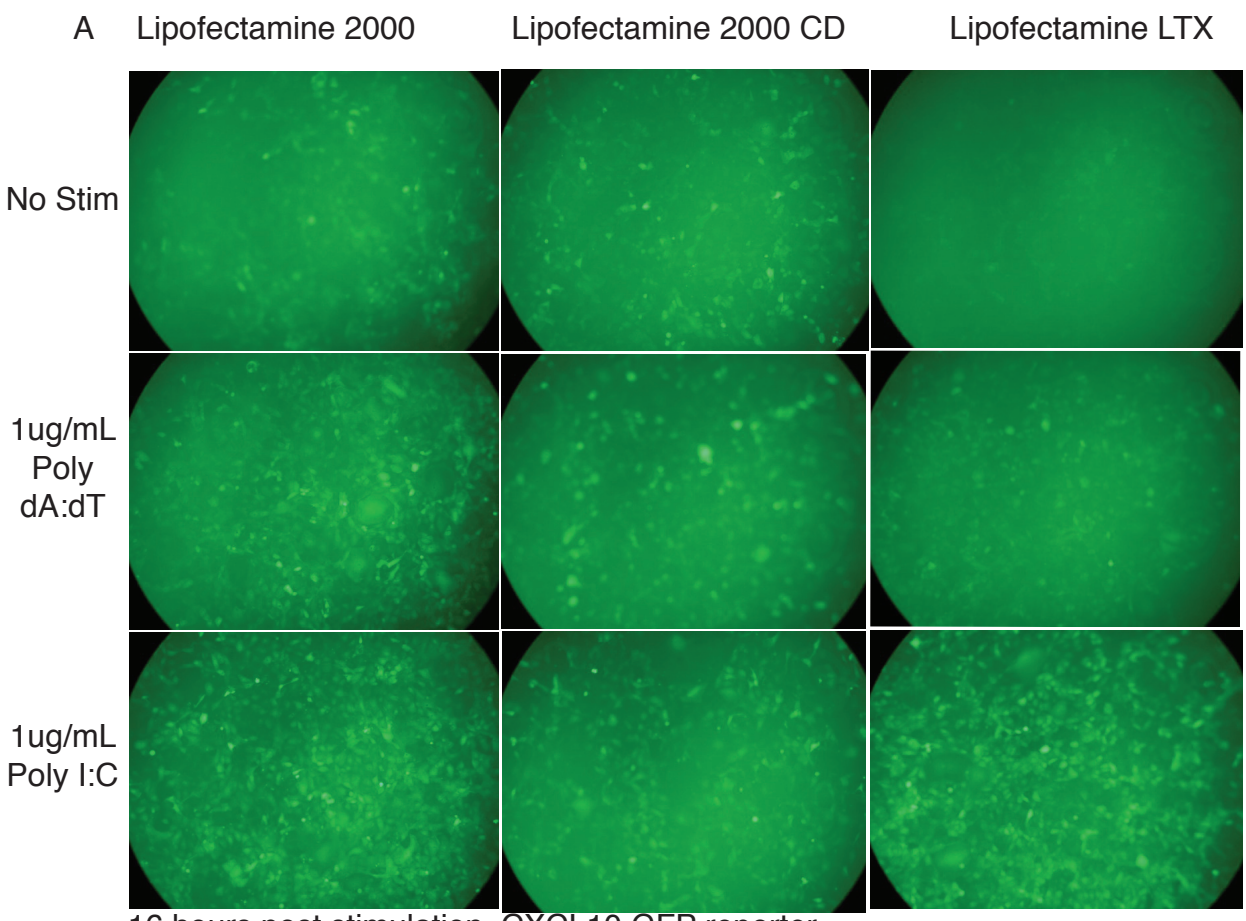
(Poly I:C) which, upon transfection, resembles the RNA of infectious viruses and elicits a consistent and strong interferon response, we used the T7 RNA polymerase to selectively synthesize single stranded RNA bearing a native 5' triphosphate, required for RIG-I mediated antiviral responses[31, 118]. We have developed and tested three in vitro transcribed (IVT) RNAs. In addition to the 19-nucleotide immunostimulatory ssRNA 9.2s (3p-RNA)[119], we generated two 50-nucleotide ssRNAs derived from the 3' end of Influenza A virus (A/Puerto Rico/8/34(H1N1)) segments 5 and 8. These sequences were selected for their potential immunostimulatory effects based on uridine content[120], containing 12 (low) and 21 (high) uridine nucleotides in segments 5 and 8 (PR8-RNA), respectively. We demonstrated the stimulatory effect of these ligands (3p-RNA and PR8-RNA) to induce *Ifn-β* and *Cxcl10* following transfection (Figure 2.2C). We continued to use Cxcl10 as a reliable readout for interferon-inducible genes because of its high induction in response to nucleic acids and its dependence on IRF3[121].

Increased concentration of dsDNA ligands correlates with increased type I IFN expression (Figure 2.2D). High concentrations (5ug/mL) of IVT-RNA are toxic as demonstrated by increased *Gapdh* values of two to three cycles (data not shown) and a decrease *Cxcl10* expression at higher concentrations. We also demonstrated that the ratio of ligand to transfection reagent influences the type I IFN response. Transfection efficiency can also be titrated by increasing the ratio of transfection reagent to ligand (Figure 2.2E).

To assess the influence of transfection reagents on transfection efficiency and toxicity, three transfection reagents were compared side-by-side using MEFs stably expressing a Cxcl10-PLJM6-GFP reporter as a proxy for type I IFN responses (provided by A. Luster, Massachusetts General Hospital, Boston), (Figure 2.3A). We considered the stimulatory capacity of transfection reagents alone, by adding two transfection

Figure 2.3

Transfection reagents influence induction of Type-I IFN response. A) B6 MEFs stably expressing a Cxcl10-PLJM6-GFP reporter were treated with the indicated transfection reagents, either alone or with dsDNA or dsRNA ligands at a 1:5 ratio of ligand to ul of transfection reagent. Sixteen hours following stimulation Cxcl10 expression was assessed by fluorescence microscopy. B) Influence of penicillin/streptomycin presence in MEF media on Cxcl10-GFP expression was compared 16 hours following transfection dsDNA complexes formed with Lipofectamine LTX. Three hours following transfection with dsDNA/LTX complexes media was replaced and compared to wells with the media unchanged.



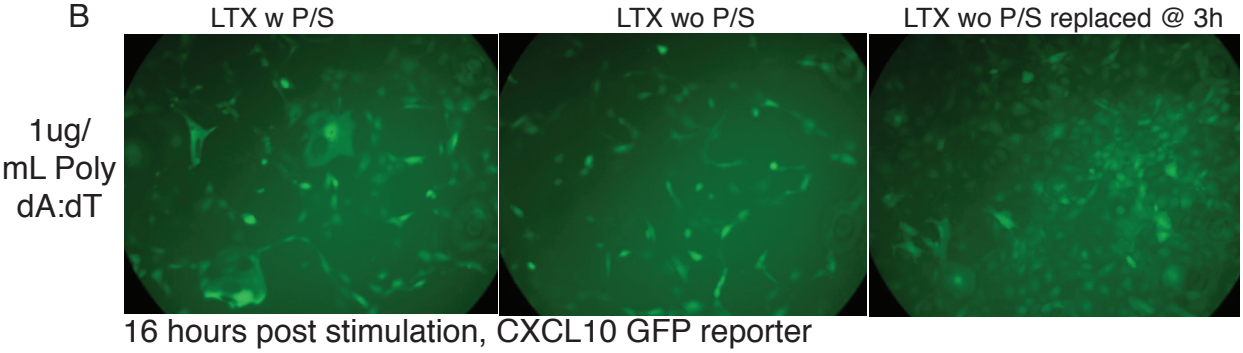


Figure 2.3 (continued)

reagents (Lipofectamine 2000, Lipofectamine 2000CD) directly to culture without the addition of nucleic acids and demonstrated increased GFP expression. An additional transfection reagent, Lipofectamine LTX, does not appear to induce GFP expression. Previous experiments failed to demonstrate type I IFN expression induction by transfection reagent alone suggesting that increased GFP expression may be toxicity-induced auto-fluorescence. Regardless of transfection reagent, Cxcl10-GFP expression increased following transfection of lipid/ligand complexes. Presence or absence of penicillin/streptomycin does not appear to effect transfection efficiency (Figure 2.3B). Furthermore, replacing dsDNA complexes three hours following transfection decreased toxicity and increased Cxcl10-GFP reporter expression.

2.3 – *Identification of cells for use in genetic and biochemical screens*

Critical to our screening system is a robust cell line that is able to withstand multiple passages, is amenable to scalability, is easily infectible, transfectable and is not sensitive to immunogenic effects of RNAi. We considered several human and murine candidate cell lines, including: a) KBM7, the chronic myeloid leukemia (CML) cell line with a haploid karyotype except for chromosome 8; b) human embryonic kidney 293 cells; c) adenocarcinomic human alveolar basal epithelial A549 cells; d) HeLa; e) primary human bronchial epithelial cells (HBEC); f) GM-CSF-differentiated bone-marrow dendritic cells (BMDCs); g) mouse embryonic fibroblasts (MEFs) from both C57BL/6 and Balb/c strains; h) monocyte/macrophage-like RAW 264.7 cells; and i) p53 deficient ($p53^{-/-}$) MEFs.

Evaluation of human haploid, immortalized and primary cells

Loss-of-function genetic screens in haploid cells have identified pathogen host factors[122]. In a method that parallels genetic approaches in haploid yeast, insertional

mutagenesis can be implemented as a screening method by generating null alleles in KBM7 cells. Nucleic acid ligand transfection induced no detectable type I IFN response, nor expression of inflammasome response genes, IL1b, MIP1a and TNFa by quantitative RT-PCR (data not shown). Infection with Adenovirus (AdV), a non-enveloped, double-stranded, linear DNA virus, produces type I IFN, interleukin-6 (IL-6) and tumor necrosis factor alpha (TNF-α) in an IRF3-dependent manner in BMDCs[123]. Infection with Vaccinia virus or AdV in KBM7 wild-type cells or a gene-trap isolate deficient in FADD, did not induce the expression of *IFN-β*, *CXCL10*, *IFIT1*, *MX1* (Figure 2.4) or Il1b, MIP1a and TNFa (data not shown). Without a detectable nucleic acid response in the haploid cells, we then considered both primary and immortalized human cell lines to study the ISD-pathway.

Immortalized human cell lines (HEK293, A549 and HELA) do not respond to transfected dsDNA (Figure 2.5 A-C). Over a time course of eight hours, only Poly I:C and IVT-RNA induced a type I IFN response in HEK293. Primary human bronchial epithelial cells (HBECs), however, responded strongly to synthetic dsDNA (Poly (dA:dT)) six hours following transfection (Figure 2.5 D). In addition to increased expression of type I IFN genes (*CXCL10*, *IFIT1*, *IFN-β* and *MX1*), there was moderate expression of both *IL-6* and *TNF-α*. DNA ligands ISD and 62.ISD, a 62 base pair double-stranded DNA oligonucleotide, a influenza PR8-based derivative of our IVT-RNA preparation, failed to induce *IFN-β* and only moderate amounts of *CXCL10* and *IFIT1*.

To further test the ISD-sensing response in HBECs we developed a Luciferase interferon stimulated response element (ISRE) reporter assay using 293T cells as a surrogate cell for transferred supernatants. Supernatants from HBECs stimulated for six hours were transferred to the ISRE reporters. We demonstrated that the type I IFN response to poly (dA:dT) stimulation in HBECs could be titrated and is sensitive to

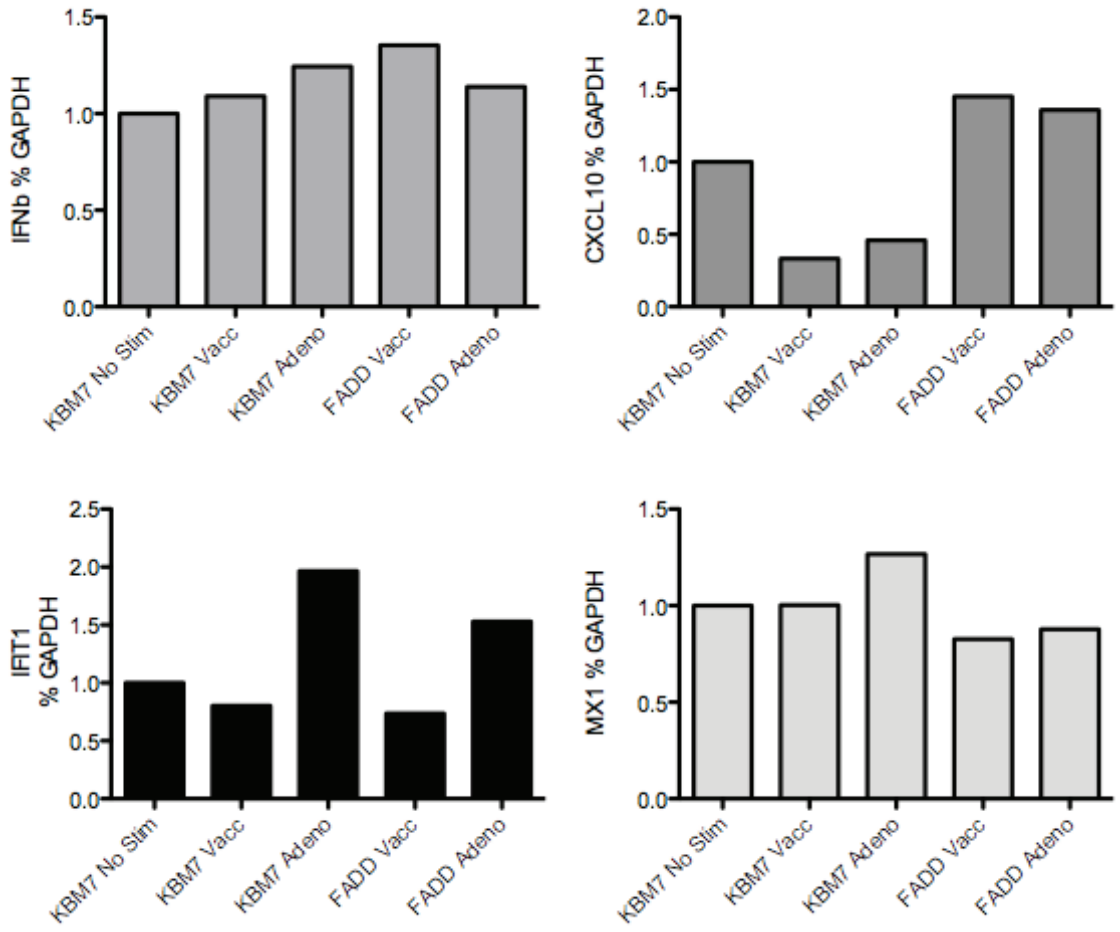


Figure 2.4

Haploid cell lines do not respond to viral stimulation. KBM7 and gene-trap isolate FADD-deficient KBM7 cells were infected with Vaccinia virus and Adenovirus for 24 hours. Type I IFN expression was measured by quantitative RT-PCR for the indicated genes.

Figure 2.5

Transfection of double-stranded DNA induces a Type I IFN response in primary human bronchial epithelial cells (HBEC) but not in human cell lines. A-C) HEK293, A549, and HELA were stimulated for 0-8 hours with 1.0ug/mL dsDNA and 0.3ug/mL RNA ligands. Lysates were collected and total RNA was isolated. Quantitative RT-PCR of Type-I interferon genes was performed from cDNAs prepared from RNA isolates. A) Expression of *IFNb*, *CXCL10* and *MX1* following stimulation was measured. B-C), Expression of Type-I IFN genes were measured as indicated. D) HBEC were stimulated for 6 hours with the indicated ligands (x-axis and legend). Expression of *CXCL10*, *IFIT1*, *IFNb*, *IL-6*, *MX1* and *TNFa* were measured by quantitative RT-PCR.

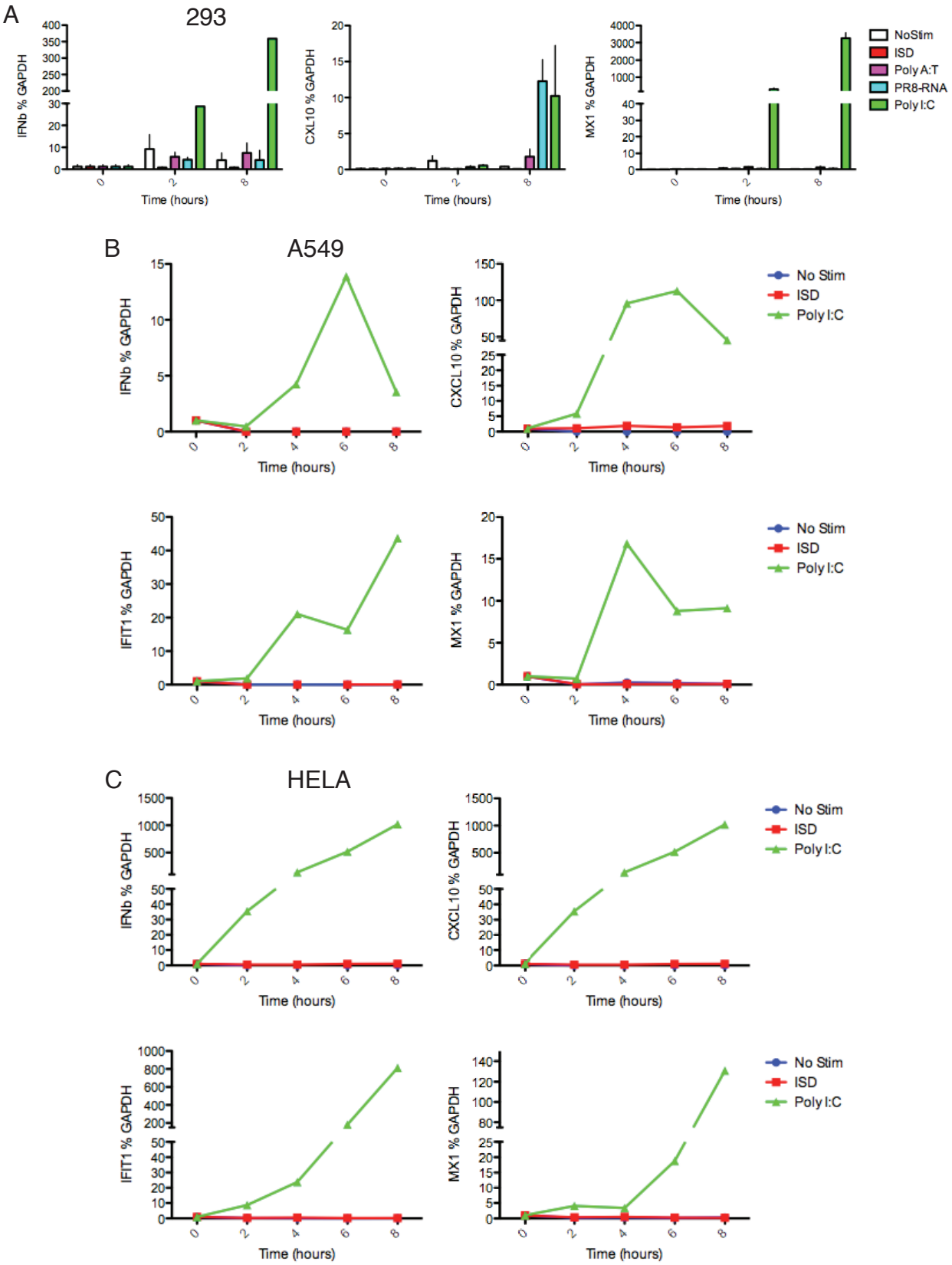


Figure 2.5 (continued)

D

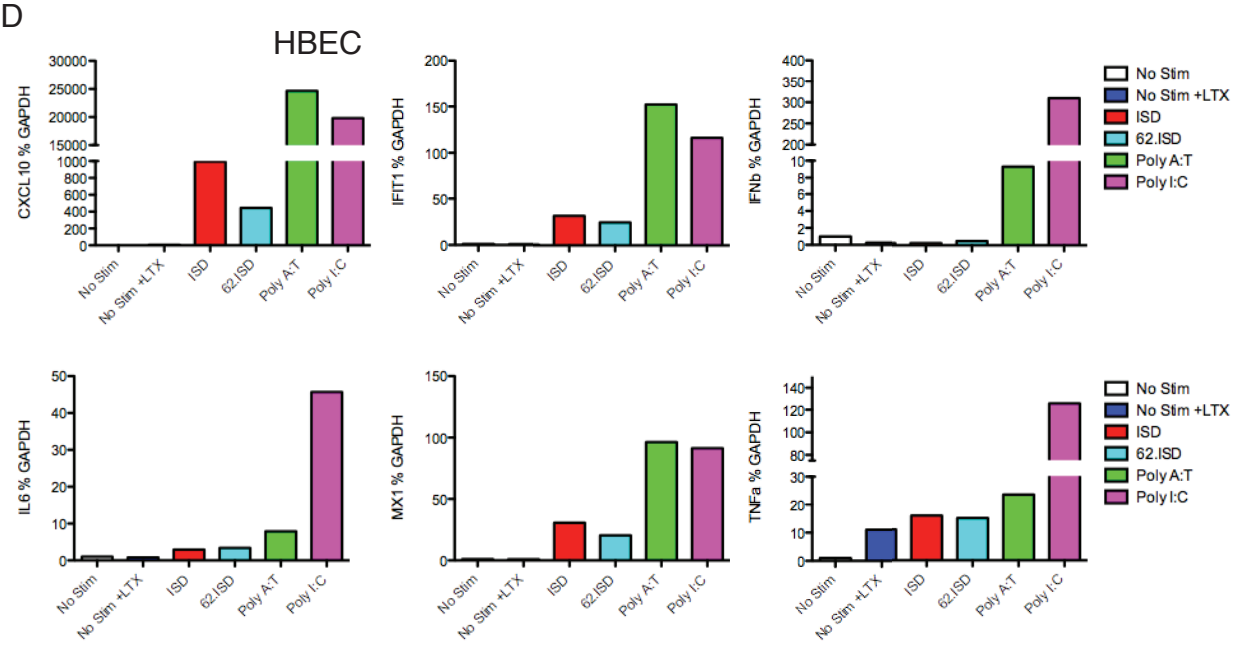


Figure 2.5 (continued)

ligand:transfection reagent rations (Figure 2.6A). Increased concentrations of Poly (dA:dT) correlated with increased ISRE response. Furthermore, transfection with both poly dA:dT and calf-thymus DNA (CT-DNA) induced CXCL10 protein in a time-dependent manner as measured by ELISA (Figure 2.6A).

To test the efficacy of HBECs as a potential screening cell, we first transfected cells with siRNAs targeting *IRF3* and *AIM2*, a gene required for the initiation of the DNA-dependent inflammasome response[48, 50, 51] (Figure 2.6B). Seventy-two hours following transfection, cells were stimulated with Poly (dA:dT). Supernatants transferred to ISRE reporter cells show a 5-fold decrease in the DNA response in *IRF3*-knockdown cells, as well as a 2-fold increase following *AIM2* knockdown, relative to non-targeting control siRNA treated cells treated HBECs. Quantitative RT-PCR of *IRF3* gene expression reveals a 20-fold decrease relative to non-targeting control (P-value <0.0001, Student's t-test). Following the optimization of siRNA conditions, we tested the response of a full panel of control genes to Poly (dA:dT) transfection in HBECs (Figure 2.6C). As expected, siRNAs targeting *IFN-β* and *IRF3* reduce the response to dsDNA by 17 and 84 fold, respectively (P-value <0.001, Student's t-test). Both *DAI* and *IRF7* knockdown resulted in moderately reduced DNA responses. Surprising, at the time, was the near complete elimination of the dsDNA response following knockdown of *MAVS* and *RIG-I*, genes required for the response to RNA ligands (31 and 28-fold, respectively, P-value <0.01). In an independent experiment, we validated these findings by knocking down *IRF3* and *RIG-I*, stimulating with Poly (dA:dT) for 6 hours and measuring the expression of interferon response genes *CXCL10* and *MX1* (Figure 2.6D).

We demonstrated that transfected Poly (dA:dT) fails to induce a type I interferon response in HBECs following knockdown of both *RIG-I* and *MAVS*. These data conflict directly with mouse models that show intact interferon responses to dsDNA in the

Figure 2.6

Human bronchial epithelial cells (HBECs) respond to DNA, are amenable to transfection and RNAi but require RIGi and MAVS to respond to dsDNA. A) HBECs were stimulated with the indicated concentration of poly (dA:dT) and ratio of transfection reagent (Lipofectamine LTX) for six hours (left panel) or 0-24 hours (right panel). Supernatants were collected and added to 293T ISRE reporter cells (left panel) or to a CXCL10 ELISA. ISRE standard curve, left inset. B) HBECs were transfected with siRNAs targeting *AIM2* and *IRF3* for 72-hours and then stimulated with 1ug/mL Poly dA:dT for 6 hours. Supernatants were transferred to 293T ISRE reporters. Knockdown efficiency over a course of indicated siRNA concentrations were measured by measured by quantitative RT-PCR (p-value <0.0001, two-way ANOVA). C-D) An siRNA panel of the indicated control genes were transfected into HBECs as described above and stimulated with poly (dA:dT) for 6 hours. Interferon response was detected by ISRE reporter (C) or CXCL10 and MX1 gene expression by quantitative RT-PCR (D), (*, p-value <0.01, **, p-value <0.001, paired t-test, siRNA control compared to stimulated non-targeting control).

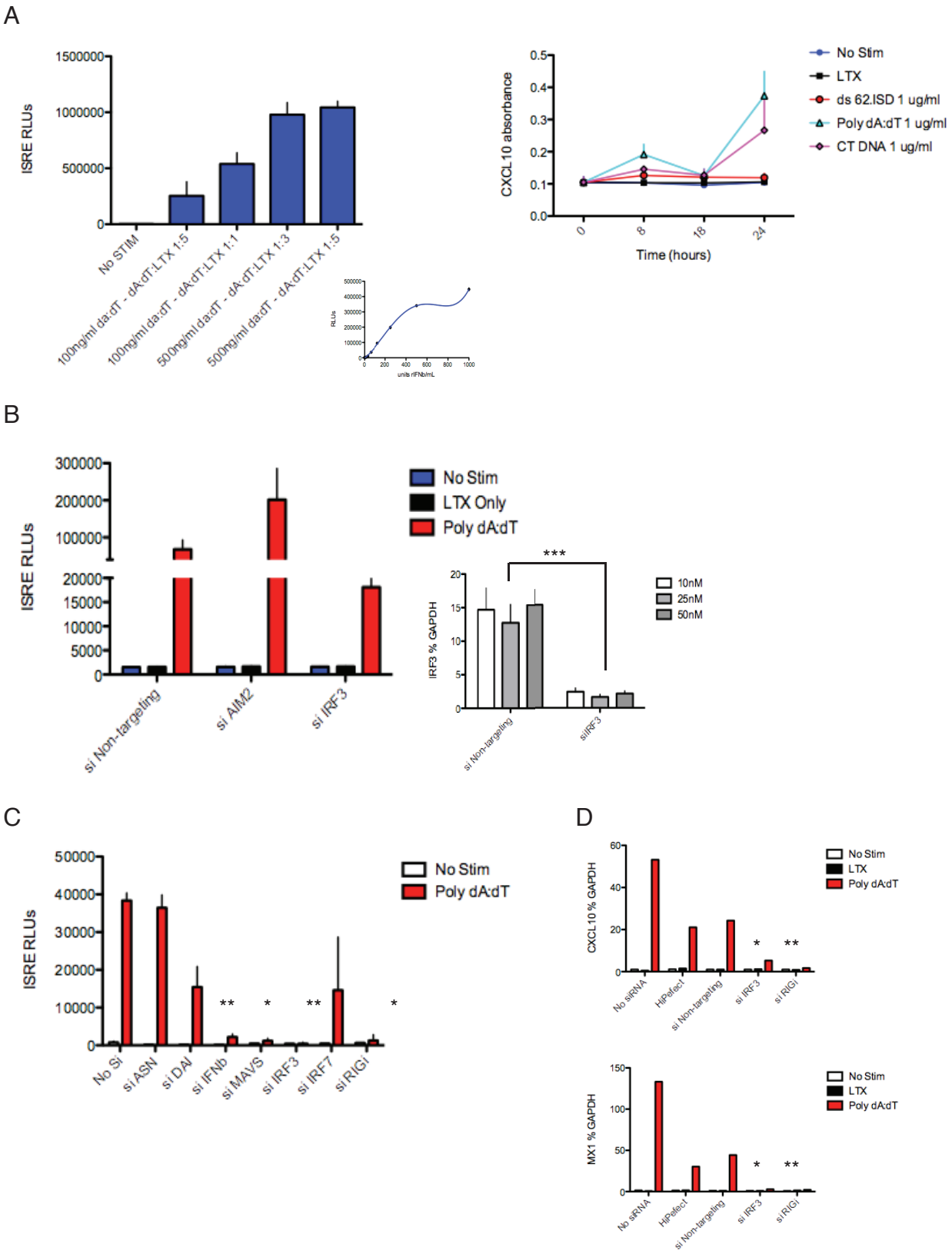


Figure 2.6 (continued)

absence of RIG-I, an RNA helicase that senses viral RNAs, and MAVS, a downstream effector of RIG-I that coordinates pathways leading to the activation of NF-κB, IRF3 and IRF7. Additionally, we have found that only the synthetic dsDNA ODN poly (dA:dT), but not any other dsDNA ligand, elicits a robust interferon signal. The data, however, were consistent with a finding that that both RIG-I and MAVS are essential for the cytosolic dsDNA-signaling pathway in a Huh-7 human hepatoma cell line that is naturally RIG-I deficient[124]. It was subsequently demonstrated that AT-rich regions in dsDNA become a template for RNA Polymerase III to generate 5' triphosphate RNA that signals through RIG-I[62, 125]. Human PBMCs produce IFN-α upon stimulation with Poly (dA:dT) and dsDNA ligands of varying lengths but only Poly (dA:dT) elicits an interferon response in 293T cells in agreement with out findings[61]. Furthermore, we learned that all human epithelial and fibroblast cells respond only to Poly (dA:dT) in a STING-dependent manner, whereas monocytes and macrophages respond to all dsDNA ligands (Y. Liu, Keystone Conference on Dendritic Cells, 2013). We therefore returned to murine models for the development of our screening system. We considered both primary cells and immortalized cell lines to assess their type I IFN response to transfected dsDNA ligands in a RIG-I-independent manner.

Evaluation of murine immortalized and primary cells

Bone marrow derived dendritic cells (BMDCs) from C57BL/6J mice were stimulated with a panel of DNA and RNA ligands for 0 to 8 hours (Figure 2.7A). Similar to our findings in B6 MEFs (Figure 2.2), BMDCs express *Cxcl10* following stimulation with DNA ligands ISD, shISD, and Poly A:T, increasing over time. CpG induced expression of *Cxcl10* occurred following four hours of stimulation consistent with the kinetics of TLR9 expression in dendritic cells[126]. Furthermore, IVT RNA and Poly I:C induced a potent *Cxcl10* response. Only IVT RNA induced *IFN-β* expression.

Figure 2.7.

Multiple murine cell types respond to dsDNA transfection by inducing Type-I IFNs in a time dependent manner. B6 Bone-marrow derived dendritic cells (A), BALB/c MEFs (B), RAW 264.7 macrophages (C) and $p53^{-/-}$ MEFs (D) were transfected with the indicated nucleic acid ligands for 0-8 hours. Lysates were collected and total RNA was isolated. Quantitative RT-PCR of Type-I interferon genes was performed from cDNAs prepared from RNA isolates. E) $p53^{-/-}$ MEFs were treated with the indicated siRNAs. Seventy-two hours following siRNA transfection cells were stimulated with ISD or IVT-RNA (3p-RNA) for 26 hours. Cxcl10 protein was measured by ELISA (*, p-value <0.05, paired t-test, aired t-test, siRNA control compared to ISD stimulated non-targeting control).

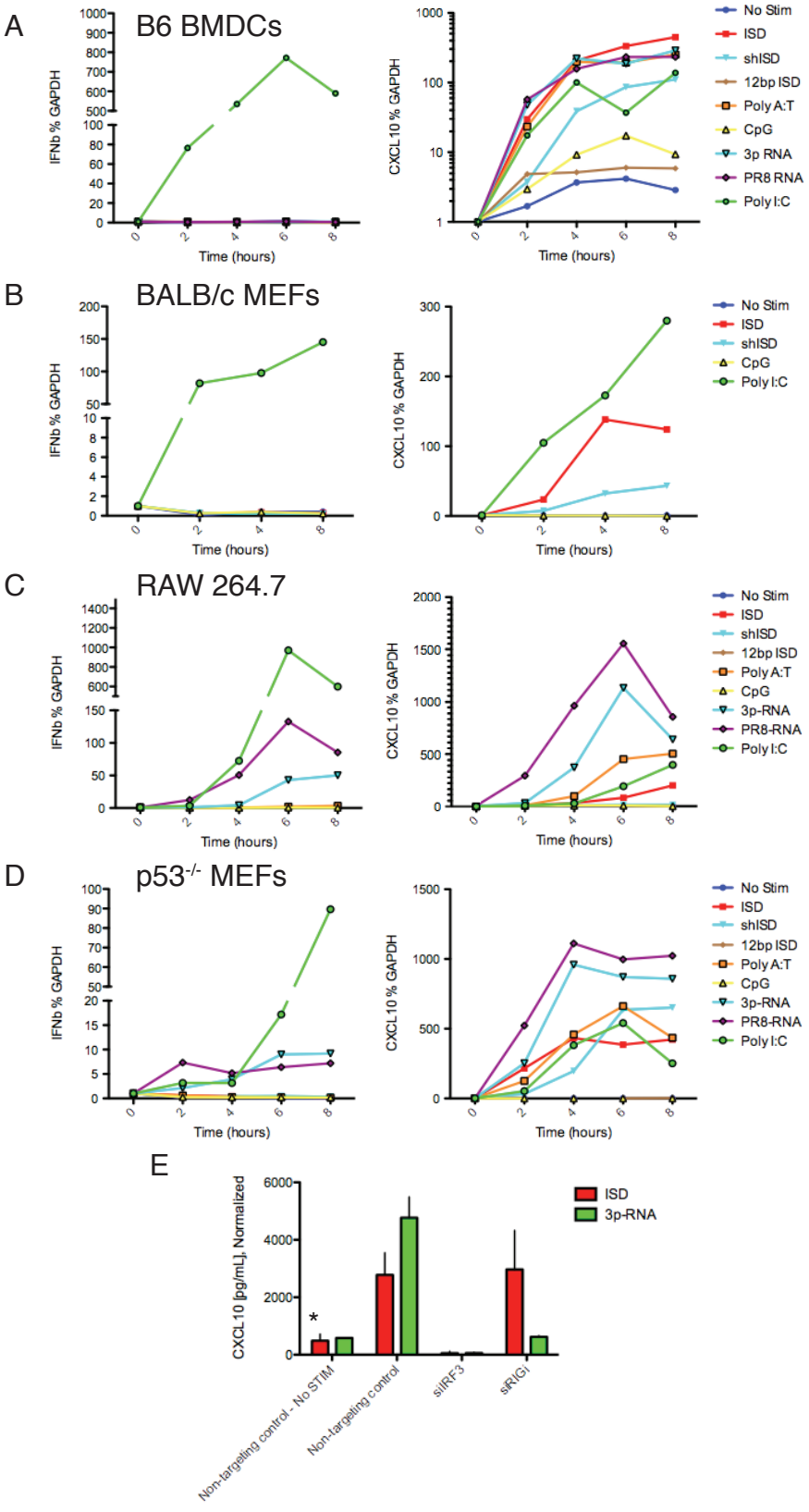


Figure 2.7 (continued)

Similar to B6 MEFs (Figure 2.2), BALB/c MEFs respond to ISD with increased *Cxcl10* expression (Figure 2.7B) with only Poly I:C inducing an *Ifn-β* response. The monocyte/macrophage-like cell line, RAW 264.7, responds strongly to both DNA and RNA ligands following transfection with increased expression of *Cxcl10* (Figure 2.7C). RNA ligands and not DNA ligands induce strong *IFN-β* responses .

We demonstrated that C57BL/6 and BALB/c MEFs, RAW264.7 macrophages, and 3T3 fibroblasts (data not shown), have intact ISD-sensing pathways, marked by the production of type I IFN following DNA stimulation. MEFs may be an useful primary cell to study innate responses, but they senesce after several passages and are thus not practical for large-scale screens (i.e. mass spectrometry experiments require as many as 10^8 cells per sample) or protocols that require multiple passages. While we have demonstrated robust responses in monocyte/macrophage-like RAW 264.7 cells they are difficult to passage and are not easily transfectable with siRNAs. Similarly, we have demonstrated robust Cxcl10 responses BMDCs. It is, however, difficult to produce pure populations of BMDCs in large numbers with GM-CSF stimulation alone and without sorting. Furthermore, BMDCs are not amenable to transfection of siRNAs with lipid-based transfection reagents, typically requiring electroporation to mediate siRNA delivery at the cost of high cell death.

As an alternative to wild-type MEFs, we tested p53-deficient MEFs ($p53^{-/-}$) (from D. Sabatini and D. Kwiatkowski, Whitehead Institute, Cambridge, MA), which are immortal and grow rapidly in culture (Figure 2.7D). To assess the ISD-sensing pathway response, $p53^{-/-}$ MEFs were stimulated with a panel of DNA and RNA ligands for 0 to 8 hours. We demonstrated that *Cxcl10* expression increased over time following stimulation with DNA ligands ISD, shISD, and Poly A:T. Additionally, IVT RNA and Poly I:C induced a potent *Cxcl10* response but only Poly I:C induced substantial IFN-β

expression. We selected p53$^{-/-}$ MEFs as our model cell for the detection of ISD-sensing pathway responses. Many of the known DNA sensors are restricted to macrophages and APCs, and although it has been reported that any nucleated cell can produce interferon in response to pathogenic stimuli, there is no strong evidence pointing to a specific sensor in the ISD-pathway in MEFs[127]. The ISD-pathway is intact in p53$^{-/-}$ MEFs, demonstrated by robust expression of Cxcl10 and other cytokines in response to multiple dsDNA ligands. Furthermore, p53$^{-/-}$ MEFs are easily passaged and, as is described in detail below, are amenable to both transfection and infection with siRNAs and shRNAs, respectively.

2.4 – *Development of a quantitative assay to detect type I interferon responses of nucleic acids*

Following the identification of a suitable cell line, we pursued the development of a robust quantitative assay to detect type I IFN responses to nucleic acids for large-scale genome-wide pooled shRNA or arrayed siRNA screens. In addition to reproducibility, specificity and sensitivity, we also factored in assay costs into the decision process. Furthermore, we considered the utility of the assay as a future resource (for example, gain-of-function (cDNA) screens).

We focused our search for reporters of the ISD-sensing pathway on the expression of type I interferon genes, including cytokines, and chemokines as a proxy of the ISD response. These inducible effector molecules are hallmarks of the antiviral response and are striking characteristic of the IFN signature described in patients with Systemic Lupus Erythematosus (SLE) [40, 42-44, 128]. To this end, we pursued a number of type I IFN reporters including a CD-tagging method, GFP-tagged Cxcl10 and Ifn-β reporters, a Luciferase interferon stimulated response element (ISRE) reporter,

intracellular staining flow assays (IC Flow) and Cxcl10 and Ifn-β ELISAs. We also developed a highly accurate dual-reporter qPCR-based system to detect the expression of IFN genes and a control gene. Additionally, we considered the contribution of cell surface markers implicated in prior studies in the immune response to nucleic acids, including MHC class I and II proteins and costimulatory markers, B7.1 (CD80)[117] and PD-L1[44].

Type I IFN-GFP reporters

To complement the proposed arrayed screens, we considered the use of genome-wide pooled shRNA screens. In this approach, a genome-wide lentiviral shRNA library is infected in a single pool into p53 MEFs, containing fluorescent type I IFN pathway reporters. Changes in these reporters are used as a proxy for genes involved in the DNA sensing process; cells exhibiting a phenotype of interest caused by specific shRNAs are isolated by FACS. The shRNAs in the isolated population can then identified using custom microarrays that display cognate probes against the entire library of shRNAs. With this method in mind, we generated $p53^{-/-}$ MEFs with stable integration of either Ifn-β or Cxcl10 promoter-GFP constructs that typically respond with 10-100-fold induction of GFP. This high response threshold would be appropriate for a pooled screens where separation of high and low GFP expressing cells is required by FACS[22, 129]. We infected p53-/- MEFs with Ifn-β and Cxcl10 promoter-GFP constructs and generated single-cell subclones identified following stimulation with recombinant Ifn-β. A summary of their responses to dsDNA and dsRNA can be found in Table 2.1. An additional round of subcloning did not improve the fold change response of transfected dsDNA compared to unstimulated cells (Figure 2.8A). As a pilot to demonstrate the potential utility of Cxcl10 GFP-reporter MEFs, we either transfected reporter subclones with positive and negative control siRNAs or infected them with mock-pools of multiple

Table 2.1

Developing a cell-autonomous assay for selecting bioactive RNAi treated cells. Generation of Type I IFN-GFP reporters. p53$^{-/-}$ MEF subclones stably expressing Cxcl10 and IFNb promoter reporters were stimulated with 5ug/mL of transfected dsDNA or dsRNA for 24 hours. GFP induction was measured by FACS.

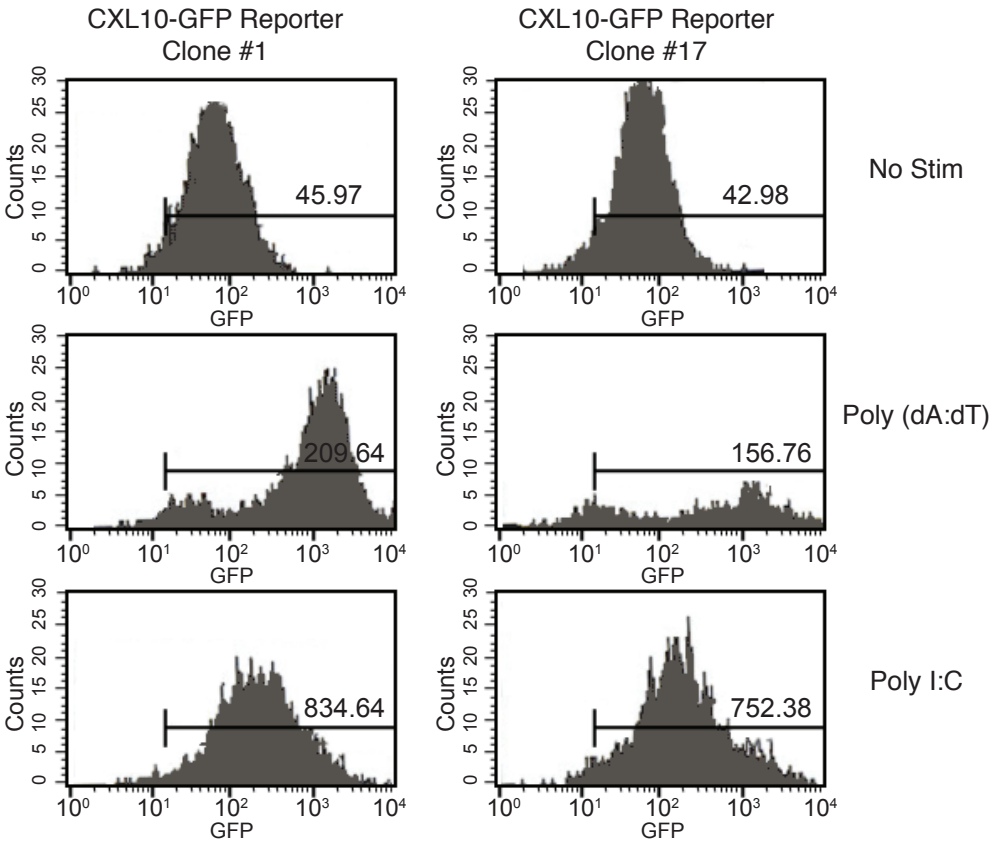
	MEFS/IP10-GFP subclones							
	GFP%			Mean			Mean folds	
	No Stim	Poly I:C	Poly A:T	No Stim	Poly I:C	Poly A:T	Poly I:C	Poly A:T
#1	95.34	98.09	98.83	63.61	1269.48	322.64	19.96	5.07
#17	95.95	88.63	97.78	60.67	860.25	287.26	14.18	4.73

	MEFs/IFNb-GFP subclones							
	GFP%			Mean			Mean folds	
	No Stim	Poly I:C	Poly A:T	No Stim	Poly I:C	Poly A:T	Poly I:C	Poly A:T
#13	1.68	72.42	14.96	7.50	93.24	12.30	12.43	1.64
#22	32.86	93.32	73.63	14.19	120.72	23.69	8.51	1.67

Figure 2.8

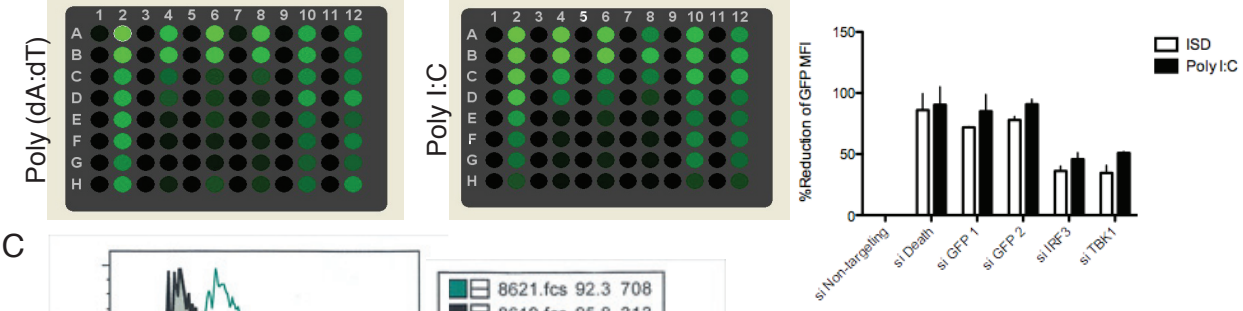
Developing a cell-autonomous assay for selecting bioactive RNAi treated cells. Generation of Type I IFN-GFP reporters. A) $p53^{-/-}$ MEF subclones stably expressing Cxcl10 promoter GFP reporters were stimulated with 5ug/mL of transfected dsDNA or dsRNA for 24 hours. GFP induction was measured by FACS. B) Increasing concentrations of control siRNA transfected Cxcl10 reporter subclones were stimulated with Poly (dA:dT) (left) or Poly I:C (right) according the indicated layout. Fluorescence was measured by microplate cytometry. Fold reduction (right panel) was measured by comparing mean florescence intensity between conditions. C) $p53^{-/-}$-CXCL10 reporter MEFs were infected was multiple shRNA control hairpins, puromycin selected and stimulated with dsDNA for 18-hours. GFP expression was measured by FACS.

A



B

A	No si	No si	No si	No si	No si	No si	No si	No si	No si	No si	No si	No si	
B	HiPerfect	HiPerfect	HiPerfect	HiPerfect	HiPerfect	HiPerfect	HiPerfect	HiPerfect	HiPerfect	HiPerfect	HiPerfect	HiPerfect	
C	AS Neg	AS Neg	AS Death	AS Death	si GFP R	si GFP R	si GFP D	si GFP D	si IRF3	si IRF3	si TBK1	si TBK1	10nM
D	AS Neg	AS Neg	AS Death	AS Death	si GFP R	si GFP R	si GFP D	si GFP D	si IRF3	si IRF3	si TBK1	si TBK1	10nM
E	AS Neg	AS Neg	AS Death	AS Death	si GFP R	si GFP R	si GFP D	si GFP D	si IRF3	si IRF3	si TBK1	si TBK1	25nM
F	AS Neg	AS Neg	AS Death	AS Death	si GFP R	si GFP R	si GFP D	si GFP D	si IRF3	si IRF3	si TBK1	si TBK1	25nM
G	AS Neg	AS Neg	AS Death	AS Death	si GFP R	si GFP R	si GFP D	si GFP D	si IRF3	si IRF3	si TBK1	si TBK1	50nM
H	AS Neg	AS Neg	AS Death	AS Death	si GFP R	si GFP R	si GFP D	si GFP D	si IRF3	si IRF3	si TBK1	si TBK1	50nM



C

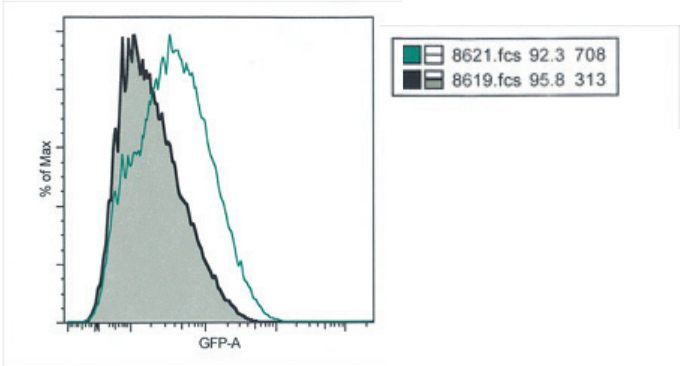


Figure 2.8 (continued)

non-targeting shRNAs. We measured the RNAi effect following ISD stimulation by automated microscopy or FACS, respectively (Figure 2.8B-C). Regardless of the RNAi method, the results failed to demonstrate significant fold change of stimulated cells relative to background.

With ISD-mediated IFN-GFP-fold induction no higher than 5-20-fold in the p53-/- reporter clones, we pursued a distinct approach that utilizes retroviral integration of a fluorescent protein at the site of endogenous genes in their normal chromosomal locations. This retroviral integration at the site of the native promoter is expected to show more dramatic induction than artificial promoter constructs.

CD Tagging

Similar to gene traps, the CD Tagging system drives the random insertion of YFP into genes via retroviral integration[130]. The presence of splice acceptor and donor sites surrounding YFP incorporates the YFP tag as part of the synthesized gene product (Figure 2.9A). We pursued a method to select for random YFP insertion at the site of ISD-pathway responsive genes. We developed the following strategy: Following infection of the pBabe-tagging vector, $p53^{-/-}$ MEFs are sorted by flow cytometry. Cells that spontaneously fluoresce are discarded. The remaining cells are cultured and stimulated with ISD. On the second round of sorting, eYFP positive cells are sorted at one cell per well into 384-well plates and expanded into clones. Cells are passaged from 384-well plates into ordinary tissue culture 96-well plates and one optical 96-well plate. The optical 96-well plate is used to image the proteins tagged following a second round of ISD stimulation. The remaining plates are used for 3' RACE to detect the gene of insertion or freezing (Figure 2.9B).

For compatibility with the puromycin-resistant lentiviral shRNA library, hygromycin and neomycin resistant CD-tagging vectors were developed (Figure 2.9C).

Figure 2.9

Developing a cell-autonomous assay for selecting bioactive RNAi treated cells. Generation of a CD-tagged protein library. Adapted from Sigal A et al. A) The CD-tagging vector pBabeAE. Splice acceptor (SA) and splice donor (SD) flanked fluorophore sequence (FL seq), with no promoter, no start codon or polyA signal, is inserted into the genome by MoMLV. Flagged mRNA translates to an internally labeled protein, with the fluorophore protein tag (FL tag) usually near the N terminus. B) Flowchart of the library generation procedure. Following infection of the pBabe tagging vector, $p53^{-/-}$ MEFs are sorted by flow cytometry. Cells that spontaneously fluoresce are discarded. The remaining cells are cultured and stimulated with dsDNA. On the second round of sorting, eYFP positive cells are sorted at one cell per well into 384-well plates and expanded into clones. Cells are passaged from the 384-well plate into two ordinary tissue culture 96-well plates and one optical 96-well plate. The optical 96-well plate was used to image the proteins tagged following a second round of dsDNA stimulation. The remaining plates are used for 3' RACE to detect the gene of insertion or freezing. C) To use the CD tagging vectors as potential screening tool for lentiviral shRNA screens, PURO resistance was replaced with the indicated resistance markers as demonstrated by gel electrophoresis. D) $p53^{-/-}$ MEFs were infected overnight with CD-TAG Null (no-resistance marker), Neomycin, or Hygromycin constructs, along with an FUGW control. Cells were sorted for eYFP positive cells.

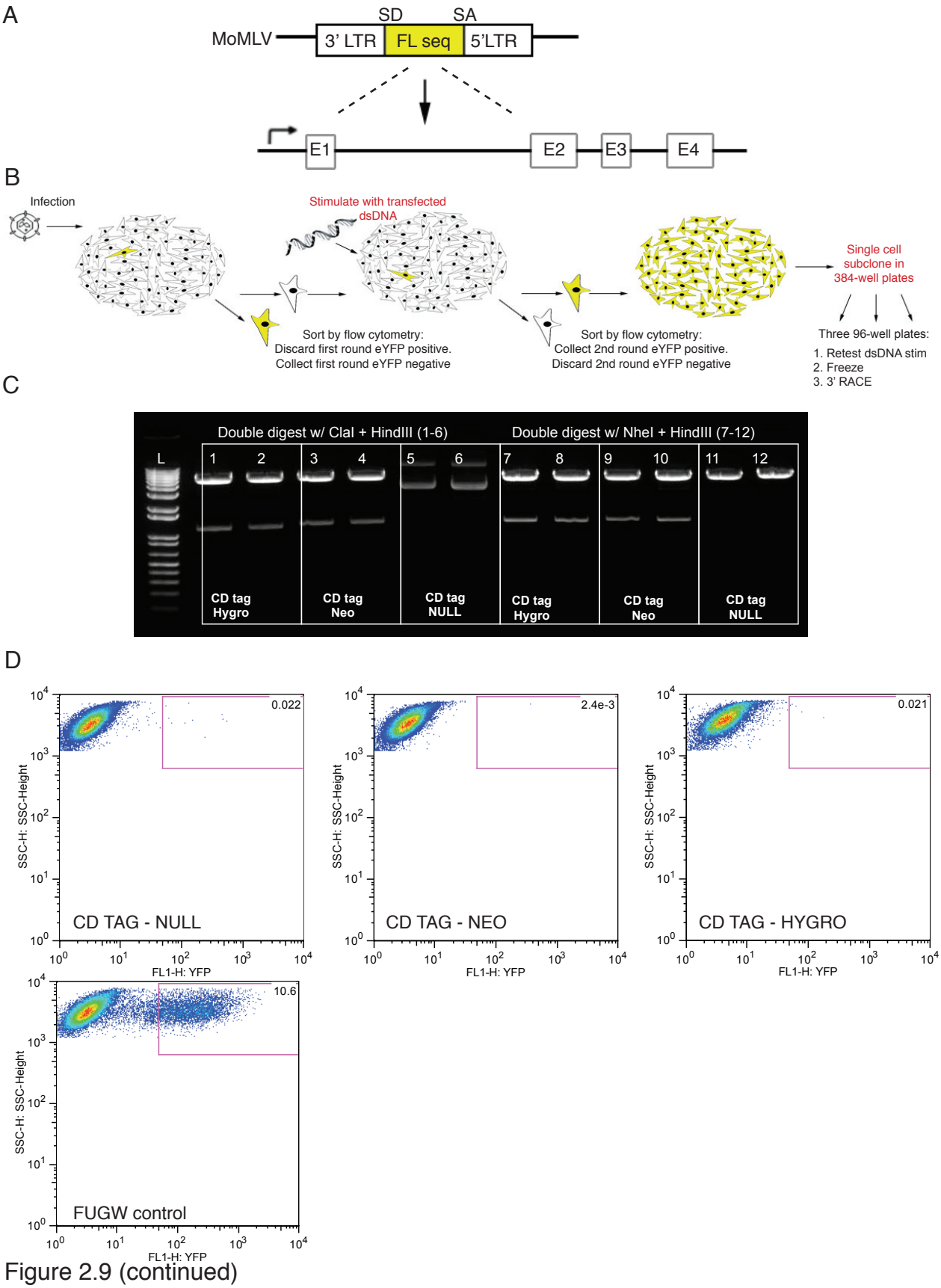


Figure 2.9 (continued)

Multiple attempts to generate an YFP-tagged type I interferon using the CD tagging approach were unsuccessful (Figure 2.9D) though we feel this may be a useful strategy for future studies.

ISD-induced cell-surface marker expression

Double-stranded nucleic acids induce expression of multiple genes related to antigen processing and presentation including B7.1 (CD80)[117], MHC Class II, CD40 and B7.2 (CD86)[44]. We therefore assessed the utility of cell-surface markers as a screening tool to interrogate the ISD-sensing pathway.

In additions to multiple IFN-gamma-inducing TLR-ligands, we transfected dsDNA into the human RS4-11 B-cell line and murine RAW264.7 cells for 24-hours and measured cell surface marker expression of CD40 and CD80 by FACS, respectively (Figure 2.10). While there were four-fold increases of CD40 positive cells in response to Poly (dA:dT) in human RS4-11 cells, we detected only a modest increase of CD80 positive cells following dsDNA stimulation in RAW264.7 monocyte/macrophages. Furthermore, we demonstrated that RS4-11 B-cells are not amenable to transfection with siRNAs, and are difficult to passage in the context of a large-scale genome-wide shRNA screen (data not shown).

Detection of type I IFN by quantitative RT-PCR

Quantitative real-time PCR is a sensitive and flexible tool for assessing gene expression. As a screening tool, quantitative RT-PCR is compatible with 384-well cell culture format and high throughput robotic workflow, including liquid handling and automated quantitative RT-PCR[131-133]. To this end, we developed a Taq-man based dual-reporter quantitative PCR system that simultaneously reports on both *Cxcl10* and a control gene (*Gapdh*). The dual reporter system simultaneously reports two unique hydrolysis probes that are labeled at the 5' end with Light-Cycler 555 Yellow or

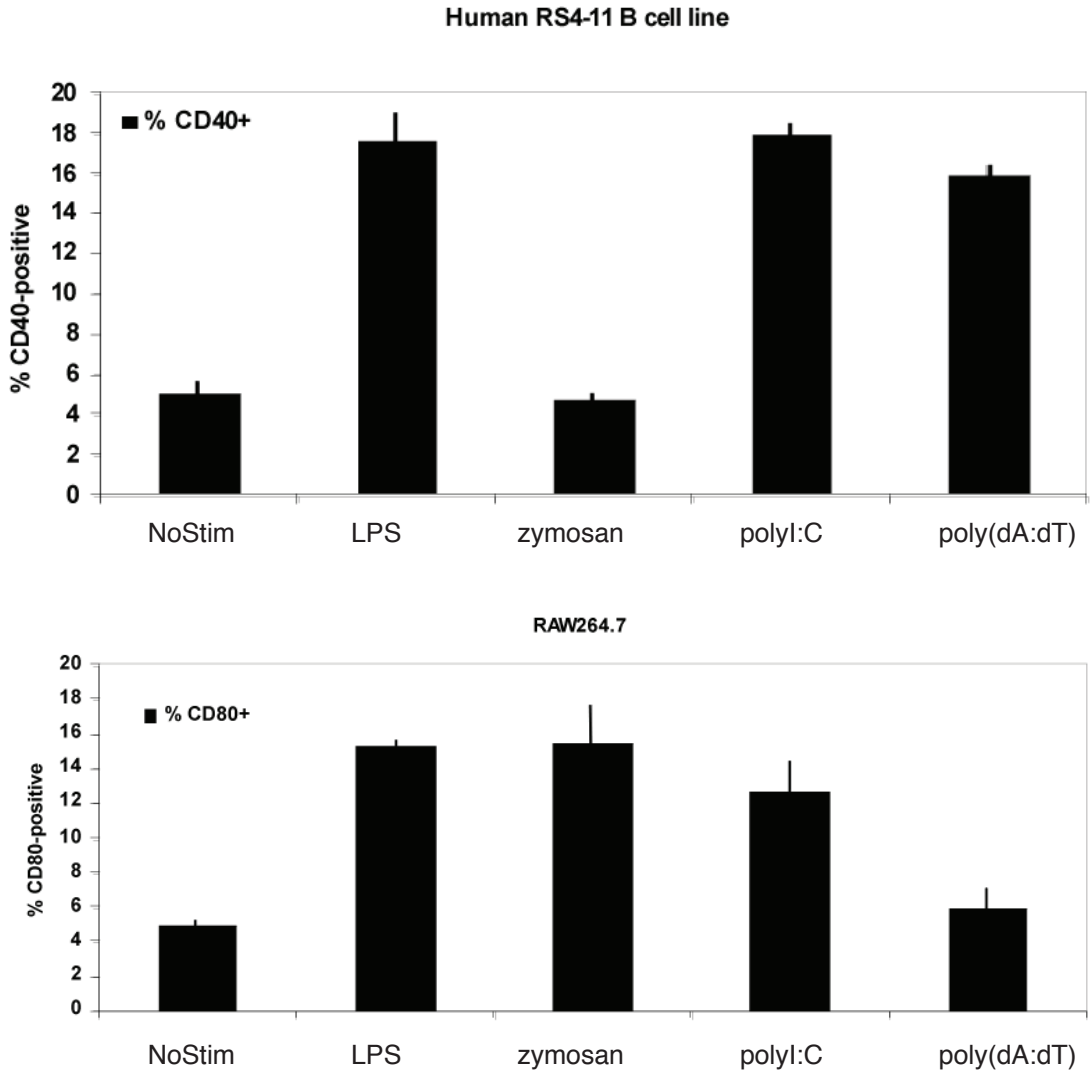


Figure 2.10

Developing a cell-autonomous assay for selecting bioactive RNAi treated cells. Assessment of cell-surface markers following TLR stimulation in human and murine cells. Human RS4-11 B-cells and murine RAW264.7 monocyte/macrophage cells were stimulated with the indicated ligands for 24 hours. CD40 and CD80 cell surface expression was measured by FACS.

fluorescein (FAM) or and at the 3' end with a dark quencher dye for the probes. Following ISD stimulation, we carried out cDNA synthesis directly on cell lysates from 384-well plates, eliminating the RNA purification step, allowing us to proceed directly to perform RT-PCR. We compared a two-PCR SYBR-green primer-based system to the dual-reporter system following stimulation of MEFs with DNA and RNA ligands (Figure 2.11). Dual reporter detection of *Cxcl10* expression demonstrated greater than 100-fold increase over unstimulated cells (p-value <0.001, unpaired t-test). Though the cost of the qPCR reagents were reasonable in the context of a genome-wide screen, the preparation of cDNA from whole cell lysates exceeded the screening budget. We explored further Taq-man-based multiplexing options, reporting up to four genes per well using lysate-direct cDNA synthesis. Though the large scale of our screen put the Taq-man based method out of reach, a multiplex system targeting the expression of a few key genes could be useful in the context of a smaller screen.

Detection of type I IFN by ISRE Luciferase reporters

To report on inflammatory responses in primary human bronchial epithelial cells, we developed a Luciferase interferon stimulated response element (ISRE) reporter assay using 293T cells as a surrogate cell for transferred supernatants (Figure 2.6)[134]. Type I interferons mediate signaling through STAT1 and STAT2 components of the JAK/STAT-signal transduction pathways. The STAT1/STAT2-responsive Luciferase construct encodes the firefly Luciferase reporter gene under the control of a mCMV promoter and tandem repeats of the ISRE. The demonstrated efficacy of the ISRE reporter in human cells persuaded us to develop an equivalent reporter for murine cells. We infected $p53^{-/-}$ MEFs with a lentivirus containing ISRE-Luciferase and, following selection, we stimulated the resulting polyclonal ISRE reporters either directly with ISD or with supernatants transferred from ISD-stimulated MEFs (Figure 2.12). ISRE Luciferase

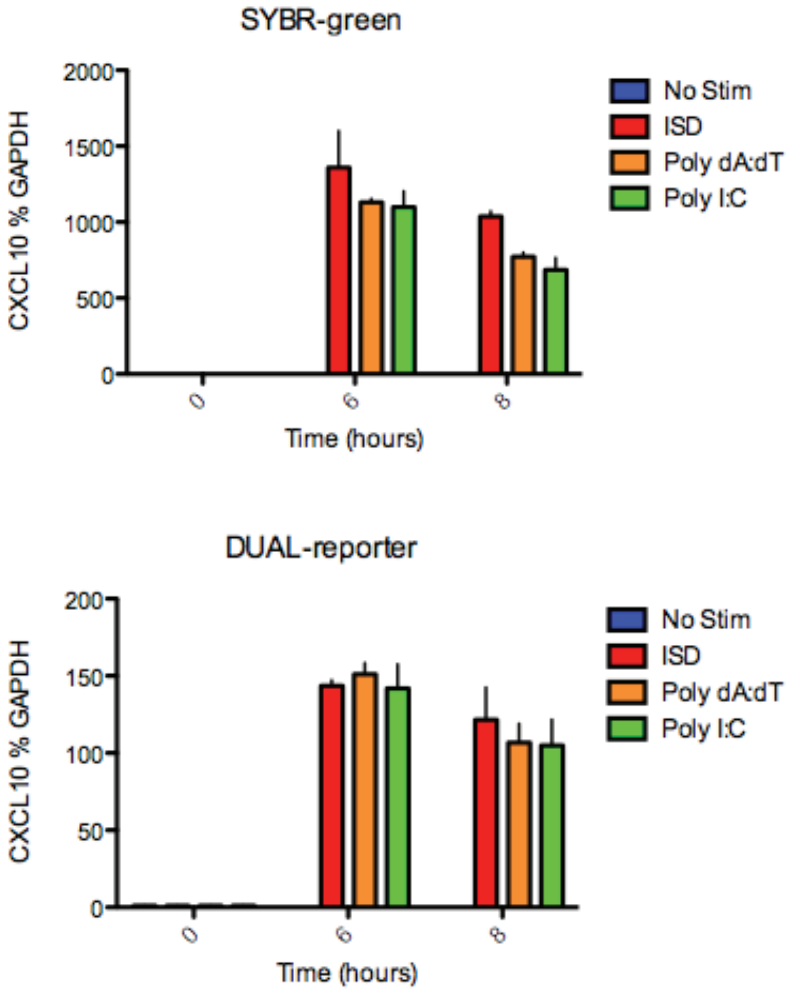


Figure 2.11

Developing a cell-autonomous assay for selecting bioactive RNAi treated cells. Quantitative-RT-PCR dual-reporter. Comparison of SYBR-green based two-step RT-qPCR with TaqMan based hydrolysis probes that are labeled at the 5' end with Light-Cycler 555 Yellow or fluorescein (FAM) or and at the 3' end with a dark quencher dye for the probes. $p53^{-/-}$ MEFs were stimulated with the indicated DNA or RNA ligands for 0, 6 or 8 hours. Lysates were prepared using one-step Cells-to-CT and directly analyzed by RT-PCR.

Figure 2.12

Developing a cell-autonomous assay for selecting bioactive RNAi treated cells. Luciferase interferon stimulated response element (ISRE) reporter development. A) p53-/- MEFs were infected with a lentivirus containing ISRE-luciferase, Puromycin-selected and then directly stimulated with 1-2 ug/mL (Lo, Hi, respectively) of dsDNA ligands for 24 hours or stimulated with supernatants for identically treated p53-/- MEFs. ISRE relative luminescent units were measured and compared to mouse IFNb stimulated cells (inset). B) Supernatant transfer induced a titratable ISRE response. C) Stably infected cells were cloned by limiting dilution and tested for responsiveness to mouse IFNβ. D) ISRE reporter cells were stimulated with the indicated amounts of Adenovirus and Sendai Virus for 24 hours. E) siRNA transfected ISRE reporter MEFs were stimulated directly with dsDNA and IVT-RNA (3p-RNA) 72-hours following siRNA treatment. ISRE RLUs were measured 24-hours following stimulation. F) ISD complexes remain active following supernatant transfer. Supernatants were collected following three-hours or 24-hours following stimulation and response was detected by Cxcl10 ELISA or ISRE reporter cells in parallel. ISRE reporters were stimulated with either fresh or frozen ISD complexes.

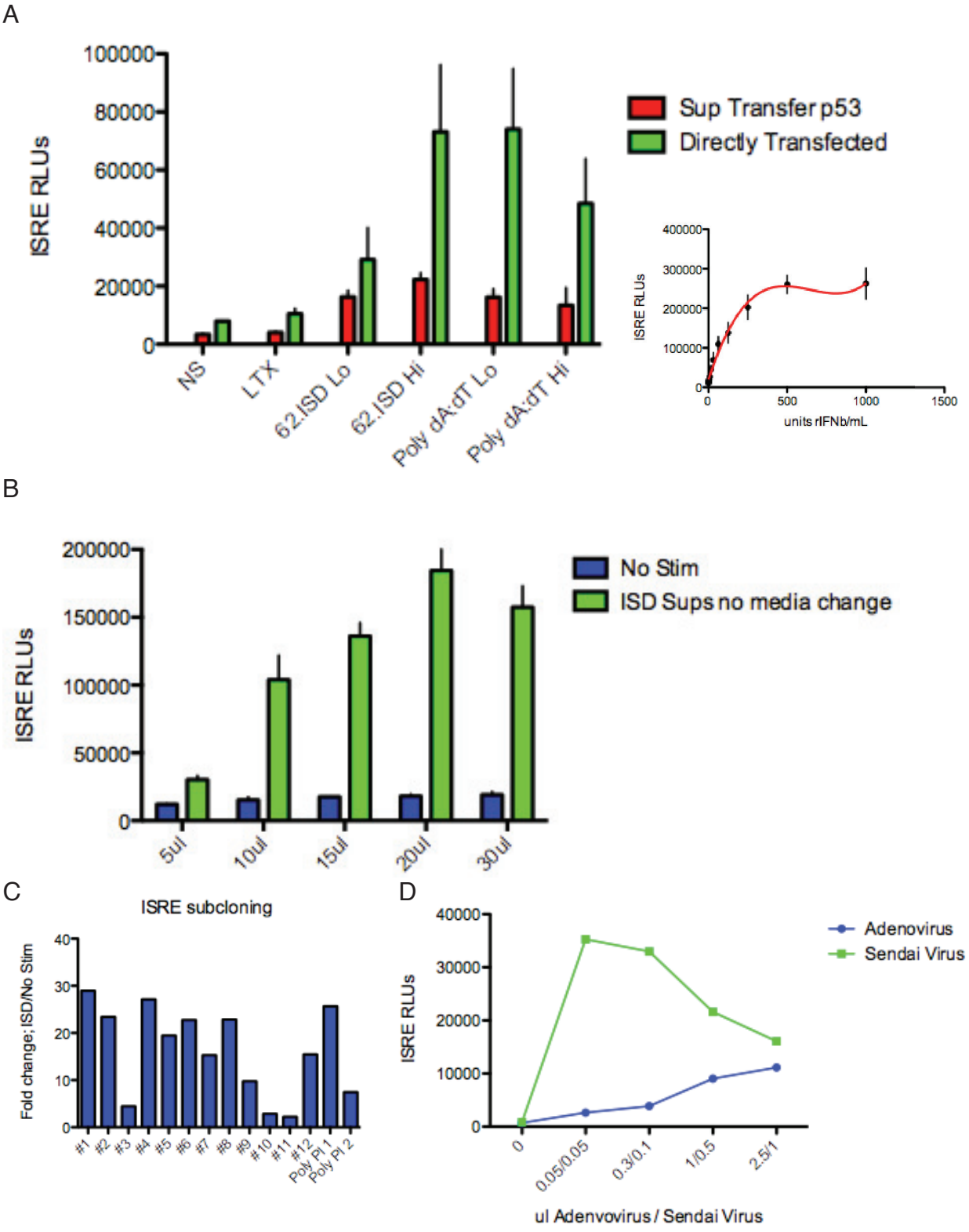


Figure 2.12 (continued)

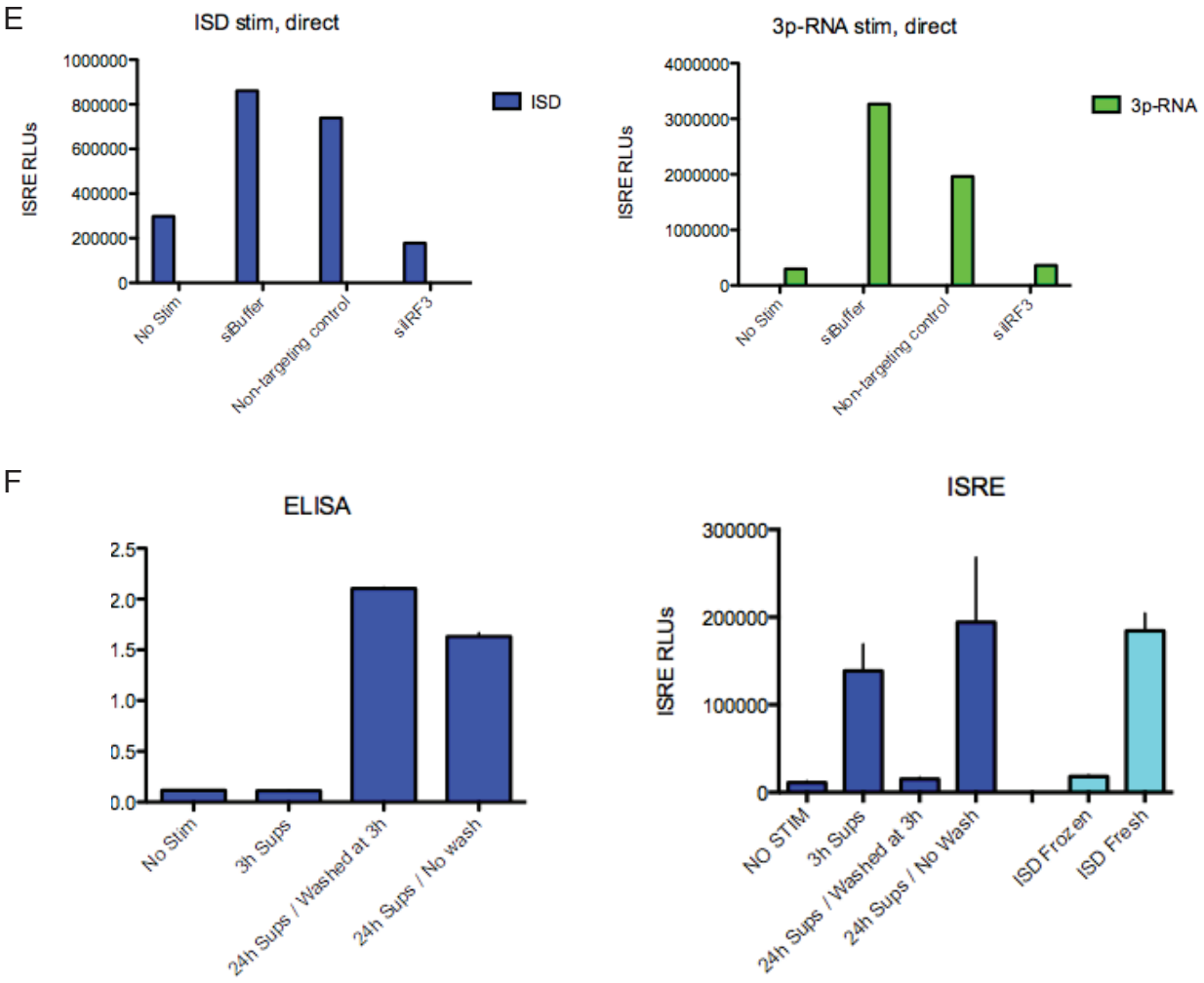


Figure 2.12 (continued)

luminescence indicated a seven and nine-fold increase relative to unstimulated cells, for transferred supernatant and direct transfection, respectively. Responses correlated directly to supernatant volume (Figure 2.12A-B). Next, we developed monoclonal ISRE reporter MEFs. Following selection with puromycin, stably infected cells were cloned by limiting dilution and tested for responsiveness to mouse Ifn-β. A clone with high signal to background ratio was selected and found to be sensitive to low levels of Ifn-β (<1U/mL) with a >50× dynamic range (Figure 2.12C). Furthermore, ISRE reporter clones were sensitive to viral infection following stimulation with either Adenovirus or Sendai virus (negative sense, single-stranded RNA virus) (Figure 2.12D).

To distinguish the effects of genetic perturbation on the ISD-sensing pathway from RNAi-associated toxicity, it is crucial to detect type I IFN responses on a per-cell basis. Cell viability can be measured directly in the context of the ISRE reporter with a dual-reporter Luciferase/Renilla system or indirectly by transferring supernatants to ISRE reporter cells and measuring viability in the siRNA treated cells by luminescent detection of ATP (CellTiter-Glo). We pursued the supernatant transfer model to fit within our screening budget. However, complexes from supernatants transferred from $p53^{-/-}$ MEFs continued to stimulate ISRE reporter cells (Figure 2.12E). Supernatants collected following three-hours of stimulation fail to induce a response in directly stimulated cells as detected by Cxcl10 ELISA but induce a potent response in ISRE reporter cells. Furthermore, supernatants from cells washed three hours following stimulation and collected 24-hours later induced Cxcl10 expression to levels similar to un-washed supernatants as detected by ELISA. Only supernatants collected at 24-hours induced a similar response in ISRE reporter cells, further supporting evidence that complexes remain active following transfer to reporter cells. Lastly, we determined that ISRE

detection in 384-well plates was not feasible due to crosstalk with neighboring wells (data not shown).

Detection of type I IFN by ELISA

Lastly, we considered enzyme-linked immunoblot staining assays as a screening tool for the detection of ISD-sensing pathway responses. CXCL10 protein detection by ELISA in ISD-stimulated primary human cells led to robust expression. We reasoned that ISD triggered responses would be similar in mice. Stimulation of p53-/- MEFs with ISD or IVT-RNA ligands induced Cxcl10 protein at levels greater than 30-fold compared to unstimulated cells (Figure 2.13A). In addition to Cxcl10, we assessed for Ifn-β, Ifn-α, and Il1b (data not shown).

We next considered the utility of an ELISA-based detection tool in the context of an siRNA-based loss-of-function screen of the ISD-sensing pathway. In MEFs treated with a non-targeting control siRNA we detected normal responses to nucleic acids. MEFs treated with an *Irf3*-directed siRNA demonstrated strongly reduced responses (more than 40-fold) to both dsDNA and IVT-RNA (p-value <0.001, Student's t-test). Furthermore, we demonstrated abrogated Cxcl10 expression in IVT-RNA stimulated MEFs with siRNA-directed knockdown of *Rig-i* and *Mavs* as well as in ISD-stimulated cells treated with *Tbk1* siRNA.

To further optimize the Cxcl10 ELISA, we considered both length of ISD complex stimulation time and volume of supernatant. We demonstrated that the detection of Cxcl10 protein expression was dependent both on the amount of time DNA complexes are present on the cell and the volume of supernatant added to the ELISA (Figure 2.13B). Additionally, we showed that cell viability was higher in wells with no media change, in spite of the continued presence of ISD-transfection complexes. Cells stimulated without additional washing had nearly identical viability as unstimulated cells

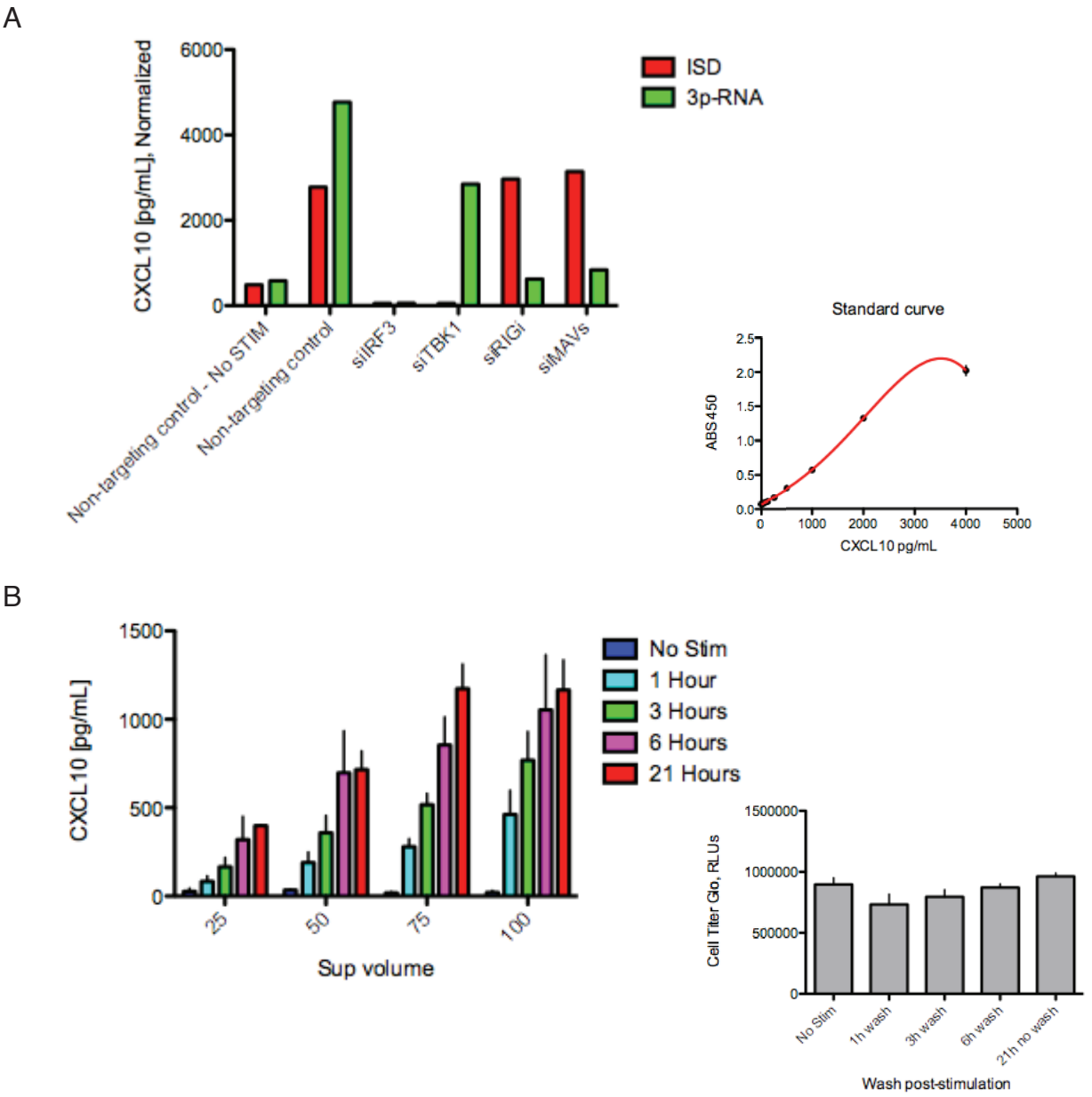


Figure 2.13

Developing a cell-autonomous assay for selecting bioactive RNAi treated cells. ELISA-based protein detection of Cxcl10. A) p53$^{-/-}$ MEFs treated with the indicated siRNAs were stimulated with 1.0ug/mL ISD and 0.1ug/mL 3p-RNA for 26 hours. Cxcl10 protein was measured by ELISA. Data analysis was performed by converting the Abs 450 to the predicted Cxcl10 value in pg/mL based on a 4n-polynomial standard curve (inset). B) ISD complexes were incubated for the indicated time and supernatant volumes were diluted in 25ul increments. Sixty-five ul of CellTiter-Glo was added to each ISD-stimulated well for 10 minutes prior to analysis.

as detected by CellTiter-Glo. To generate a normalized value, we divided pg/mL of Cxcl10 by the number of cells per well, reported in relative luminescence units (RLUs). By normalizing CXCL10 protein expression to cell number, the ELISA/CellTiter-Glo combination is an effective and scalable screening tool with a high signal to background ratio that is sensitive to low levels of stimulation yet has large dynamic range. In the context of our screen, we developed a more precise mode of normalization. To more accurately predict the effect of siRNA on the ISD-pathway we seeded cells in a dilution curve of increasing cell numbers. Subsequent stimulation with ISD provides expected Cxcl10 expression for a given cell number. In conjunction with cell viability measurements via CellTiter-Glo, we demonstrated that Cxc10 ELISA effectively captures ISD-sensing pathway responses. Furthermore, detection of proteins also provides a level of detection that mRNA expression cannot provide; nominally that mRNA does not represent the final gene product.

2.5 – *Genetic perturbation using RNA interference*

With the identification of a model cell and type I Interferon-detection method, we considered two screening strategies; pooled genome-wide lentiviral shRNA and candidate-based siRNA screens (Figure 2.14).

Genome-wide lentiviral shRNA-based screens

Initial efforts focused on lentiviral-based shRNA pooled screens, developed in-house at the Broad Institute[135, 136]. To enable arrayed RNAi screening in a broad range of cells, our laboratory, in collaboration with several others, founded the Broad Institute RNAi Consortium (TRC) in 2004. The goal of TRC was to generate a genome-wide RNAi library in a high-titer lentiviral vector that can stably infect most cell types and effectively knockdown gene expression. The library consists of >180,000 sequence-verified

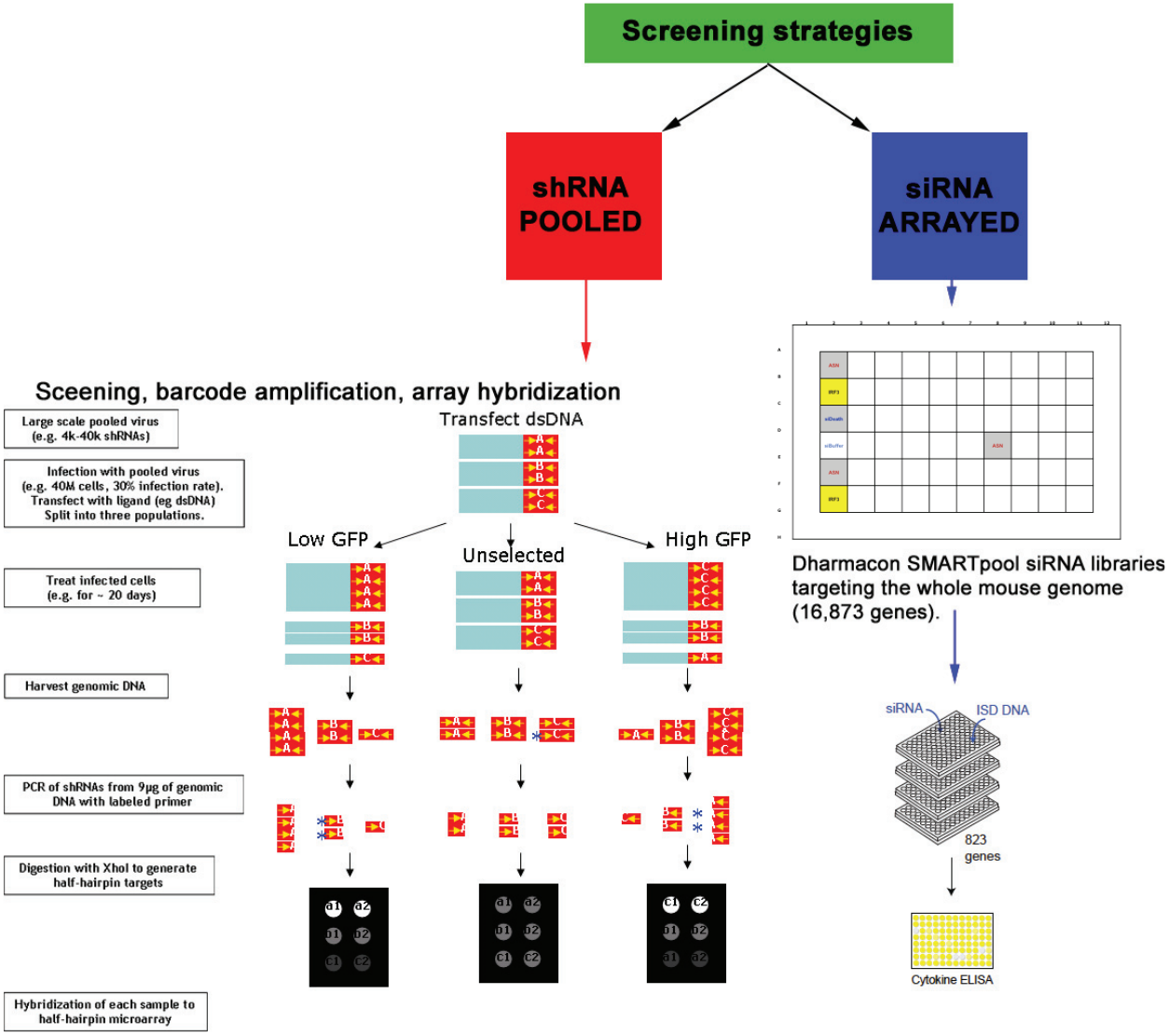


Figure 2.14

Genetic perturbation using RNAi interference. Schematic of shRNA and siRNA screen workflow. Left, pooled shRNA screen: cells are infected with a pool of 15,000 shRNAs and separated into experimental and control populations. ShRNAs are amplified from each population, cut to avoid hairpin structures, and fluoresceinated. Custom microarrays are used to detect the enrichment of specific shRNAs in the experimental vs. control population. Right, candidate siRNA screen: cells are transfected with siRNA pools targeting 1003 genes and stimulated with ISD 72-hours following treatment. Cxcl10 response was detected by ELISA.

lentiviral constructs targeting most human and mouse gene (TRC public portal: www.broad.mit.edu/genome_bio/trc/). There are at least five independent shRNAs per gene and producing high-titer virus that can infect a wide range of primary and immortalized cell lines. Knockdown efficiency of the entire shRNA library will be available to the public shortly.

Using a streamlined lentiviral vector, we can generate high-titer lentivirus particles for each library construct in a high-throughput fashion, and thus perform arrayed screens in 96 or 384-well plates, in which each well contains many virus particles targeting a single gene or pooled screens where each cell is infected with a single shRNA-containing virus. Using quantitative RT-PCR, we have demonstrated that 90% of genes have ≥1 shRNA that knocks down mRNA levels by ≥70% in A549 cells[135].

Selecting a subset of the human lentiviral RNAi library targeting all human kinases and phosphatases, we performed a screen using automated fluorescence microscopy to identify genes that regulate mitosis as determined by DNA content and phospho-H3 staining. We were able to recover many of the known cell cycle regulators (such as *CDC2*, *AURKB* and *PLK1*) and identify new ones, thereby demonstrating the utility of the library in identifying critical genes in a cellular process.

A major concern with the use of RNAi in mammalian cells is off-target effect. To mitigate this problem, shRNAs were designed to have several mismatches to all known human or mouse cDNAs. As this does not eliminate the possibility of off-target effects with shorter stretches of identity, we also produced five distinct shRNAs targeting each gene and routinely require that screen hits have ≥2 constructs inducing the same phenotype. Because distinct shRNAs are expected to have non-overlapping off-target effects, this criterion is expected to filter out most off-target effects.

Pooled screens are carried out by infecting a large population of cells with a single pool of lentiviral particles targeting thousands of genes, such that each cell is infected with a single virus targeting one gene. Cells are then selected with puromycin to eliminate uninfected cells. Finally, cells that exhibit the desired phenotype are isolated, and the identity of the shRNAs within those cells is determined using microarrays that contain probes for all library shRNAs (Figure 2.14).

To test the pooled screening approach, Luo *et al*[137] infected human Jurkat T cells with ~45,000 shRNAs in duplicate infections. One infected population was treated with FASL to kill cells through the FAS pathway. The other population was not treated. By this design, shRNAs that cause resistance to FASL were enriched in the treated population, and their target genes to be essential for the FAS killing pathway. Next, DNA was isolated from the infected population, amplified the shRNAs, labeled with fluorescent nucleotides and hybridized to a custom Affymetrix array containing probes complementary to all the shRNAs in our library. When cells were treated with FASL for one week, a group of shRNAs was enriched in cells that were resistant to FASL compared to untreated cells. shRNAs conferring FASL resistance were re-infected individually into Jurkat cells and were able to induce resistance to FASL similarly to the pooled screen results. Of the 14 genes identified with multiple targeting hairpins, several known genes were identified, including FAS, FAD and caspase 8. Eleven additional genes were identified, four out of five of which showed correlation of phenotype and knockdown, suggesting that the hit shRNAs were acting through the intended target. The powerful capability of pooled library screening coupled with custom microarrays to detect enriched or depleted shRNAs enables rapid identification of essential genes, and makes it feasible to consider parallel genome-wide screens to test specific hypotheses and explore multiple experimental conditions.

Compared to genome-wide arrays, pooled shRNA screens are ideal for rapidly finding a small number of robust hits and for conducting comparative screens across conditions and cell types. In contrast, arrayed screens are best for identifying many hits because of their high sensitivity, but are not suited to comparative screening due to their higher cost of reagents and labor. Initially, we proposed to use both approaches to dissect nucleic acid responses in order to isolate the most dramatic hits using pooled libraries but to test more deeply a high-value candidate set using arrayed screens.

Arrayed candidate siRNA screens

Using HBECs optimized for siRNA transfection (Figure 2.6), we demonstrated the efficacy of siRNA-based arrayed screens in a dissection of host-influenza interactions[134]. In this screen, 1745 candidate genes identified through physical interactions, transcriptional responses and associated pathways were interrogated for a role in influenza infection regulation. To assess the functional contribution of the candidate genes on viral replication and type I IFN production, three functional assays were used to measure the effect following genetic perturbation with siRNA pools. First, siRNA-transfected primary HBECs were infected with PR8 virus and virus production was measured after 48 hours using a cellular reporter system that is analogous to conventional plaque assays. Additionally, a reporter cell line detecting IFN-β was used to detect changes in siRNA-transfected HBECs in response to either ΔNS1 virus infection or viral RNA transfection. The resulting data point to potential roles for some unanticipated host and viral proteins in viral infection and the host response, including a network of RNA binding proteins, components of WNT signaling and viral polymerase subunits.

The success of our host-influenza interaction screen, coupled with our difficulty in generation an effective reporter cell and inherent genome-wide screening costs,

influenced our decision to continue with a candidate-based siRNA screen. The candidate list of ~800 genes of both experimentally derived and hypothesis-based candidates discussed are discussed in detail in Chapter 3.

2.6 – *Screening Strategy*

The final phase of screening development was divided into two parts: siRNA transfection optimization and ISD stimulation. For the transfection of siRNAs we considered siRNA concentration, multiple transfection reagents, transfection application methods (forward vs. reverse transfection methods) and siRNA-to-transfection reagent ratio. Additionally, we tested cell seeding density, media/serum concentration conditions, time to optimal knockdown of a panel of positive, negative, and transfection control siRNAs. Secondly, in the context of our siRNA screens we reevaluated ISD transfection conditions including ligand concentration and time of supernatant collection. Furthermore, with the goal of reducing edge effects associated with long-term incubation of tissue culture microplates, we considered plate layout design and investigated outer well contents and incubator plate stacking variables. Finally, we tested experimental workflow by testing a combination of robotic versus by-hand setups and assessed our screening performance using the Z-factor scoring method[138].

Optimization of siRNA conditions

First, we sought out to optimize siRNA transfection conditions in p53-/- MEFs. MEFs were transfected with increasing concentrations of a non-targeting control and *Irf3* siRNAs (10-50nM) then stimulated with ISD 72-hours following treatment (Figure 2.15A). Cxl10 protein expression reduction correlated with *Irf3* knockdown efficiency at an optimal 25nM concentration (>12-fold CXCL10 reduction compared to ISD and >60% *Irf3* knockdown, p-value <0.001, Student's t-test). We compared a large panel of transfection

Figure 2.15

Optimization of siRNA conditions in p53$^{-/-}$ MEFs: siRNA transfection conditions. A-B) Cxcl10 protein expression was measured by ELISA 72-hours following siRNA transfection and 6-hours after ISD stimulation across a range of indicated siRNA concentrations. Reduction in IRF3 expression was measured by quantitative RT-PCR. siRNA complexes were generated with the indicated transfection reagents and stimulated with ISD 72-hours following treatment. p53$^{-/-}$ MEFs that constitutively express GFP were treated with transfection complexes added to cells as they were being seeded (Forward) or 24-hours following seeding (Reverse) (C) or with the indication volume (in ul) of transfection reagent per well (D). GFP expression was detected by quantitative RT-PCR. A matrix of the indicated siRNA transfection reagent volume and siRNA concentration were assessed by RT-qPCR (E) and by automated microplate cytometry plotted as a ratio GFP mean fluorescent intensity to Hoecsht mean intensity for the indicated conditions (F). G) Fluorescent microscopy images of optimal knockdown conditions.

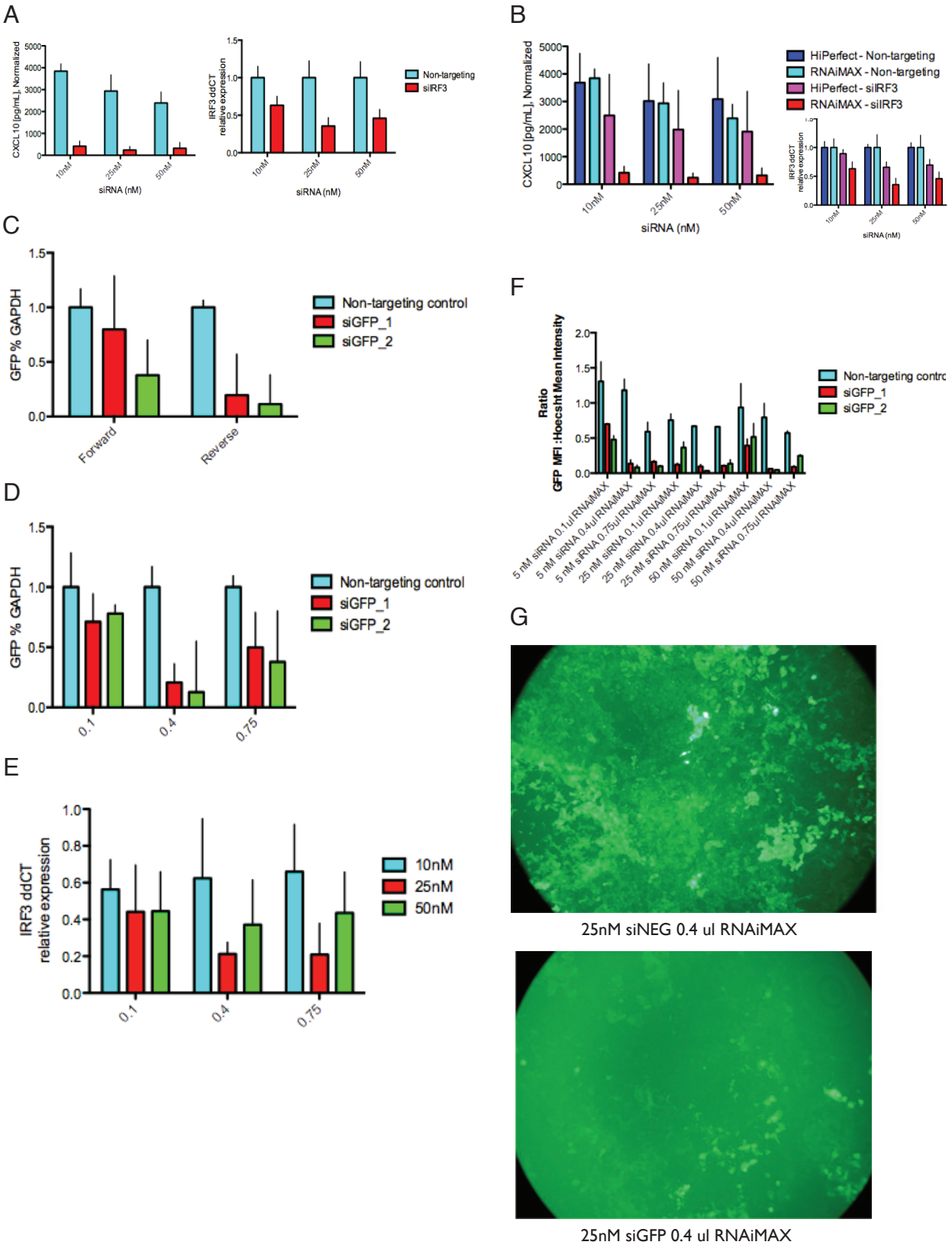


Figure 2.15 (continued)

reagents (DharmFECT1-4 (Dharamcon), Lipofectamine LTX, Lipofectamine 2000 and Lipofectamine RNAiMAX (Invitrogen), and HiPerfect, Attractene and Effectene (Qiagen) transfection reagents) (data not shown) and demonstrated that one reagent (Lipofectamine LTX) resulted in the largest fold-change reduction of Cxcl10 response (>12-fold) and knockdown efficiency (>60% IRF3 relative to non-targeting control) (Figure 2.15B).

To better demonstrate siRNA transfection conditions, we generated p53-/- MEFs that constitutively expressed GFP and compared independent GFP-targeting siRNAs across a number of conditions (Figure 2.15C-D,F-G). To optimize the order in which transfection occurred we initiated knockdown in p53-/- MEFs at the same time as cell seeding (Forward) or 24-hours after cells were seeded (Reverse) (Figure 2.15C). GFP expression was reduced three-times more in reverse-transfected cells than in forward-transfected cells compared to non-targeting control (p-value <0.001, Student's t-test). Additionally, we determined that transfection reagent volume was inversely correlated with knockdown (Figure 2.15D). We observed, however, increased cell death with increased transfection reagent volume. In an effort to find the ideal ratio of siRNA concentration with transfection reagent, we tested a matrix of *Irf3*-knockdown conditions and determined that 0.4ul of transfection reagent per well in conjunction with 25uM of siRNA reduced *Irf3* expression greater than other conditions and with the least effect on cell viability (Figure 2.15E). As independent confirmation, we tested a matrix of knockdown conditions and assessed knockdown efficiency by microplate fluorescent cytometry and normalized GFP mean fluorescent intensity to Hoecsht mean intensity (Figure 2.15 F-G). We identified a condition (0.4ul RNAiMAX per well, with 25nM siRNA) that reproducibly maximized knockdown efficiency and reduced siRNA-induced toxicity,

resulting in more than 20-fold reduction of GFP expression in siGFP treated cells compared to non-targeting control (p-value <0.0001, Student's t-test).

Lastly, we considered cell-seeding density, serum percentage in media and optimal time to knockdown (Figure 2.16). Additionally, we tested a panel of positive and negative siRNA controls to demonstrate the optimal knockdown condition following stimulation with ISD. First, p53-/- MEFs were seeded at 500 or 1000 cells per well in 100ul of media and transfected with control siRNAs 24-hours later. Cells seeded at a lower density had a greater fold change of Cxcl10 protein expression of non-targeting to *Irf3* treated wells possibly due to the fact the siRNA transfection was more efficient at lower cell density (>6-fold reduction compared to <3-fold reduction, p-value <0.0001, Student's t-test) (Figure 2.16A). Transfection complexes are formed in serum-free media and allowed to form for 15-minutes prior to addition to cells. We demonstrated that dilution of culture media with serum-free media improved fold-change responses in non-target to *Irf3* and *Tbk1*-knockdown wells (26-fold difference in media diluted 50 percent compared to 9-fold in undiluted media, p-value <0.0001, two-way ANOVA) (Figure 2.16B). Over a course of four-days following siRNA transfection, we assessed knockdown efficiency in the context of ISD stimulation (Figure 2.16C) and expression of a panel of control genes (Figure 2.16D). First we compared fold-reduction of *Cxcl10* in response to ISD following 48 or 72-hours of transfection with siRNAs targeting *Irf3*, *Tbk1* and *Rig-I* and noted improved reduction at 72-hours in *Irf3* and *Tbk1* treated cells. Furthermore, we analyzed expression of eight control genes (*Aim2*, *Dai*, *Ikke*, *Mavs*, *Irf3*, *Irf7*, *Rig-i* and *Tbk1*). Cells transfected with siRNAs for 72-hours consistently showed the greatest reduction in gene expression as measured by quantitative RT-PCR. Lastly, in a test of screening conditions, p53-/- MEFs were transfected with control panel siRNAs and stimulated with ISD 72-hours following RNAi treatment (Figure 2.16E). *Cxcl10*

Figure 2.16

Optimization of siRNA conditions in p53$^{-/-}$ MEFs: Cell culture conditions effect knockdown efficiency. Control gene panel validation. A) MEFs were seeded with the indicated number of cells per well in 96-well plates, transfected with siRNAs for 72-hours and then stimulated with ISD for 26-hours. Cxcl10 expression was measured by ELISA. B) Cell culture media was diluted with Opti-MEM serum-free media to decrease the percent FBS during transfection of siRNAs. Cells were stimulated with ISD for 26-hours and Cxcl10 response was measured by ELISA. C) p53-/- MEFs were transfected with the indicated siRNAs for 48 or 72-hours, then stimulated with ISD for 26-hours. Fold reduction in Cxcl10 expression was determined by ELISA and compared to non-targeting control response to ISD. D) A panel of siRNAs targeting known nucleic acid sensing components. Lysates were collected following transfection for the indicated times. Gene expression was detected by quantitative RT-PCR. E) Control panel siRNAs were transfected for 72-hours and stimulated with ISD for 26-hours. Cxcl10 response was measure by ELISA.

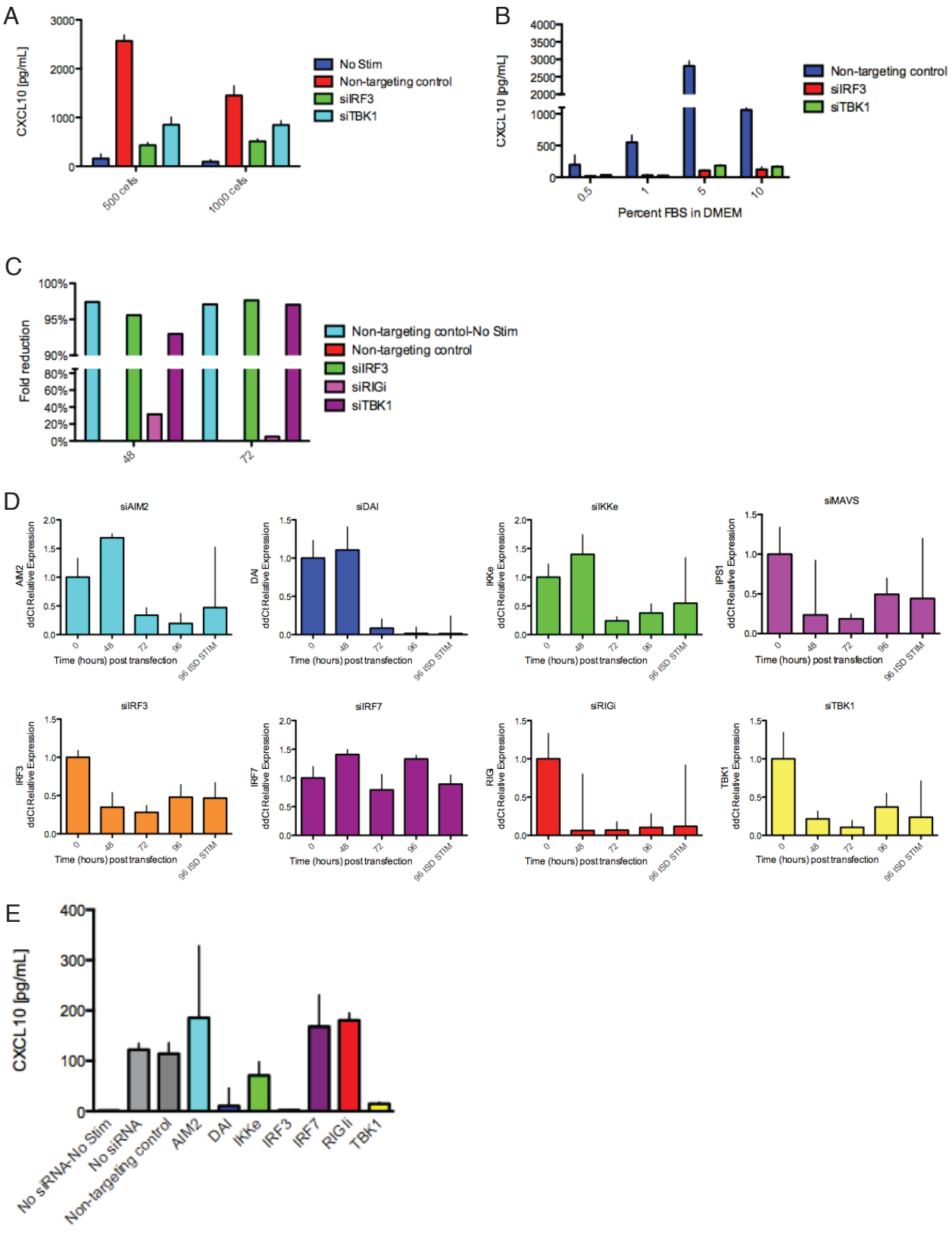


Figure 2.16 (continued)

expression was reduced by >50-fold in *Irf3* treated MEFs compared to non-targeting controls (p-value <0.001, Student's t-test).

Optimization of ISD stimulation conditions

Following the optimization of siRNA conditions, we revisited ISD transfection variables to validate stimulation conditions in the context of our RNAi system. Following knockdown with non-targeting and *Irf3* siRNAs, cells were stimulated with a range of ISD from 0 to 1.5ug/mL. We demonstrated maximal fold change of *Cxcl10* between non-targeting and *Irf3* treated was at 1ug/mL ISD (>13-fold reduction, p-value <0.001, Student's t-test) (Figure 2.17A). Furthermore, increased ISD stimulation time correlated with increased *Cxcl10* expression in non-targeting control treated cells (Figure 2.17B). Over a time course with 3-hour increments beginning at 12-hours, *Cxcl10* expression increased up to a point of saturation (>30-hours, data not shown). Furthermore, fold reduction in *Cxcl10* expression increased with longer incubation time.

We determined that by adding media to the outer wells our sample plates and stacking PBS filled plates on the top and bottom of each stack we significantly reduced edge effect. Finally, we tested experimental setup by testing a combination of robotic versus by-hand setups. Plate-to-plate variation was further reduced with the inclusion of these methods (data not shown).

Assay workflow

In brief, the final assay workflow is described below (Figure 2.18):

- Day 0 - Seed 750 cells/well (96-well plate) by hand in 60% DMEM in D10/OptiMEM (60ul DMEM with 10%FBS w/out Pen/Strep plus 40ul OptiMEM), entire plate is seeded, outer wells are supplemented with an additional 130ul of media for a final volume of 230ul for outer wells. In the incubator, three plates are sandwiched between PBS filled plates.

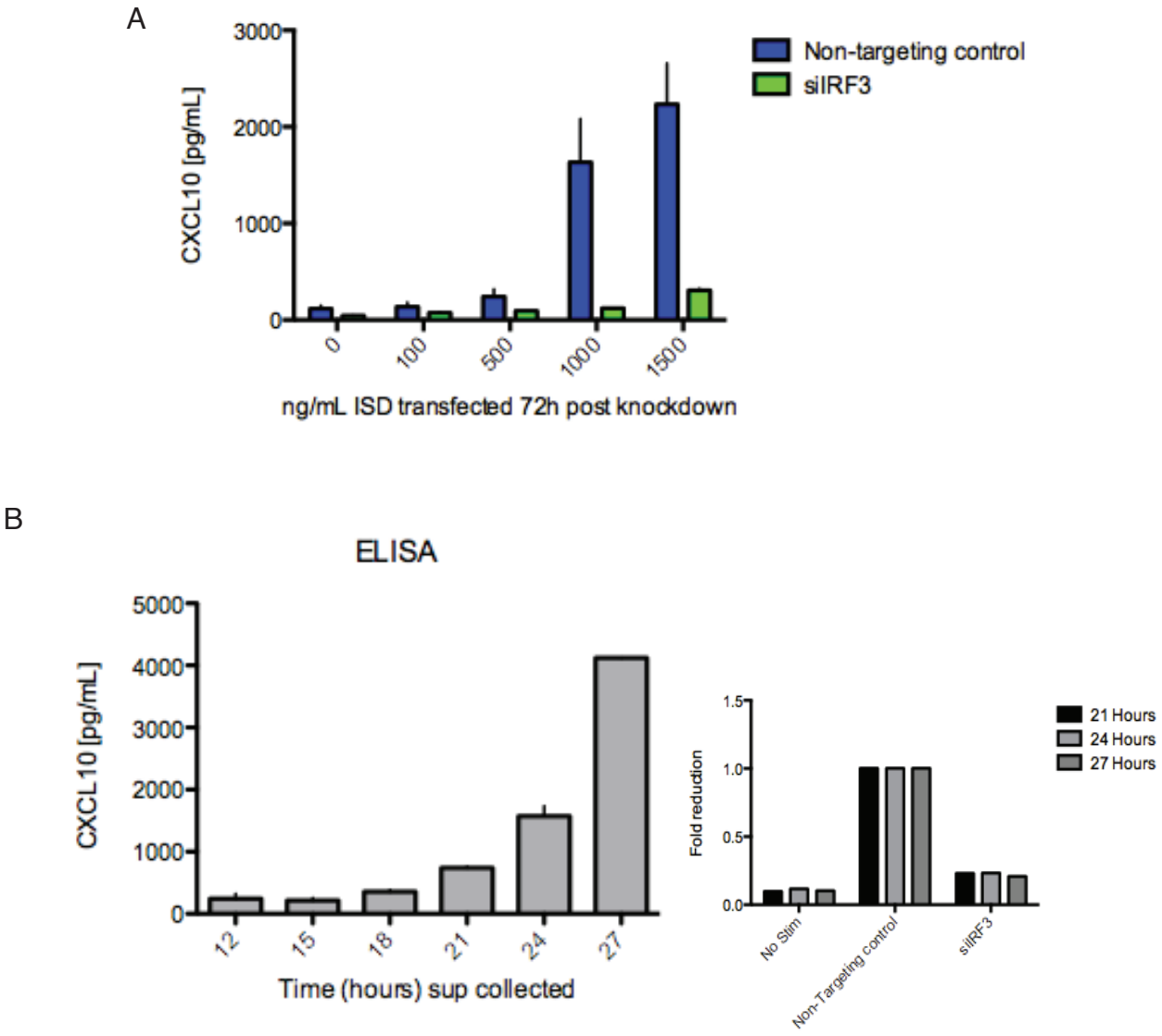


Figure 2.17

ISD stimulation following siRNA knockdown A) Following siRNA treatment, p53-/- MEFs were stimulated with the indicated amounts of ISD for 26-hours. Cxcl10 expression was detected by ELISA. B) Cells were stimulated with ISD for the indicated times following transfection with non-targeting control or *Irf3* siRNAs. Supernatants were collected following stimulation, frozen and then assayed for Cxcl10 protein expression by ELISA.

Figure 2.18

Assay workflow. 750 p53-/- MEFs per well were seeded in 96-well plates in 60% DMEM and 40% Opti-MEM. Each plate included three non-targeting controls (ASN, Dharmacon All-star negative control), two *Irf3* positive controls, an siDeath control and a buffer control. Outer edges were seeded with cells and supplements with an additional 130ul of media (total volume 230ul) but not transfected with siRNAs. 25 nM siRNA was complexed with 0.5 uL Lipofectamine RNAiMax (Life Technologies) in Opti-MEM, incubated for 12 min at 22°C, and added to the wells. 72-hours later, cells were transfected with 1ug/mL of ISD. 26-hours following stimulation, supernatants were collected and Cxcl10 was quantified by ELISA. Cell viability was estimated by the CellTiter-Glo Luminescent Cell Viability Assay (Promega).

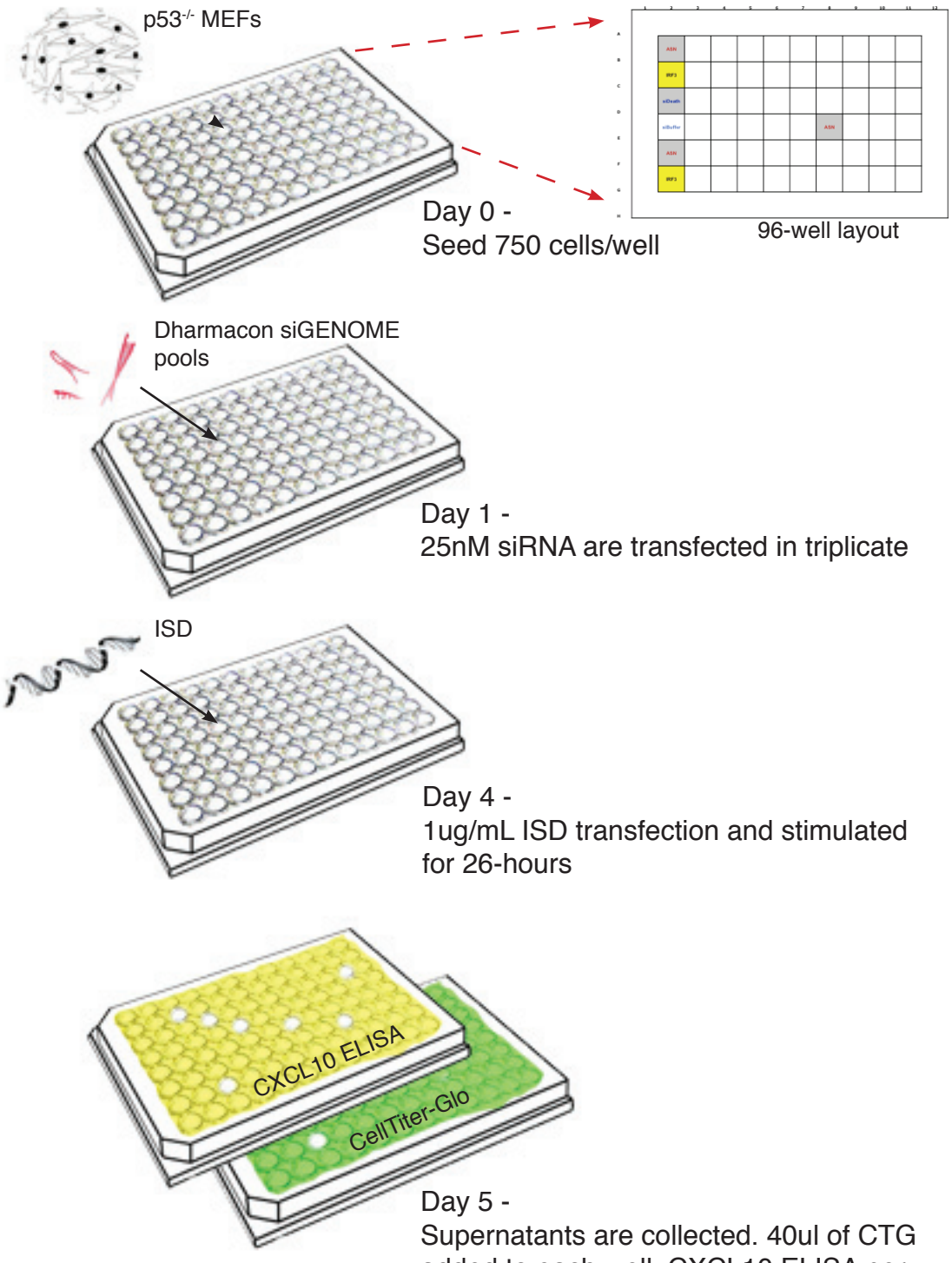


Figure 2.18 (continued)

- Day 1 - Transfect siRNAs in triplicate. 25nM final in 20ul OptiMEM with 0.5 ul RNAiMAX. Complexes are formed for 12 minutes at room temperature then slowly added drop-wise to each well, 10 wells at a time by hand.
- Day 4 - 72-hours post siRNA transfection, change media, add ISD transfection complexes at 1ug/mL final in 20ul OptiMEM with 0.36 ul Lipofectamine LTX. Transfection complexes are formed for 30 minutes at room temperature and then slowly added drop-wise to each well, 10-wells at a time by hand.
- Day 5 - 26-hours post ISD transfection, collect and freeze supernatants. Add 40ul CellTiter-Glo, cover and incubate for 8-10 minutes. Read on EnVision plate reader.
- Day 6 - Pre-Elisa. From pre-selected control wells, determine optimal supernatant volumes to prevent saturation of luminescent signal.
- Day 7 – Perform CXCL10 ELISA using calculated supernatant volumes.

Screening performance assessment

To assess screening performance we used the Z-factor scoring method[138]. The Z-factor reflects both the dynamic range of the assay and the data variation associated with the signal measurements, in this case, the variation of the siRNA treatment, tissue culture, media changes, etc., and is a critical quality assessment of the screening fitness. The closer the Z' calculation approaches 1, the more robust the assay. To perform this test, we transfected one half of a test plate with a non-targeting control (Qiagen All-Star Negative) and the second half with siRNAs against IRF3 (Dharamcon On-Target-Plus (OTP) Smart Pools) (n=60), (Figure 2.19). Prior to stimulation media was replaced in half of the wells 72-hours following siRNA treatment (n=30). We considered

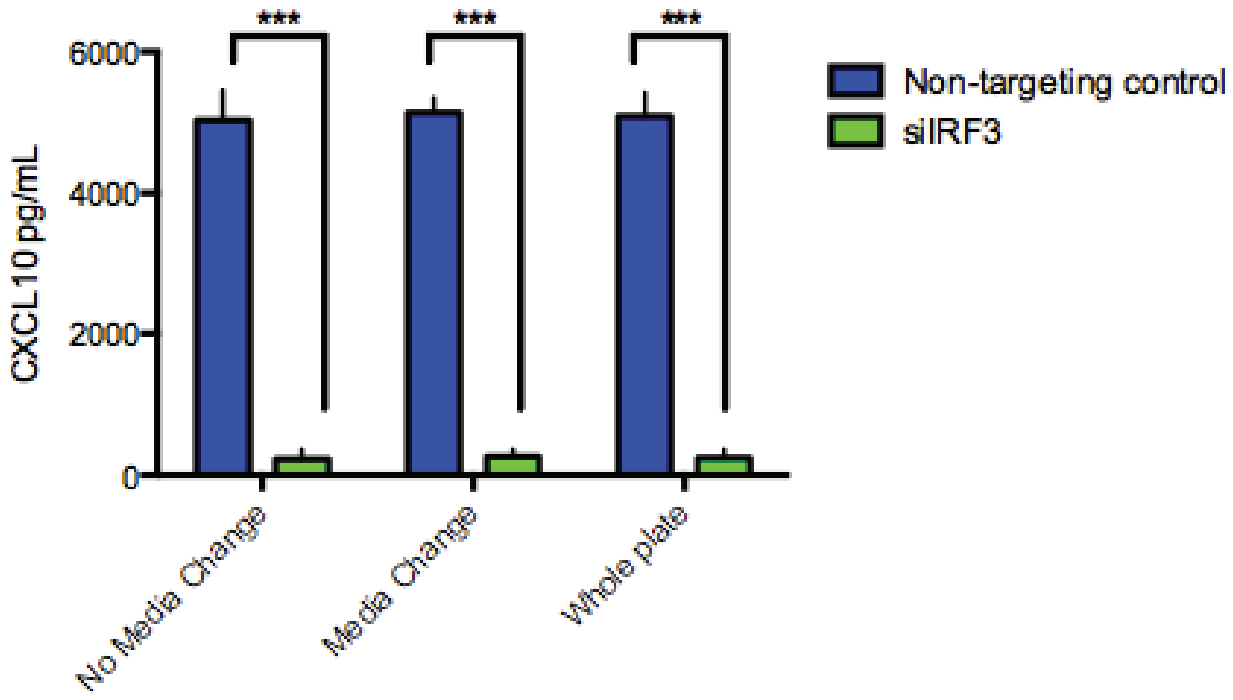


Figure 2.19

Z-factor test of control siRNAs confirms a robust screening method. p53-/- MEFs were transfected with the indicated siRNAs and then stimulated with ISD as described previously. Z-factor for pilot tests comparing no media change, media change or whole-plate analysis. N=30/sample x 3 replicates. P-value <0.0001, Student's t-test.

media change prior to stimulation to eliminate well-to-well difference caused by evaporation during the siRNA-treatment phase.

Following the workflow described above, cells were then stimulated with 1ug/mL for 26-hours. A range of supernatants was tested from positive and negative control samples to determine the appropriate supernatant volume to screen the samples with. Cxcl10 values were normalized to CellTiter-Glo values. The screen performance score, $Z=1-(3xSD_{neg}+3SD_{pos})/abs(mean_{neg}-mean_{pos})$, was calculated for each condition and for the entire plate. The Z-scores for each condition (no media change, media change and whole plate) were 0.775, 0.836, and 0.802, respectively, each representing a 20-fold reduction in Cxcl10 expression in IRF3-treated samples compared to non-targeting control (p-value <0.0001, Student's t-test).

2.7 – *Conclusion*

We developed a robust and repeatable high-throughput loss-of-function screening tool to aid in the dissection and identification of novel ISD-sensing pathway components. In the process of development we assessed multiple nucleic acid DNA and RNA ligands, investigated the ISD response in multiple human and murine primary and established cells lines, tested multiple type I IFN reporter systems and considered multiple genetic perturbation techniques. Our siRNA-based screening system will be used to interrogate the function of 1003 candidate genes discussed in the following chapters.

Chapter 3:

Generation of a Candidate Gene Set by Curation and Quantitative Proteomics

3.1 – *Introduction: candidate gene selection*

Following the completion of an siRNA-based screening strategy, we set out to generate a set of candidate genes from genomic, proteomic, and domain-based studies that we hypothesized contain unidentified ISD-sensing pathway components. We selected 1003 ISD-sensing pathway candidates using the following criteria. First, we used previously published array data and our own gene expression experiments to select genes regulated in response to IFN-β or DNA. Second, we conducted SILAC-based mass spectrometry experiments using ISD as bait in IFN-β stimulated cells. Third, because of the well-established role of helicases in nucleic acid sensing, we hypothesized that any annotated helicase could have a potential role in dsDNA sensing and included all available helicases as candidates. Fourth, in addition to nucleases, we identified genes that had known DNA binding properties and were putatively localized to the cytoplasm based on established localization and published predictive algorithms. Lastly, we focused on annotated phosphatases and deubiquitinases as part of our pilot screen to identify potential negative regulators of the ISD pathway. We supplemented the annotated phosphatases and deubiquitinases with putative negative regulators of the RIG-I pathway identified in our recent dissection of host-influenza interactions[134]. Our final candidate selection integrates genes from genomic, proteomic and domain-based data sets that we hypothesized contain unidentified ISD-sending pathway components (Figure 3.1).

3.2 – *Candidate gene selection: interferon-regulated, DNA-stimulated genes from published arrays*

First we selected 355 DNA- and interferon-stimulated genes (ISGs) from existing microarray datasets based on the hypothesis that a subset of components of this

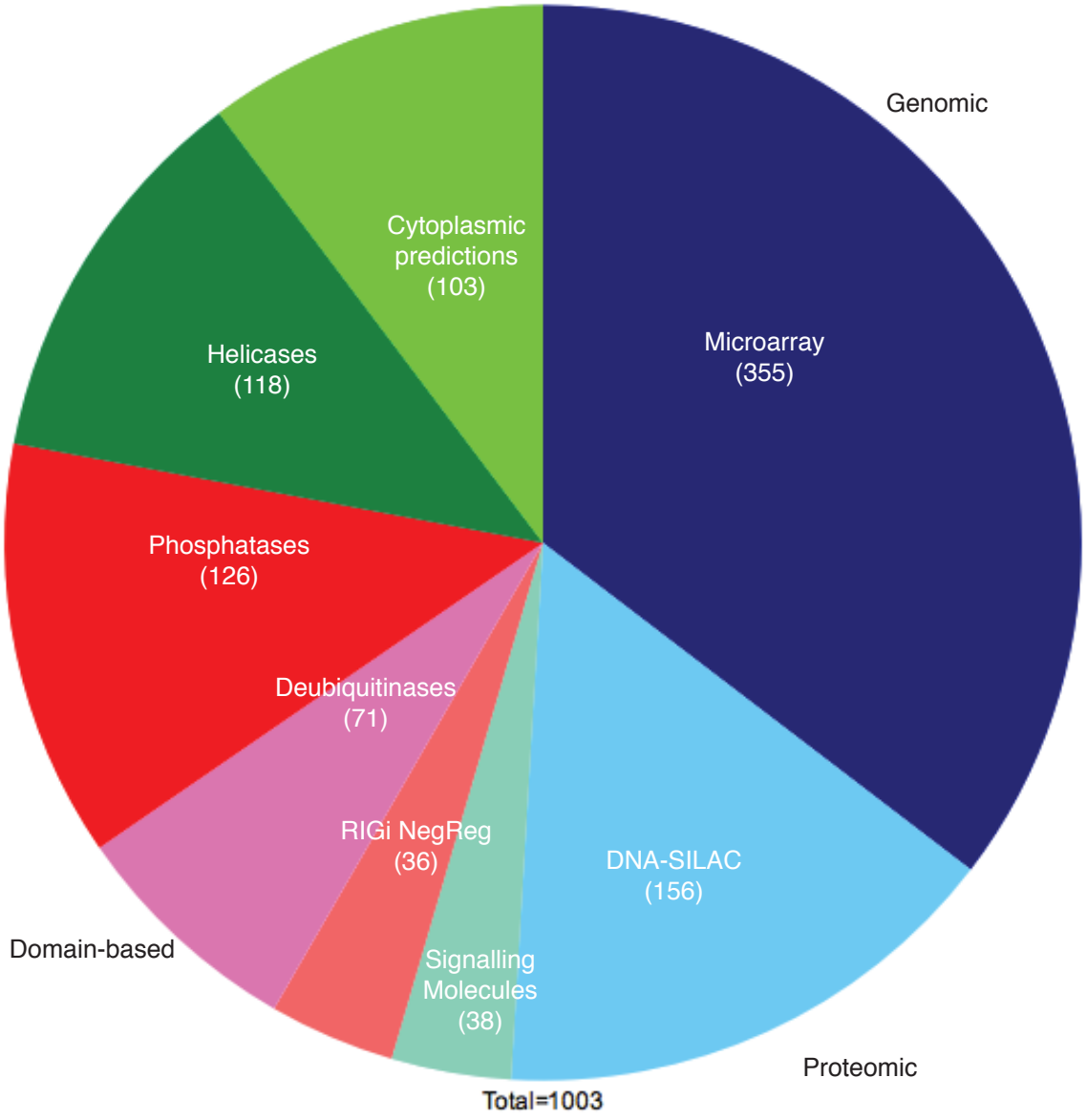


Figure 3.1

Generation of a candidate gene set by curation and quantitative proteomics. Sources of the 1003 ISD-sensing pathway candidates from genomic, proteomic and domain-based datasets.

pathway is feedback-regulated (Table 3.1)[43, 44, 48]. For example, in an expression-based screen of a mouse carcinoma cDNA library reporting IFN-β, TRIM56 was identified as an interferon-inducible regulator of the dsDNA-mediated type I IFN response[139]. TRIM56 interacts with STING, promoting STING ubiquitination and subsequent TBK1 recruitment and type I IFN induction. Additionally, biotinylated dsDNA 70-mers derived from Vaccinia virus were used to affinity purify DNA-binding proteins from cytosolic extracts of THP-1 human monocytes[65]. Among the proteins identified was the AIM2-like IFI16. Recent studies ascribe a role of IFI16 in the direct association with viral DNA upon stimulation, driving recruitment of STING and subsequent IFN induction.

We cross-referenced data from three independent microarrays and selected the strongest hits among them. Genes with available siRNA pools were included as candidates. First, we surveyed arrays from a screen to identify genes that were transcriptionally regulated by IFN-β[48]. NIH3T3 and L929 cells were stimulated for four hours with recombinant IFN-β, RNA was isolated and global changes in gene expression were analyzed by microarray analysis. We selected 225 genes with greater than 3.9-fold upregulation after IFN-β stimulation of NIH3T3 cells, and greater than 6.25-fold upregulation after IFN-β stimulation of L929 cells (Gene Expression Omnibus database: submission #GSE14413), including many well-known interferon-induced genes.

In a seminal report providing a dissection of the DNA signaling pathway, MEFs from wild-type, *Tbk1*$^{-/-}$, *Ikki*$^{-/-}$ and *Tbk1*$^{-/-}$/*Ikki*$^{-/-}$ mice were stimulated with Poly (dA:dT) for four hours and expression profiles were determined by microarray[43]. DNA included the upregulation of many interferon-inducible antiviral genes in a mostly *Tbk1*-dependent manner. Amongst the wild-type stimulated cells, we selected 196 genes with greater 6.1-fold upregulation after poly(dA-dT)–poly(dT-dA) stimulation (Gene Expression Omnibus database: submission #GDS1773). Additionally, in a comparison of the ISD pathway to

Table 3.1

Candidate Gene List: Published Array Curation

			AKIRA		Superti-Furga		Medzhitov	
Gene	GeneID	Accession	Fold Change	Rank	Fold Change	Rank	Fold Change	Rank
0610033I05RIK	71684	NM_030243	4.84	485	1.32	177	2.60	3221
0610039C21RIK	66853	NM_025802	7.00	279	0.99	22457	1.02	16271
1110001A05RIK	56376	NM_019808	0.82	9589	1.36	152	1.90	5484
1110001A05RIK	56376	NM_019808	0.82	9589	1.36	152	1.90	5484
1110003E01RIK	68552	NM_133697	0.82	9649	1.32	178	0.88	21566
1110007F12RIK	68487	NM_197986	4.49	545	1.04	11113	4.89	1116
1110049F12RIK	66193	NM_025411	0.69	14030	1.34	158	0.53	36870
1200013F24RIK	66880	NM_025822	1.06	5444	1.40	129	0.63	33026
1700010I14RIK	66931	NM_025851	6.22	329	0.96	31014	0.81	24816
1700010O16RIK	67504	NM_026205	8.72	207	1.02	17002	1.90	5482
1700019G17RIK	75541	NM_029331	25.33	48			1.86	5702
2210412K09RIK	76959	NM_029814	0.79	10648	1.33	171	1.32	10107
2310010I16RIK	66371	NM_025519	6.09	338	1.02	15211	2.93	2690
2310015I10RIK	69550	NM_198095			2.45	57		
2310042N02RIK	71913	NM_024246	24.28	51	1.07	6025	1.65	6969
2310046K10RIK	108670	NM_029495	9.13	183	4.27	29	3.66	1859
2310056K19RIK	75689	NM_080846	9.83	165	1.11	2635	0.78	26205
2310061N23RIK	76933	NM_029803	0.56	18299	1.04	10993	2.64	3147
2310079P12RIK	50724	XM_203577	0.52	19227	1.31	186	0.79	25372
2510004L01RIK	58185	NM_021384	101.18	3	7.20	19	257.11	8
2510038N07RIK	67138	XM_355768			2.96	43	20.71	102
2610007K22RIK	67040	NM_199079	0.81	9883	1.20	605	1.52	7988
2810485I05RIK	72826	NM_176836	1.16	4464	1.36	147	1.06	15183
3110049J23RIK	67307	NM_026085	6.39	315	1.10	3064	1.15	12776
3632410F03RIK	74025	NM_028721	1.03	5843	1.31	180	0.65	32014
4921522K05RIK	67981	NM_026489	13.13	106	1.06	8323	0.42	40300
4930547C10RIK	68274	NM_026652	10.50	147			1.32	10151
4932437H03RIK	219103	XM_127861			1.07	6877	1.93	5376
4933411K20RIK	66756	NM_025747	10.24	155	1.09	4073	0.96	18255
4933430F08RIK	74481	NM_028967	5.26	424	1.53	100	11.49	230
5330431N24RIK	101187	NM_181402			1.09	4127	3.87	1700
5430413I02RIK	56742	NM_019976	7.20	267	1.04	11010	2.36	3820
5830458K16RIK	67775	NM_023386	8.57	215	0.92	36146	82.23	28
5830484A20RIK	109032	NM_175397			2.33	59		
6330442E10RIK	268567	NM_178745	6.43	311	0.99	23150	2.72	3014
8430423A01RIK	98415	NM_175294	1.65	2257	1.40	127	1.40	9133
9130002C22RIK	74558	NM_029000						
9130009C22RIK	71586	NM_027835	37.88	28	13.57	7	21.03	100
9230105E10RIK	319236	NM_175677			1.97	68	3.48	2012
9830137M10RIK	237886	NM_172796			1.03	13492	3.89	1684
9830147J24RIK	229900	NM_145545	19.25	72			7.88	432
9930111J21RIK	245240	NM_173434						
A130009K11RIK	384309	NM_201373			1.46	110		
A130072J07	240832	NM_172843	1.94	1729	0.99	22823	5.62	864
A230050P20RIK	319278	NM_175687	1.75	2025	1.59	93	2.42	3640
A430056A10RIK	99899	NM_133871	22.38	54	8.57	17	51.26	42
A630026L20	668139	NM_172906						
A730024F05RIK	67988	NM_198295			1.32	176		
A930026L03RIK	243382	NM_175523			1.36	148	4.82	1146
AA175286	209086	NM_010156	8.96	189	1.19	639	6.39	686
ABCF1	224742	NM_013854	6.80	287	1.00	20430	2.31	3971
ABHD3	106861	NM_134130	17.83	78	1.02	16797	0.69	30403
ADAR	56417	NM_019655	3.81	702	1.74	80	5.81	807
AGRN	11603	NM_021604	1.05	5590	1.42	122	4.38	1383
AGTRL1	23796	NM_011784	9.39	176	0.89	39382	1.91	5457
AI447904	236312	NM_175026	2.44	1254			24.05	87
AI481100	54396	NM_019440	49.36	13			21.32	97
AI595338	100702	NM_194336	113.95	2	0.96	29932	58.29	38
AIM1	11630	NM_172393	27.85	43	0.91	37362	2.61	3204
AKT3	23797	NM_011785	1.78	1985	0.97	27366	3.04	2542
AL033326	24105	NM_019705	0.78	10859	1.36	149	1.28	10687
APOBEC1	11810	NM_031159	1.84	1884	1.03	13537	1.05	15467
APOBEC3	80287	NM_030255	2.68	1119	1.18	705	3.96	1632
ARHGAP9	216445	NM_146011	2.41	1270	1.34	155	0.57	35467
ARIDSA	214855	NM_145996	1.26	3749	0.99	22361	8.58	373
ASB13	142688	NM_080857	9.76	166	0.99	23425	7.22	538
ASB14	142687	NM_080856	7.38	260	0.99	23569	0.52	37048
ATP1B4	67821	NM_133690	10.86	134	0.98	26239	0.53	36715
AW111922	60440	NM_021792	38.69	26			173.76	13
AW539457	99382	NM_178890			1.02	16715	17.75	126
AY036118	170798	NM_133243						
B2M	12010	NM_009735	1.03	5849	1.47	106	1.20	11866
BAZ2A	116848	NM_054078	1.53	2572	1.07	6803	0.86	22195
BC003281	80285	NM_030253	8.15	229	3.90	34	7.64	470
BC004022	80750	NM_030563			1.12	2076	2.48	3503
BC006779	229003	BC026386	6.68	294	3.29	38	17.93	123
BC021340	547253	NM_145481	19.23	73	1.01	19190	15.24	152
BC022145	217203	NM_144830	1.82	1916	0.99	24167	2.06	4786
BC028975	242584	NM_146254	0.70	13658	1.43	120	0.44	39643
BC032925	242122	NM_178594			0.96	30125	0.92	19728
BC062109	231503	NM_182841			1.45	112	1.12	13475
BRD4	57261	NM_020508	6.28	324	1.09	4432	5.13	1011
BTN1A1	12231	NM_013483	9.22	179	1.09	4125	1.30	10417
C330003B14RIK	105594	NM_175342	8.33	223			1.64	7040
CABP1	29867	NM_013879	6.86	285			0.78	26277
CARHSP1	52502	NM_025821	1.08	5235	1.02	15487	1.46	8504
CASP7	12369	NM_007611	2.11	1517	1.32	173	2.23	4199
CCNB3	209091	NM_183015	6.67	295	0.97	27398	0.56	35868
CDADC1	71891	XM_127813	9.40	175	1.11	2719	1.07	14785
CHEK1	12649	NM_007691	7.19	269	1.02	15062	1.85	5742
CHN2	69993	NM_023543	6.43	310			0.71	29550
CNTN3	18488	NM_008779	6.24	328	1.04	11905	0.67	31176
COPG2	54160	NM_017478	0.99	6370	1.36	150	1.00	16935
COPS5	26754	NM_013715	6.60	301	1.08	4725	0.84	23141
CPEB3	208922	NM_198300	1.51	2640	1.10	3233	4.48	1310
CSDA	56449	NM_011733	0.75	11780	1.31	188	0.83	23776
CSF1	12977	NM_007778	0.97	6695	1.32	174	1.94	5307
CSPRS	114564	NM_033616	4.25	587			3.39	2103
CXCL10	15945	NM_021274	28.55	42	0.94	33461	183.08	12
D11LGP2E	80861	NM_030150	63.81	7	1.14	1605	25.77	77
D14ERTD668E	219132	NM_199015						
D15ERTD366E	65970	NM_023063	0.92	7525	1.43	119	1.62	7173
D1PAS1	110957	NM_033077	12.85	108	1.07	5951	3.53	1963
D430033A06RIK	100978	NM_177007	0.69	14053	1.03	13600	6.52	660
D630045D17RIK	213311	NM_178674			1.31	184	1.15	12809
D7ERTD458E	52118	NM_027514	8.94	192			3.74	1796
DAB1	13131	NM_010014	10.44	149	0.96	30115	2.74	2975
DAF2	13137	NM_007827	20.50	65			1.82	5899
DAXX	13163	NM_007829	3.88	686	0.99	24281	9.12	323
DCPP	13184	NM_019910	0.75	11863			0.55	36298
DDB2	107986	NM_028119	9.13	182	1.00	22087	0.61	33868

Table 3.1

Candidate Gene List: Published Array Curation (continued)

			AKIRA		Superti-Furga		Medzhitov	
Gene	GeneID	Accession	Fold Change	Rank	Fold Change	Rank	Fold Change	Rank
DHX29	218629	NM_172594			1.37	141	0.83	23643
DNASE1L3	13421	NM_007870	1.16	4467	1.12	2440	16.46	133
E430029F06	236573	NM_172777			0.96	29932		
EAR5	54159	NM_019398	8.98	188	0.93	34885	1.11	13676
ELF1	13709	NM_007920	2.57	1165	1.30	205	1.46	8524
EPM2AIP1	77781	NM_175266	0.39	21097	1.31	189	0.64	32666
eva	14012	NM_007962	0.83	9576	1.34	160	7.24	537
FALZ	207165	XM_126724	6.20	331			0.89	21241
FLN29	231712	NM_172275	3.32	841	1.04	10836	5.60	874
FOXA1	15375	NM_008259	11.95	115	1.42	121	1.98	5141
G430041M01RIK	101214	NM_198102			1.15	1300	1.87	5651
GALNT15	67909	NM_026449	12.26	114	1.01	19227	2.83	2846
GAS7	14457	NM_008088	8.69	209	1.01	19429	3.91	1674
GBP1	14468	NM_010259	23.37	53			2.40	3705
GBP2	14469	NM_010260	25.70	47	0.98	25445	7.87	436
GBP3	55932	NM_018734	17.67	79	0.94	33068	7.42	497
GBP5	229898	NM_153564			0.98	26050		
GCA	227960	NM_145523	7.89	238	1.08	4854	7.31	519
GCH	14528	NM_008102	0.90	7813	1.07	6826	9.29	314
GDAP10	14546	NM_010268						
GIMAP4	107526	NM_174990	10.24	153	1.02	15289	1.75	6313
GLUD	14661	NM_008133	0.76	11498	1.37	140	0.73	28659
GNG1	14699	NM_010314	8.76	201	0.98	26697	2.79	2894
GPR34	23890	NM_011823	11.00	129	1.03	12897	0.39	40831
HAK	225638	XM_128981	8.77	200	1.13	1750	3.77	1771
HAP1	15114	NM_010404	1.24	3850	1.07	6995	2.46	3541
HMG20A	66867	NM_025812	6.50	305	0.98	25294	2.25	4127
HMX3	15373	NM_008257	8.56	217			7.13	550
HOD	74318	NM_175606	6.39	314	0.99	22662	1.42	8961
HOXD13	15433	NM_008275	6.26	327	1.17	946	1.78	6149
IFI1	15944	NM_008326	8.64	212	1.02	15640	27.10	71
IFI16	15951	NM_008329	21.10	60			12.15	204
IFI202B	26388	NM_008327	12.68	110			0.51	37579
IFI203	15950	NM_008328	43.41	22			20.33	104
IFI205	226695	NM_172648	52.16	11			18.04	120
IFI35	70110	NM_027320	5.65	392	3.02	41	11.70	221
IFI47	15953	NM_008330	30.70	37			20.22	106
IFIT1	15957	NM_008331	30.16	38			269.84	7
IFIT2	15958	NM_008332	60.40	9	49.85	3	171.05	14
IFIT3	15959	NM_010501	47.33	16	80.26	1	132.84	20
IFITM1	68713	NM_026820	0.60	17204			0.97	17942
IFITM2	80876	NM_030694	0.85	8857	2.17	63	0.97	17760
IFITM3	66141	NM_025378	1.74	2047	3.91	33	2.64	3156
IGTP	16145	NM_018738	33.06	32			21.29	98
IHPK1	27399	NM_013785	6.09	339	1.02	15927	1.78	6124
IL6	16193	NM_031168	4.08	627	0.97	27184	72.25	33
INSL6	27356	NM_013754	6.00	348	1.04	10937	2.02	4937
INSM1	53626	NM_016889	12.65	111	0.95	31421	0.67	31268
IPF1	18609	NM_008814	7.71	245	1.10	3562	0.24	43589
IRF2	16363	NM_008391	3.60	760	1.69	83	2.15	4448
IRF7	54123	NM_016850	14.56	92	3.33	37	67.56	36
ISG20	57444	NM_020583	4.12	618	1.05	9385	47.72	46
ISGF3G	16391	NM_008394	2.06	1579	4.53	28	2.63	3176
JAM2	67374	NM_023844	0.91	7605	1.40	128	1.30	10424
JARID2	16468	NM_021878	0.74	12184	1.08	5219	2.07	4723
KCNE3	57442	NM_020574	8.07	232	0.98	24777	0.16	44510
KLHL10	66720	NM_025727	6.43	309	0.95	31402	1.31	10272
KLK13	13647	NM_010115	7.94	235			0.92	19899
KLK16	16615	NM_008454	7.12	274	0.97	28882	0.37	41304
KLK5	16622	NM_008456	9.04	185	0.97	28882	0.76	26915
KLRA16	27424	NM_013794	9.17	181			1.62	7137
KLRA8	16639	NM_010650	6.40	312			0.87	21867
LAMP3	239739	NM_177356			1.61	90		
LIN7B	22342	NM_011698	7.00	280	1.11	2894	2.45	3559
LMYC1	16918	NM_008506	8.93	193	0.96	29695	1.53	7864
LOC209387	209387	NM_199146	33.78	31			16.36	136
MCL1	17210	NM_008562	0.65	15521	1.37	139	1.49	8206
MEF2C	17260	NM_025282	3.59	763	1.05	9285	1.03	16006
MELA	17276	NM_008581						
MGC6357	208263	NM_144791	3.54	770	1.06	7705	4.26	1446
MOV10	17454	NM_008619	2.73	1096	1.13	1986	5.69	847
MPA2	17472	NM_008620	1.00	6201			3.36	2138
MRPS15	66407	NM_025544	0.71	13139	1.42	124	0.86	22464
MS4A4D	66607	NM_025658	1.80	1948			4.17	1489
MTAC2D1	74413	NM_028924	10.29	152	1.20	529	1.43	8821
MUP4	17843	NM_008648	15.90	87			0.50	37917
MX1	17857	NM_010846	63.07	8			25.10	81
MX2	17858	NM_013606	59.30	10			221.88	10
MYD88	17874	NM_010851	2.41	1275	1.05	9462	2.98	2621
MYEF2	17876	NM_010852	6.44	308	1.12	2414	1.34	9872
MYT1	17932	NM_008665	6.97	282	1.01	18657	0.51	37690
NLGN2	216856	XM_147559	1.48	2742	0.96	29447	2.21	4268
NMI	64685	NM_019401	7.85	240	2.46	55	6.05	761
NR0B1	11614	NM_007430	10.16	158	0.95	31711	5.02	1059
NRTN	18188	NM_008738	1.02	5908	1.40	126	1.43	8807
NTNG2	171171	NM_133500	10.75	139			1.05	15220
NUDT13	67725	NM_026341	4.98	464	1.07	6981	1.13	13181
NUMB	18222	NM_010949	9.27	177	1.07	5863	1.03	15849
OAS1A	246730	NM_145211			3.49	36		
OAS1B	23961	NM_011853	13.30	105			11.68	222
OAS1C	114643	NM_033541	10.21	157			2.44	3590
OAS1G	23960	NM_011852	6.11	337	3.49	36	8.44	382
OAS2	246728	NM_145227	2.84	1055	2.45	56	19.04	113
OAS3	246727	NM_145226	2.77	1077	2.59	49	24.40	84
OASL1	231655	NM_145209	27.85	44	1.60	92	161.95	15
OASL2	23962	NM_011854	46.75	17			52.45	41
OCIL	93694	NM_053109	2.64	1133			2.22	4223
OGFR	72075	NM_031373	1.99	1670	1.51	102	4.54	1279
OLFM3	229759	NM_153157	44.00	21	1.06	8325	0.56	35824
OOG1	193322	NM_178657	6.66	296			3.19	2334
OPLAH	75475	NM_153122	7.18	270	1.00	21654	0.97	17910
PANK2	74450	NM_153501			1.73	81	1.06	15137
PAX6	18508	NM_013627	34.00	30	1.07	6675	2.41	3696
PCBP2	18521	NM_011042	3.71	727	1.01	18352	0.89	20939
Pdcd1lg1	60533	NM_021893	31.03	35	0.87	40199	6.59	643
PELI1	67245	NM_023324	2.35	1319	1.00	22047	9.87	285
PEX13	72129	NM_023651	0.64	15954	1.34	156	1.18	12381
PHC3	241915	NM_153421			1.37	136	0.83	23874
PHEX	18675	NM_011077	8.60	214	0.99	24149	0.39	41002
PHF11	219131	NM_172603					19.49	111
PHIP	83946	XM_358384	10.53	146	1.07	6736	2.83	2845

Table 3.1

Candidate Gene List: Published Array Curation (continued)

			AKIRA		Superti-Furga		Medzhitov	
Gene	GeneID	Accession	Fold Change	Rank	Fold Change	Rank	Fold Change	Rank
PLAGL1	22634	NM_009538	0.82	9759			15.69	142
PLEC1	18810	NM_011117	5.92	353	1.06	7995	8.75	356
PLEKHA4	69217	NM_148927	3.65	743			4.37	1385
PLK2	20620	NM_152804	0.82	9742	1.42	123	4.61	1243
PLSCR1	22038	NM_011636	2.50	1214	4.23	30	1.74	6367
PLSCR2	18828	NM_008880	14.40	94			16.40	135
PML	18854	NM_008884	7.31	262	2.75	46	5.84	803
PNP	18950	NM_013632	3.32	842	1.06	8119	2.54	3343
PNPT1	71701	NM_027869	3.80	707	1.95	70	5.47	908
PNRC1	108767	XM_131355	20.67	63	1.02	15118	1.31	10304
PODXL	27205	NM_013723	8.95	190	1.02	16679	2.44	3602
POU4F3	18998	NM_138945	11.33	126	0.84	41222	1.03	16042
PPM1A	19042	NM_008910	17.50	80	1.05	9267	2.02	4964
PRG2	19074	NM_008920	6.69	293	1.09	4228	1.12	13425
Prkr	19106	NM_011163			2.98	42	7.12	553
PSIP1	101739	NM_133948	1.11	4888	1.08	4964	1.19	12086
PSMB9	16912	NM_013585	17.34	82	1.47	107	7.41	501
PSME1	19186	NM_011189	2.40	1281	1.33	164	1.63	7093
PTK9	19230	NM_008971	0.76	11564	1.46	109	1.25	11104
QK	19317	NM_021881	0.83	9495	1.22	450	1.28	10637
RAB27A	11891	NM_023635	11.50	122	1.02	16442	2.20	4302
RAB3C	67295	NM_023852	10.10	160	1.06	7724	2.08	4703
RAI2	24004	NM_198409	7.22	265	0.88	39994	2.31	3945
RASGEF1B	320292	NM_145839			0.96	29861	12.11	205
RNF31	268749	NM_194346			0.97	28895	2.75	2956
RNF36	70928	NM_080510	0.62	16459	1.90	72	1.18	12303
RS1H	20147	NM_011302	11.33	125	1.05	8907	1.77	6179
RSHL1	83434	NM_031255	11.54	120	0.91	38008	0.40	40732
SAMHD1	56045	NM_018851	8.67	210	1.75	77	3.42	2066
SCOTIN	66940	NM_025858	1.19	4242	1.52	101	1.48	8278
SEC1	56546	NM_019934	9.19	180			0.74	28081
SERPINA3M	20717	NM_009253	7.64	247			1.48	8298
SERPINA6	12401	NM_007618	10.12	159	1.03	13716	1.53	7893
SERPINB9	20723	NM_009256	2.50	1204	1.07	6249	2.38	3755
SERPINI1	20713	NM_009250	8.56	216	1.13	1630	2.37	3777
SHH	20423	NM_009170	10.67	143	1.02	17011	1.04	15768
SLC1A3	20512	NM_148938	0.71	13370	1.37	135	1.25	11039
SLC25A22	68267	NM_026646	3.16	912	0.98	26788	8.14	410
SLC25A28	246696	NM_145156	2.38	1303	1.31	183	1.51	8011
SLC3A1	20532	NM_009205	0.89	8062	1.39	130	2.50	3442
SLC6A14	56774	NM_020049	10.83	136	1.11	3026	1.52	7978
SLC9A8	77031	NM_148929	7.79	243	1.04	12061	1.20	11880
SLCO1A6	28254	NM_023718	7.56	251			1.89	5541
SLFN10	237887	XM_204665			1.03	13492	2.36	3810
SLFN2	20556	NM_011408	1.67	2227			2.87	2778
SLFN3	20557	NM_011409	26.27	45	0.90	38716	68.35	35
SLFN4	20558	NM_011410	2.31	1348	0.90	38716	25.33	79
SLFN5	327978	NM_183201			1.38	132	32.67	62
SLFN8-PENDING	276950	NM_181545	47.90	14	1.03	13492	18.30	119
SOCS2	216233	NM_007706	1.91	1762	1.66	87	1.55	7681
SP100	20684	NM_013673	3.16	910			7.99	424
SRI	109552	NM_025618	0.74	12248	1.33	168	1.05	15486
STAT1	20846	NM_009283	10.22	156	8.84	15	8.86	346
STAT2	20847	NM_019963	7.68	246	2.05	65	9.32	311
STK31	77485	NM_029916	34.50	29	0.94	34391	0.67	31055
STK32C	57740	NM_021302	8.52	219	1.02	16037	1.36	9615
STK4	58231	NM_021420	0.77	11133	1.33	165	1.05	15438
T2BP	211550	NM_145133	1.52	2614	1.10	3721	2.99	2608
TAP1	21354	NM_013683			2.23	61		
TAP2	21355	NM_011530	3.19	893	1.33	166	2.83	2837
TAPBP	21356	NM_009318	2.75	1087	1.43	118	2.59	3237
TBC1D10	103724	NM_134023	0.46	20250	1.37	138	0.77	26544
TCIRG1	27060	NM_016921	3.02	971	0.96	29776	1.18	12219
TCTE3	21647	NM_011560	6.59	302	1.08	5064	1.38	9349
TDRD3	219249	NM_172605	0.77	11279	0.99	22931	0.86	22190
TDRD7	100121	NM_146142	3.61	756	1.96	69	3.95	1642
TGTP	21822	NM_011579	73.70	6			137.20	19
THRSP	21835	NM_009381	8.90	195	1.03	13434	3.40	2084
TIPARP	99929	NM_178892	1.81	1932	1.07	6934	4.41	1356
TLR3	142980	NM_126166	45.49	18	1.01	17230	13.47	173
TM7SF1	83924	NM_031999	11.00	130	1.04	11872	0.71	29326
TNC	21923	NM_011607	0.79	10575	1.67	86	2.48	3504
TNFAIP3	21929	NM_009397	1.42	2982	1.05	8693	12.61	192
Tnfrsf5	21939	NM_170704	3.16	907	1.16	1121	43.58	51
TNFRSF6	14102	NM_007987	2.69	1115	1.32	175	3.09	2476
TNP1	21958	NM_009407	6.30	320	0.93	34632	4.72	1194
TOR3A	30935	NM_023141	4.79	496	1.04	11996	10.30	267
TRA1	22027	NM_011631	0.79	10679	1.31	191	1.03	15858
TREX1	22040	NM_011637	3.60	758	0.96	29420	9.94	281
TRIM14	74735	NM_029077			1.93	71	2.51	3420
TRIM21	20821	NM_009277	7.24	264	2.48	53	5.75	827
TRIM25	217069	NM_009546	6.08	340	2.55	51	7.39	505
TRIM26	22670	NM_030698			1.10	3057		
TRIM27	19720	NM_009054	8.66	211	1.07	6750	2.54	3335
TRIM30	20128	NM_009099	19.61	69			8.39	394
TRIM34	94094	NM_030684	7.07	278			7.11	557
TYKI	22169	NM_020557	97.21	4	11.18	10	601.82	3
UBE1L	74153	NM_023738	10.29	151	1.06	8115	5.60	875
UBE2L6	56791	NM_019949	2.76	1082	2.30	60	4.66	1222
UNC93B	54445	NM_019449	0.74	12374	1.31	181	1.01	16635
USP18	24110	NM_011909	26.10	46	2.16	64	57.69	39
USP25	30940	NM_013918	2.87	1036	1.10	3609	3.77	1770
V1RD2	81016	NM_030741	11.20	127			1.05	15493
VDAC1	22333	NM_011694	9.61	169	1.03	13065	0.69	30154
VNN1	22361	NM_011704	1.09	5068	0.84	41093	3.17	2362
WBSCR21	68758	NM_145215			1.34	161	0.71	29622
XMR	22526	NM_009529	19.17	74			1.74	6370
XRN1	24127	NM_011916	2.06	1574	1.25	320	2.42	3635
ZBP1	58203	NM_021394	45.37	19	1.09	4104	73.17	31
ZC3HAV1	78781	NM_028864			2.59	50	2.40	3710
ZC3HDC1	243771	NM_172893			0.97	28045		
ZFP313	81018	NM_030743	2.89	1023	1.07	6021	3.12	2432
ZFP99	67235	NM_023322	6.16	334	1.04	11518	0.55	35941
Znfx1	98999		3.08	948	1.20	585	15.10	156

the CpG/TLR9-dependent response, classical Dendritic Cells (cDCs) were stimulated with ISD or CpG[44]. Four hours following stimulation, microarray analysis was performed to identify the gene-expression programs activated with each response. As a further confirmation of DNA-directed antiviral response, only ISD specifically induced type I IFNs. Amongst these, we selected 124 genes with greater than 5.1-fold increases relative to unstimulated cells (Gene Expression Omnibus database: submission # GSE2197). The resulting list of 355 candidate genes represents the intersection of genes upregulated following stimulation with IFN-β (in NIH3T3 and L929 cells) or dsDNA (B6 MEFs and cDCs) and includes many uncharacterized ISGs.

3.3 – *Candidate gene selection: DNA SILAC*

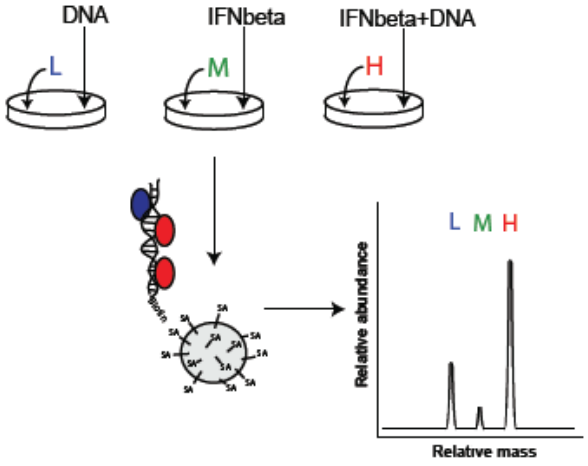
Next, we selected an additional 156 candidates from our own mass spectrometry-based list of putative STING-interacting proteins[140]. To directly identify DNA sensors and their binding partners, we used biotinylated DNA to pull down cytosolic binding partners through comparative proteomic screens with stable isotope labeling by amino acids in cell culture (SILAC). SILAC, a method developed by a scientist at the Broad Proteomics Platform, relies on the incorporation of amino acids with substituted stable isotopic nuclei[141, 142]. We utilized this three-state SILAC method to label and quantitate peptides via mass spectrometry, with medium isotope-labeled cells used for a negative control (beads alone), light isotope-labeled cells for bead-DNA precipitation, and heavy isotope-labeled cells for bead-DNA precipitation preceded by IFN-β stimulation to upregulate pathway components (Figure 3.2A).

We identified 184 proteins with SILAC ratios that showed enrichment for DNA binding following mass spectrometry (Figure 3.2B). Among the 184 identified proteins, 121 (64.2%) were classified by Gene Ontology as having nucleic acid binding function (*P*

Figure 3.2

Quantitative mass spectrometry identifies known components of the ISD sensing pathway. A) Schematic of DNA-interacting SILAC experiments. MEFs were labeled with light (L)-, medium (M)-, or heavy (H)-isotope SILAC solutions. Cells were pre-treated with Ifnβ or left unstimulated. Cytoplasmic extracts were prepared and incubated with or without biotinylated ISD. ISD was precipitated with streptavidin beads, and precipitated proteins were trypsinized and subjected to mass spectrometry. B) Quantitative mass spectrometry analysis showing DNA-binding proteins precipitated from cytoplasmic extracts of MEFs; proteins were precipitated with biotinylated DNA immobilized on streptavidin beads with streptavidin beads alone used as a negative control. DNA-interacting proteins with colors signify corresponding to pathways; white (all significant proteins) red (Aim2 inflammasome), purple (proteins encoded by genes mutated in AGS), blue (RNA polymerase III complex), green (HMGB proteins) and orange (SET complex) circles, DNA-interacting proteins with colors corresponding to pathways in c; yellow dots, nonsignificant precipitated proteins; A, abundance; H, M and L, isotope-labeled samples. Ratio of DNA-binding (DNA pull-down, +Ifnβ; AH) to bead-binding (empty bead pull-down, +Ifnβ; AM) per protein on the x axis is plotted against ratio of DNA binding with Ifnβ prestimulation (DNA pull-down, +Ifnβ; AH) to DNA binding without Ifnβ prestimulation (DNA pull-down, –Ifnβ; AL) per protein on the y axis. Full SILAC results, inset, lower right.

A



B

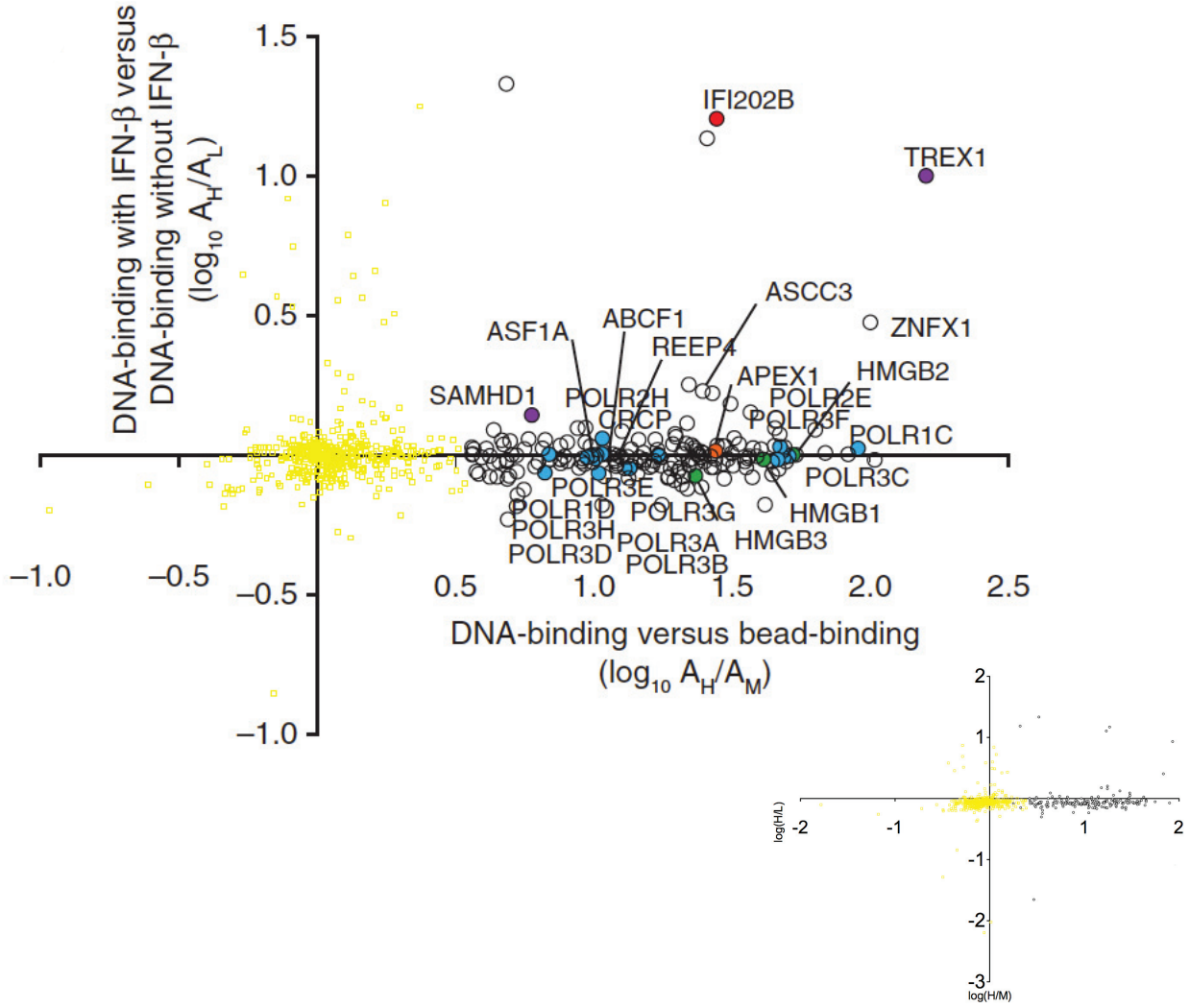


Figure 3.2 (continued)

= 5.95x10^{-58}; GO:0003676), and others were components of known DNA-binding complexes (Table 3.2).

Of the identified proteins, 20 (10.9%) represent the majority of known players involved in the immune sensing of cytosolic DNA. We identified known components of DNA sensing pathways including: the HMGB family proteins (HMGB1, HMGB2, HMGB3)[143], components of the AIM2 inflammasome (IFI202B and the HMGB proteins)[49, 143], and the cytosolic RNA polymerase III complex (POLR3A, POLR3B, POLR3C, POLR3D, POLR3E, POLR3F, POLR3G, POLR3H, POLR1C, POLR1D, POLR2E, POLR2H, and CRCP). Additionally, we identified three members of the SET complex (TREX1, APEX1, and HMGB2) that regulate the ISD pathway as well as HIV-1 detection and infection[104, 106, 143, 144]. We also identified associated proteins responsible for the autoimmune disease, Aicardi-Goutières syndrome (SAMHD1 and TREX1)[100, 145], which are involved in regulating retroviral and retroelement detection[104, 146]. Our findings validate the utility of quantitative mass spectrometry as an approach to find candidate components of cytosolic DNA sensing pathways. Of the 184 identified protein SILAC candidates, we found 156 matching siRNAs for inclusion in our arrayed screen.

3.4 – *Candidate gene selection: helicases*

The role of helicases in nucleic acid sensing is well established[29, 68, 71, 73, 147, 148]. The discovery that TLR3 and TLR7 deficient animals are able to produce type I IFN in response to RNA virus infection led to investigations to find TLR-independent mechanisms of viral-RNA mediated immune signaling, the RIG-I-like receptor family of helicases (RLRs)[2, 149, 150]. A key finding was the identification of three DExD/H box RNA helicases, retinoic acid-inducible gene-I (RIG-I), melanoma differentiation-associated gene 5 (MDA5), and LGP2, instrumental in mediating viral responses[29]. RIG-I and

Table 3.2
Candidate Gene List: DNA SILAC

Gene Name	GeneID	log(H/M)	log(H/L)
1110007L15Rik	67604	0.938	-0.092
1810029B16Rik	66282	0.595	-0.104
2810432D09Rik	69961	0.698	-0.154
8430406I07Rik	74528	1.040	-0.068
8430410A17Rik	232210	1.089	-0.152
A630055G03Rik	223970	1.309	0.054
Abcf1	224742	0.879	-0.076
Alkbh2	231642	0.826	-0.077
Anp32b	67628	1.405	-0.067
Anxa4	11746	1.134	-0.005
Anxa5	11747	1.317	-0.058
Apex1	11792	1.149	-0.058
Aplf	72103	0.894	-0.064
Aptx	66408	1.288	-0.100
Ascc1	69090	0.979	-0.095
Ascc2	75452	1.243	0.160
Ascc3	77987	1.180	0.177
Asf1a	66403	0.959	-0.091
Asf1b	66929	1.232	-0.149
Bst2	69550	0.322	1.186
Ccdc124	234388	0.616	-0.128
Cep170	545389	0.547	-0.302
Cggbp1	106143	0.553	-0.134
Cirbp	12696	1.016	-0.073
Cnn2	12798	1.383	-0.025
Crcp	12909	0.968	-0.001
Creb1	12912	1.078	0.023
Creb3l1	26427	1.234	-0.019
Creb3l2	208647	0.559	-0.193
Csda	56449	0.540	-0.152
Dazap1	70248	1.147	0.031
Ddb1	13194	0.864	-0.073
Ddb2	107986	1.085	-0.026
Ddx47	67755	0.327	-0.149
Ddx49	234374	0.450	-0.077
Dek	110052	1.090	-0.038
Dhx36	72162	0.476	-0.005
Dnajc9	108671	1.481	-0.066
Dr1	13486	1.461	-0.027
Drap1	66556	1.651	-0.028
Drg1	13494	0.426	-0.037
E430004N04Rik	210757	0.467	-1.653
Eef1d	66656	1.107	-0.039
Eif2s1	13665	0.511	-0.087
Eif2s2	67204	0.518	-0.077
Eif2s3x	26905	0.539	-0.082
Eif5b	226982	0.855	-0.066
Fam49b	223601	0.421	0.002
Fbxo18	50755	0.524	-0.076
Fen1	14156	1.241	-0.047
Frg1	14300	0.669	-0.117
Gata4	14463	0.661	0.032
Gnb2l1	14694	1.291	-0.030
H1fx	243529	0.606	-0.139
Hist1h1a	80838	1.561	-0.096
Hist1h1b	56702	1.580	-0.085
Hist1h1c	50708	1.460	-0.079
Hist1h1e	50709	1.008	-0.073
Hmga1	15361	1.444	-0.033
Hmga2	15364	1.603	-0.029
Hmgb1	15289	1.403	-0.081
Hmgb2	97165	1.550	-0.067
Hmgb3	15354	1.337	-0.132
Hmgn1	15312	1.044	-0.155
Hnrnpa0	77134	1.080	-0.008
Hnrnpa3	229279	1.346	-0.036
Hnrnpd	11991	1.316	-0.039
Hnrnpu	51810	1.083	-0.082
Hnrpab	15384	1.478	-0.035
Ifi202b	26388	1.267	1.171
Ifit1	15957	0.520	1.334
Kif22	110033	1.327	-0.059
Kif2a	16563	0.444	-0.075
Ldha	16828	1.559	-0.038
Ldhb	16832	1.408	-0.145
Lig3	16882	0.900	-0.100
LOC100044068	100044068	1.232	1.105
Mdh1	17449	1.641	-0.110
Mdh2	17448	1.435	-0.054
Morf4l2	56397	0.674	-0.069
Mpg	268395	1.082	-0.006
Msh2	17685	0.486	-0.094
Msh3	17686	0.848	-0.087
Msh6	17688	1.047	-0.233
Msi2	76626	0.821	0.026
Mtap1b	17755	1.118	-0.115
Mtap4	17758	1.025	-0.044
Mybbp1a	18432	1.100	-0.069
Nagk	56174	1.136	-0.069
Neil1	72774	1.131	-0.165
Nfib	18028	0.885	-0.228
Nthl1	18207	0.824	-0.078
Nuak2	74137	1.151	-0.105
Obfc1	108689	0.770	-0.101
Ogg1	18294	1.341	-0.087
Parp1	11545	1.184	-0.134
Parp2	11546	1.270	-0.082
Parp3	235587	1.244	0.203
Pcna	18538	0.759	-0.027
Pdia3	14827	0.471	-0.029
Phf6	70998	1.008	-0.088
Pkm2	18746	1.216	-0.050
Pnkp	59047	1.320	-0.072
Polb	18970	1.149	-0.104
Pold2	18972	0.738	0.039
Pole3	59001	1.155	-0.081
Polr1c	20016	1.748	-0.052
Polr1d	20018	0.809	-0.057
Polr2e	66420	1.652	-0.033
Polr2h	245841	0.934	-0.060
Polr3a	218832	1.043	-0.071
Polr3b	70428	0.779	-0.065
Polr3c	74414	1.582	-0.053
Polr3d	67065	0.805	-0.063
Polr3e	26939	0.836	-0.067
Polr3f	70408	1.558	-0.080
Polr3g	67486	0.964	-0.098
Polr3h	78929	0.776	-0.130
Prkar2b	19088	0.431	-0.057
Prrx1	18933	1.479	0.092
Psip1	101739	1.193	-0.073
Pycr2	69051	1.054	-0.132
Rabggtb	19352	1.359	-0.065
Rbm28	68272	0.816	-0.020
Rbm39	170791	0.691	-0.029
Rbms1	56878	1.100	-0.008
Rbms2	56516	1.483	0.051
Rbms3	207181	1.108	-0.157
Rbpj	19664	1.523	-0.076
Recql	19691	0.977	-0.081
Reep3	28193	1.211	-0.143
Reep4	72549	1.041	-0.098
Rfc1	19687	1.035	-0.097
Rfc2	19718	0.966	-0.082
Rfc3	69263	0.873	-0.100
Rfc4	106344	1.150	-0.088
Rfc5	72151	1.000	-0.077
Rpa1	68275	1.214	-0.070
Rpa2	19891	0.985	-0.066
Rpa3	68240	0.709	-0.077
Rpp38	227522	0.486	-0.135
Rrbp1	81910	1.133	-0.053
Rsl1d1	66409	0.556	-0.245
Samhd1	56045	0.641	0.092
Serbp1	66870	0.716	-0.049
Sfrs2	20382	0.534	-0.056
Sfrs3	20383	0.530	-0.066
Sfrs7	225027	0.494	0.021
Skp1a	21402	0.428	-0.082
Smarcal1	54380	0.880	-0.085
Snrpa	53607	0.805	-0.091
Snx9	66616	0.755	-0.067
Srm	20810	0.475	-0.114
Ssbp1	381760	0.909	-0.029
Sub1	20024	1.651	-0.052
Tbp	21374	1.209	-0.054
Tcfe3	209446	1.513	-0.057
Tcfeb	21425	1.900	-0.072
Tdg	21665	1.437	-0.178
Tdp1	104884	1.007	-0.095
Tead1	21676	0.493	-0.005
Tfam	21780	1.323	-0.031
Thex1	67276	0.616	-0.082
Thoc4	21681	0.546	-0.053
Thyn1	77862	1.408	-0.089
Top1	21969	1.210	-0.072
Top2a	21973	1.298	-0.004
Trex1	22040	1.932	0.932
Trip4	56404	1.375	0.124
Trmt6	66926	0.831	-0.004
Trmt61a	328162	1.334	-0.114
Tsnax	53424	1.528	-0.072
Ttf2	74044	1.627	0.026
Ubtf	21429	0.999	-0.084
Usp39	28035	0.676	-0.081
Vrk1	22367	1.042	-0.041
Wdr76	241627	0.618	0.003
Wrn	22427	1.471	0.012
Xbp1	22433	1.241	-0.156
Xpa	22590	1.439	-0.145
Xrcc1	22594	1.396	-0.148
Xrcc5	22596	1.113	-0.103
Xrcc6	14375	1.276	-0.100
Ybx1	22608	0.250	-0.083
Zc3h15	69082	0.610	-0.053
Znfx1	98999	1.835	0.404
Zranb3	226409	1.274	-0.126

MDA5 helicases drive production of type I IFNs in all cell types in response to RNA virus infection, with the exception of TLR7 dependent pDCs[29, 32, 151]. LGP2 functions both as a negative regulator of the RIG-I/MDA5 pathway[152] and, more recently, as a possible co-receptor for some RIG-I and MDA5 ligands[33]. Additionally, the DExD/H box RNA helicase family member DDX3 was identified as another RLR[147, 148, 153]. It was reported that DDX3 binds both to transfected Poly I:C and viral RNA and associates with MAVS to induce type I IFN production. DDX3 has also been implicated in sensing cytomegalovirus whereby phosphorylated DDX3 binds to the IFN-β promoter following CMV infection[154].

The aspartate-glutamate-alanine-histidine box (DEAH)/RNA helicases DHX36 and DHX9 were recently identified as specific sensors for CpG-A and CpG-B, respectively[68]. In pDCs treated with CpG, DHX36 and DHX9 are localized in the cytosol and bind to Toll-IL receptor (TIR) domain of myeloid differentiation primary response gene 88 (MyD88) leading to activation of IRF7 and NF-κB. Subsequently, DHX36 was identified as TLR3/MDA5-independent sensor of Poly I:C, that, in a complex with DDX1 and DDX21, pairs with the adaptor TRIF to trigger type I IFN responses[71]. Furthermore, DHX9 was found to pair with MAVS to sense dsRNA in myeloid dendritic cells (MDCs). Additionally, the helicase DDX41 was identified as an intracellular DNA sensor in MDCs[73]. In a focused shRNA screen targeting 59 members of the DExD/H-box helicase family, DDX41 knockdown led to impaired type I IFN and pro-inflammatory cytokine production in response to various dsDNA stimuli. Co-immunoprecipitation experiments suggest interaction of DDX41 and STING in resting and stimulatory conditions. We therefore generated a comprehensive list of helicases via the PANTHER classification system (Table 3.3)[155-157]. We identified 174 proteins with known or predicted RNA and DNA helicase activity, 118 of which were present in the siRNA library.

Table 3.3
Candidate Gene List: Helicases

Gene	GeneID	Accession	Candidate List	Panther Molecular Function	Panther Biological Process
0610007P08RIK	76251	NM_023507	Helicase	DNA helicase	DNA repair;DNA recombination
1210002B07RIK	67997	NM_026500	Helicase	RNA helicase	Nucleoside, nucleotide and nucleic acid metabolism
1810014J18RIK	109151	XM_284439	Helicase	DNA helicase	mRNA transcription regulation
2310061O04RIK	69663	NM_027156	Helicase	RNA helicase	Nucleoside, nucleotide and nucleic acid metabolism
2610007K22RIK	67040	NM_199079	Helicase	RNA helicase	Nucleoside, nucleotide and nucleic acid metabolism
2610528A15RIK	72198	NM_028151	Helicase	DNA helicase;Hydrolase	Nucleoside, nucleotide and nucleic acid metabolism;Meiosis
2610528E23RIK	66497	NM_025599	Helicase	RNA helicase	Nucleoside, nucleotide and nucleic acid metabolism
2810457M08RIK	234733	NM_172284	Helicase	RNA helicase	Nucleoside, nucleotide and nucleic acid metabolism
4930422G04RIK	71643	NM_197997	Helicase	Nuclease;RNA helicase;Hydrolase	RNA catabolism
5430439G14RIK	71389	NM_173368	Helicase	DNA helicase	mRNA transcription regulation
6030422M02	240697	NM_177722	Helicase	DNA helicase;Hydrolase	DNA replication;DNA replication
6330505F04RIK	236790	NM_172779	Helicase	RNA helicase	Tumor suppressor
A330009G12RIK	330149	NM_177873	Helicase	DNA helicase;Hydrolase	Nucleoside, nucleotide and nucleic acid metabolism;Meiosis
A330064G03RIK	320632	NM_177214	Helicase	RNA helicase;mRNA splicing factor;Hydrolase	mRNA splicing
A930037J23RIK	269254	NM_177365	Helicase	Helicase	Muscle contraction
AI449441	208084	NM_172453	Helicase	DNA helicase	Biological process unclassified
ATRX	22589	NM_009530	Helicase	DNA helicase	mRNA transcription regulation
AW494914	106794	NM_198942	Helicase	RNA helicase	mRNA splicing
AW540478	269400	XM_194070	Helicase	DNA helicase	DNA repair
B830009D23RIK	96957	NM_175285	Helicase	RNA helicase	RNA catabolism
BAT1A	53817	NM_019693	Helicase	RNA helicase	Nucleoside, nucleotide and nucleic acid metabolism
BC004701	236930	NM_146235	Helicase	DNA helicase	DNA repair;DNA recombination
BC019206	216161	NM_183426	Helicase	Nuclease;Helicase	Nucleoside, nucleotide and nucleic acid metabolism;Developmental processes;Other metabolism
BLM	12144	NM_007550	Helicase	DNA helicase	DNA replication;DNA repair;DNA recombination;DNA replication
BRIP1	237911	NM_178309	Helicase	DNA helicase	DNA repair;Oncogene
C130058G22RIK	319955	NM_177043	Helicase	DNA helicase	DNA repair
CHD1	12648	NM_007690	Helicase	DNA helicase	mRNA transcription regulation
CHD1L	68058	NM_026539	Helicase	DNA helicase	mRNA transcription regulation
CHD2	244059	XM_145698	Helicase	DNA helicase	mRNA transcription regulation
CHD3	216848	NM_146019	Helicase	DNA helicase	mRNA transcription regulation
CHD4	107932	NM_145979	Helicase	DNA helicase	mRNA transcription regulation
Chd8	67772	NM_201637	Helicase	DNA helicase	mRNA transcription regulation
DDX1	104721	NM_134040	Helicase	RNA helicase	Nucleoside, nucleotide and nucleic acid metabolism
DDX11	320209	XM_128714	Helicase	DNA helicase	Chromosome segregation
DDX18	66942	NM_025860	Helicase	RNA helicase	Nucleoside, nucleotide and nucleic acid metabolism
DDX19	13680	NM_007916	Helicase	RNA helicase	Nucleoside, nucleotide and nucleic acid metabolism
DDX20	53975	NM_017397	Helicase	RNA helicase	Nucleoside, nucleotide and nucleic acid metabolism
DDX21	56200	NM_019553	Helicase	RNA helicase	Nucleoside, nucleotide and nucleic acid metabolism
DDX24	27225	NM_020494	Helicase	RNA helicase	Nucleoside, nucleotide and nucleic acid metabolism
DDX25	30959	NM_013932	Helicase	RNA helicase	Nucleoside, nucleotide and nucleic acid metabolism
DDX26	18130	NM_008715	Helicase	RNA helicase	Tumor suppressor
DDX27	228889	NM_153065	Helicase	RNA helicase activity;translation factor activity	nucleobase;nucleoside;nucleotide and nucleic acid metabolic process translation
DDX28	71986	NM_028038	Helicase	RNA helicase	Nucleoside, nucleotide and nucleic acid metabolism
DDX39	68278	NM_197982	Helicase	RNA helicase	Nucleoside, nucleotide and nucleic acid metabolism
DDX3X	13205	NM_010028	Helicase	RNA helicase	Nucleoside, nucleotide and nucleic acid metabolism
DDX41	72935	NM_134059	Helicase	RNA helicase	Nucleoside, nucleotide and nucleic acid metabolism
DDX42	72047	NM_028074	Helicase	RNA helicase	Nucleoside, nucleotide and nucleic acid metabolism
DDX46	212880	NM_145975	Helicase	RNA helicase	Nucleoside, nucleotide and nucleic acid metabolism
DDX48	192170	NM_138669	Helicase	RNA helicase;Translation initiation factor	Protein biosynthesis;Translational regulation
DDX5	13207	NM_007840	Helicase	RNA helicase	Nucleoside, nucleotide and nucleic acid metabolism
DDX50	94213	NM_053183	Helicase	RNA helicase	Nucleoside, nucleotide and nucleic acid metabolism
DDX52	78394	NM_030096	Helicase	RNA helicase	Nucleoside, nucleotide and nucleic acid metabolism
DDX54	71990	NM_028041	Helicase	RNA helicase	Nucleoside, nucleotide and nucleic acid metabolism
DDX55	67848	NM_026409	Helicase	RNA helicase	Nucleoside, nucleotide and nucleic acid metabolism
DDX56	52513	NM_026538	Helicase	RNA helicase	Nucleoside, nucleotide and nucleic acid metabolism

Table 3.3
Candidate Gene List: Helicases (continued)

Gene	GeneID	Accession	Candidate List	Panther Molecular Function	Panther Biological Process
DDX6	13209	NM_007841	Helicase	RNA helicase	Nucleoside, nucleotide and nucleic acid metabolism
DHX15	13204	NM_007839	Helicase	RNA helicase	mRNA splicing
DHX16	69192	NM_026987	Helicase	RNA helicase	mRNA splicing
DHX29	218629	NM_172594	Helicase	RNA helicase	mRNA splicing
DHX30	72831	NM_133347	Helicase	RNA helicase	mRNA splicing
DHX32	101437	NM_133941	Helicase	RNA helicase	mRNA splicing
DHX33	216877	NM_178367	Helicase	RNA helicase	mRNA splicing
DHX34	71723	NM_027883	Helicase	RNA helicase	mRNA splicing
DHX35	71715	NM_145742	Helicase	RNA helicase	mRNA splicing
DHX37	208144	NM_203319	Helicase	RNA helicase	mRNA splicing
DHX38	64340	NM_178380	Helicase	RNA helicase	mRNA splicing
DHX40	67487	NM_026191	Helicase	RNA helicase	mRNA splicing
DHX8	217207	NM_144831	Helicase	RNA helicase	mRNA splicing
DHX9	13211	NM_007842	Helicase	RNA helicase	mRNA splicing
DQX1	93838	NM_033606	Helicase	RNA helicase	mRNA splicing
E130016E03RIK	623474	NM_177285	Helicase	DNA helicase	DNA repair;DNA recombination
E130315B21RIK	327762	NM_177372	Helicase	DNA helicase;Hydrolase	DNA replication;DNA replication
E430027O22RIK	107182	XM_129248	Helicase	DNA helicase	mRNA transcription regulation
EIF4A1	13681	NM_144958	Helicase	RNA helicase;Translation initiation factor	Protein biosynthesis;Translational regulation
EIF4A2	13682	NM_013506	Helicase	RNA helicase;Translation initiation factor	Protein biosynthesis;Translational regulation
EP400	75560	NM_029337	Helicase	DNA helicase	mRNA transcription regulation
ERCC2	13871	NM_007949	Helicase	DNA helicase	DNA repair;mRNA transcription;Oncogenesis
ERCC3	13872	NM_133658	Helicase	DNA helicase;Hydrolase	DNA repair
FIN14	13205	NM_008015	Helicase	RNA helicase	Nucleoside, nucleotide and nucleic acid metabolism
HELLS	15201	NM_008234	Helicase	DNA helicase	mRNA transcription regulation
IGHMBP2	20589	NM_009212	Helicase	DNA helicase;Single-stranded DNA-binding protein	DNA metabolism
MCM2	17216	NM_008564	Helicase	DNA helicase;Hydrolase	DNA replication;DNA replication;Cell proliferation and differentiation
MCM3	17215	NM_008563	Helicase	DNA helicase;Hydrolase	DNA replication;DNA replication
MCM4	17217	NM_008565	Helicase	DNA helicase;Hydrolase	DNA replication;DNA replication
MCM5	17218	NM_008566	Helicase	DNA helicase;Hydrolase	DNA replication;DNA replication
MCM6	17219	NM_008567	Helicase	DNA helicase;Hydrolase	DNA replication;DNA replication
MCM7	17220	NM_008568	Helicase	DNA helicase;Hydrolase	DNA replication;DNA replication
MCM8	66634	NM_025676	Helicase	DNA helicase;Hydrolase	DNA replication;DNA replication
MOV10L1	83456	NM_031260	Helicase	Helicase	Biological process unclassified
PEO1	226153	NM_153796	Helicase	DNA helicase	DNA replication;DNA replication
POLQ	77782	NM_029977	Helicase	DNA helicase;Hydrolase	Nucleoside, nucleotide and nucleic acid metabolism;Meiosis
RAD54L	19366	NM_009015	Helicase	DNA helicase	DNA repair;DNA recombination
RECQL4	79456	NM_058214	Helicase	DNA helicase	DNA repair
RECQL5	170472	NM_130454	Helicase	DNA helicase	DNA metabolism
RENT1	19704	NM_030680	Helicase	RNA helicase;Hydrolase	RNA catabolism
RUVBL1	56505	NM_019685	Helicase	Transcription cofactor;DNA helicase	mRNA transcription regulation;Embryogenesis;Mesoderm development
RUVBL2	20174	NM_011304	Helicase	Transcription cofactor;DNA helicase	mRNA transcription regulation;Embryogenesis;Mesoderm development
SHPRH	268281	NM_172937	Helicase	DNA helicase	mRNA transcription regulation
SKIV2L	108077	NM_021337	Helicase	DNA helicase;Hydrolase	Nucleoside, nucleotide and nucleic acid metabolism;Meiosis
SMARCA1	93761	NM_053123	Helicase	DNA helicase	mRNA transcription regulation
SMARCA2	67155	NM_011416	Helicase	DNA helicase	mRNA transcription regulation
SMARCA3	20585	NM_009210	Helicase	DNA helicase	mRNA transcription regulation
SMARCA4	20586	NM_011417	Helicase	DNA helicase	mRNA transcription regulation
SMARCA5	93762	NM_053124	Helicase	DNA helicase	mRNA transcription regulation
SMARCAD1	13990	XM_132597	Helicase	DNA helicase	mRNA transcription regulation;Chromatin packaging and remodeling
SRISNF2L	81000	NM_030730	Helicase	DNA helicase	mRNA transcription regulation
SUPV3L1	338359	NM_181423	Helicase	RNA helicase	Nucleoside, nucleotide and nucleic acid metabolism

3.5 – *Candidate gene selection: cytoplasmic DNA-binding proteins*

In an extension of our hypothesis that any of the annotated nucleases could have a role in dsDNA sensing we subsequently curated a list of proteins that had known DNA binding properties and were putatively localized to the cytoplasm. First, 842 nucleic-acid binding proteins were identified with the PANTHER classification system, 174 of which were the previously described helicases. The balance of the list included 25 ssDNA binding proteins, 197 nucleases, 30 dsDNA binding proteins, 321 proteins with other DNA binding designations and 108 proteins with uncharacterized nucleic acid binding properties. To assess protein localization, we first cross-referenced our list with annotated databases and secondly, assessed localization with published predictive algorithms. We utilized three annotated databases that provide curated cellular localization data, UniPROT, LOCATE and TFCat. LOCATE is a database that houses data describing the membrane organization and subcellular localization of proteins derived from a genome-wide mouse proteome study[158, 159]. TFCat is a curated catalog of mouse and human transcription factors (TF) based on a core collection of annotations obtained by review of the scientific literature[160]. Annotated genes are assigned to a functional category and confidence level. Together, the databases indicated cytoplasmic location data for 17% of the identified nucleic acid binding proteins.

To complement the curated database, we assessed cellular location utilizing two protein subcellular localization prediction algorithms, CELLO and WoLF PSORT[161-163]. CELLO, or Subcellular Localization, utilizes a supervised learning model, or support vector machine (SVM)-based classification system. CELLO uses four types of sequence coding schemes: amino acid composition, dipeptide composition, partitioned amino acid composition and the sequence composition based on the physicochemical properties of amino acids. Combined votes from these classifiers and SVM-jury votes are used to

determine the final assignment. Alternatively, WoLF PSORT converts protein amino acid sequences into numerical localization features based on sorting signals, amino acid composition and functional motifs such as DNA-binding motifs. After conversion, the k-nearest neighbor algorithm, a pattern-recognition method for classifying objects based on closest training examples, is applied to make predictions.

We assessed the utility of the protein prediction methods and annotated databases by selecting 20 proteins with known subcellular localization, 15 of which are cytosolic (Table 3.4). While the annotated databases agreed with the known prediction more than 90 percent of the time, the protein localization algorithm methods predicted the correct cellular compartment as one of the top two cellular compartments roughly 55 percent of the time. In addition to the 78 proteins identified through the annotated databases, we added another 25 proteins based on a composite score of the prediction algorithms (Table 3.5). First, proteins that were predicted to be cytoplasmic by both localization predictions were included. Second, we added proteins that were predicted with high confidence by one algorithm as the most likely cytoplasmically localized (greater than 60% confidence) and proteins for which cytoplasmic localization was the second predicted compartment. The resulting list of 103 proteins includes candidates with annotated or predicted cellular localization available in the siRNA library.

3.6 – *Candidate gene selection: putative negative regulators and signaling molecules*

Finally, we added annotated phosphatases and deubiquitinases as part of a pilot screen to identify potential regulators of the ISD pathway. Activation of signaling molecules critical to the ISD pathway requires phosphorylation and ubiquitination[60, 139, 164]. For example, TBK1, directly phosphorylates IRF3, and thus mice deficient in Tbk1 failed to induce type I IFN following stimulation with Poly (dA:dT)[43]. The E3 ubiquitin

Table 3.4

Cytoplasmic protein prediction matrix. Subcellular localization prediction algorithms (CELLO and W_PSORT) are assessed on the indicated genes along with annotated localization data from UniProt and LOCATE databases. The top two cellular compartments are presented for each prediction tool. CELLO values represent the composite prediction fraction of five possible compartments. Wolf PSORT values roughly indicate the number of nearest neighbors to the query that localize to each site, adjusted to account for the possibility of dual localization. Green boxes agree with known protein localizations. Predictive accuracy is presented as percentage. Cytoplasm = cyt. Nucleus = nuc. Mitochondria outer membrane = mit out mem. Endoplasmic reticulum = ER. Nuclear envelope = nuc env. Plasma membrane = plasm. Extracellular = extr.

gene_name	gene_id	actual	UniProt	LOCATE	CELLO1	CELLO2	W_PSORT1	W_PSORT2
ddx58	230073	cyt	cyt	cyt	cyt:2.448	nuc:1.985	cyt:24	nuc:3
ifih1	71586	cyt	cyt. Nuc.	cyt	nuc:3.355	nuc:1.089	cyt:25.5	cyt_nucl:13.5
Trim25	217069	cyt	-		nuc:2.385	extr:1.416	nucl:29.5	cyt_nucl:16.5
TBK1	56480	cyt	cyt		cyt:2.860	nuc:1.064	cyt:18	cyt_nucl:12
IKBKE	56489	cyt	cyt	cyt	nuc:2.144	cyt:1.484	nucl:19	cyto_nucl:17
ticam1	106759	cyt	-		nuc:3.662	plasm:0.659	nucl:15	cyto:9.5
MyD88	17874	cyt	cyt	cyt	nuc:2.604	cyt:0.841	cyto:14	mito:10
Nod1	107607	cyt	cyt	cyt	nuc:2.002	extr:1.218	cyto_nucl:10	nucl:9.5
Nod2	257632	cyt	cyt	cyt	plasm:1.703	extr:1.084	cyto_nucl:10.2	cyto:10
nlrp3	216799	cyt	cyt	cyt	nuc:2.329	extr:1.463	cyto_nucl:10.8	nucl:10.5
mavs	228607	mit out mem	mit out mem		nucl:3.373	extr:0.647	extr:7.5	extr_plas:6.5
gapdh	14433	cyt	cyt	cyt	cyt:3.993	mito:0.418	cyto:16.5	mito:13
trex1	22040	cyt	nuc	er, nuc env	nucl:2.742	extr:1.667	cyto:8	extr:7.5
zbp1	58203	cyt	-	cyt	nuc:3.961	cyt:0.732	cysk:15	nucl:15
aim2	383619	cyt	nuc		cyt:2.009	nucl:1.958	cyto:15	nucl:14
ifi204	15951	nuc	nuc. Cyto	nucleolus	nuc:3.230	mit:0.713	cyto_nucl:16	nucl:15
ifi205	226695	nuc	nuc		nucl:4.192	cyt:0.372	nucl:21	cyto:8
rela	19697	cyt. Nuc.	nuc. Cyto.		nuc:4.610	plasm:0.124	nucl:28	cyto_nucl:18
myc	17869	nuc	nuc	nucleus	nuc:4.826	cyt:0.103	nucl:32	-
smarcb1	20587	nuc	nuc	nucl, cyt	nuc:2.952	cyt:0.747	cyt:18	cyt_nucl:15
			93.75%	**100%**	**55.00%**		**60.00%**	

Table 3.5

Candidate Gene List: Cytoplasmic proteins

Gene	GeneID	Accession	Candidate List	CELLO Organelle	CELLO Rank	Wolfpsort Organelle	Wolfpsort Rank
0610010I17RIK	66847	NM_025798	Cytoplasmic - Uniprot	nuc	2.67	nucl	32
1700051E09RIK	67338	NM_026097	Cytoplasmic - Uniprot	nuc	4.77	nucl	32
2310005K03RIK	69537	NM_027109	Cytoplasmic - Predicted	extra	1.93	cyto	20
2410006F12RIK	71957	NM_028020	Cytoplasmic - Uniprot	nuc	2.94	nucl	19.5
2810028N01RIK	72662	XM_127907	Cytoplasmic - Uniprot	cytop	2.50	cyto	13
4930517K23RIK	403171	NM_207275	Cytoplasmic - Uniprot	extra	2.66	cyto_nucl	12.5
4932442K20RIK	231464	NM_144910	Cytoplasmic - Uniprot	cytop	1.99	nucl	29.5
4933406L09RIK	74430	NM_028934	Cytoplasmic - Locate	nuc	3.89	nucl	32
5830483C08RIK	209334	NM_177331	Cytoplasmic - Predicted	nuc	4.00	cyto	21.5
A230103N10RIK	104625	NM_212484	Cytoplasmic - Uniprot	cytop	1.65	nucl	22
ANG4	219033	NM_177544	Cytoplasmic - Predicted	cytop	3.18	cyto	22.5
ANKRD3	72388	NM_023663	Cytoplasmic - Uniprot TFCat				
ARC	11838	NM_018790	Cytoplasmic - Uniprot TFCat				
BAT2	53761	NM_020027	Cytoplasmic - Uniprot	nuc	3.79	cyto_nucl	18
BC034753	234258	NM_146208	Cytoplasmic - Predicted	nuc	4.85	cyto	26
BC052360	231999	NM_001001335	Cytoplasmic - Uniprot				
BIN1	30948	NM_009668	Cytoplasmic - Uniprot TFCat				
BZW1	66882	NM_025824	Cytoplasmic - Predicted	cytop	3.39	nucl	13.5
BZW2	66912	NM_025840	Cytoplasmic - Predicted	cytop	2.95	nucl	32
CDC5L	71702	NM_152810	Cytoplasmic - Uniprot TFCat				
CSDA	56449	NM_011733	Cytoplasmic - Uniprot TFCat				
D11ERTD497E	52626	NM_029976	Cytoplasmic - Predicted	cytop	2.47	nucl	32
D630024B06RIK	218973	NM_172598	Cytoplasmic - Predicted	nuc	4.01	cyto	19
D7WSU87E	360216	NM_207302	Cytoplasmic - Uniprot	nuc	4.37	nucl	23
DFFB	13368	NM_007859	Cytoplasmic - Uniprot	nuc	2.63	nucl	25.5
DPF2	19708	NM_011262	Cytoplasmic - Uniprot TFCat				
DSIP1	14605	NM_010286	Cytoplasmic - Locate TFCat				
EAR1	13586	NM_007894	Cytoplasmic - Uniprot	nuc	4.71	nucl	31
EAR2	13587	NM_007895	Cytoplasmic - Uniprot	extra	4.08	cyto	17.5
EAR3	53876	NM_017388	Cytoplasmic - Predicted	extra	4.05	cyto_nucl	18.8
EAR4	53877	NM_017389	Cytoplasmic - Uniprot	nuc	4.28	nucl	31.5
ENDOG	13804	NM_007931	Cytoplasmic - Predicted	nuc	4.42	cyto	25.5
EPS15	13858	NM_007943	Cytoplasmic - Uniprot TFCat				
ERCC1	13870	NM_007948	Cytoplasmic - Locate	plas	2.74	nucl	32
EXOSC2	227715	NM_144886	Cytoplasmic - Uniprot	cytop	1.73	cyto	21.5
EXOSC3	66362	NM_025513	Cytoplasmic - Uniprot	cytop	1.94	nucl	29.5
EXOSC4	109075	NM_175399	Cytoplasmic - Uniprot	nuc	4.08	nucl	32
EXOSC5	27998	NM_138586	Cytoplasmic - Locate	cytop	1.81	nucl	25
EXOSC6	72544	NM_028274	Cytoplasmic - Uniprot	nuc	1.90	nucl	27
EXOSC8	69639	NM_027148	Cytoplasmic - Locate	nuc	2.54	extr	15
EXOSC9	50911	NM_019393	Cytoplasmic - Uniprot	cytop	2.57	nucl	29.5
HIC2	58180	NM_178922	Cytoplasmic - Predicted	cytop	2.84	nucl	32
JARID1D	20592	NM_011419	Cytoplasmic - Locate TFCat				
JTV1	231872	NM_146165	Cytoplasmic - Predicted	nuc	3.26	cyto	20.5
LIN28	83557	NM_145833	Cytoplasmic - Uniprot	nuc	3.69	mito	20
LSM8	76522	NM_133939	Cytoplasmic - Locate	nuc	2.20	nucl	26.5
MAPK14	26416	NM_011951	Cytoplasmic - Uniprot TFCat				
MAPK7	23939	NM_011841	Cytoplasmic - Uniprot TFCat				
MATR3	17184	NM_010771	Cytoplasmic - Predicted	nuc	4.12	cyto	21
MBD1	17190	NM_013594	Cytoplasmic - Predicted	nuc	4.52	cyto	18
MBD3L1	73503	NM_028557	Cytoplasmic - Locate	nuc	3.71	nucl	14.5
MBD4	17193	NM_010774	Cytoplasmic - Predicted	nuc	2.32	cyto	19.5
MBNL2	105559	NM_175341	Cytoplasmic - Uniprot	nuc	3.39	nucl	24
MLH1	17350	NM_026810	Cytoplasmic - Predicted	nuc	4.42	cyto	18
MRE11A	17535	NM_018736	Cytoplasmic - Predicted	cytop	3.18	cyto	22
MRPS28	66230	NM_025434	Cytoplasmic - Predicted	nuc	2.72	cyto	20
NANOS1	332397	NM_178421	Cytoplasmic - Uniprot	nuc	4.43	nucl	32
NDEL1	83431	NM_023668	Cytoplasmic - Uniprot TFCat				
Ndn	17984	NM_010882	Cytoplasmic - Uniprot	nuc	3.91	nucl	32
NFKBIA	18035	NM_010907	Cytoplasmic - Uniprot TFCat				
NFKBIE	18037	NM_008690	Cytoplasmic - Uniprot TFCat				
Otud7a	170711	NM_130880	Cytoplasmic - Uniprot	nuc	4.83	cyto	20
PA2G4	18813	NM_011119	Cytoplasmic - Uniprot				
PAWR	114774	XM_125814	Cytoplasmic - Uniprot TFCat				
PCBP3	59093	NM_021568	Cytoplasmic - Uniprot TFCat				
PELO	105083	NM_134058	Cytoplasmic - Uniprot	cytop	1.76	cyto	29
PER3	18628	NM_011067	Cytoplasmic - Uniprot TFCat				
PLEKHA3	83435	NM_031256	Cytoplasmic - Uniprot				
POGK	71592	NM_175170	Cytoplasmic - Locate				
PPP1R13B	21981	NM_011625	Cytoplasmic - Uniprot TFCat				
Ptrf	19285	NM_008986	Cytoplasmic - Uniprot TFCat				
PURG	75029	NM_152821	Cytoplasmic - Predicted	nuc	3.12	cyto	18
RABGEF1	56715	NM_019983	Cytoplasmic - Uniprot TFCat				
RAD1	19355	NM_011232	Cytoplasmic - Locate	nuc	3.52	nucl	32
REX3	19716	NM_009052	Cytoplasmic - Uniprot TFCat				
ROCK2	19878	NM_009072	Cytoplasmic - Locate TFCat				
RPP21	67676	NM_026308	Cytoplasmic - Predicted	cytop	3.25	mito	16
RXRG	20183	NM_009107	Cytoplasmic - Locate TFCat				
SART3	53890	NM_016926	Cytoplasmic - Uniprot	nuc	2.10	nucl	32
SBDS	66711	NM_023248	Cytoplasmic - Uniprot	nuc	3.51	nucl	28
SIRT2	64383	NM_022432	Cytoplasmic - Uniprot TFCat				
SNAPC3	77634	NM_029949	Cytoplasmic - Locate	nuc	2.90	nucl	32
SSBP2	66970	NM_024186	Cytoplasmic - Predicted	cytop	2.03	cyto	14.5
STATIP1	58523	NM_021448	Cytoplasmic - Uniprot TFCat				
SWAP70	20947	NM_009302	Cytoplasmic - Uniprot				
TCEB2	67673	NM_026305	Cytoplasmic - Locate TFCat				
TDRD1	83561	NM_031387	Cytoplasmic - Uniprot	nuc	4.04	nucl	32
THAP11	59016	NM_021513	Cytoplasmic - Uniprot	cytop	2.22	nucl	31.5
TRERF1	224829	NM_172622	Cytoplasmic - Locate	nuc	4.80	nucl	32
TSN	22099	NM_011650	Cytoplasmic - Uniprot	nuc	3.84	cyto	11
TTF1	22130	NM_009442	Cytoplasmic - Predicted	cytop	3.25	nucl	32
TXK	22165	NM_013698	Cytoplasmic - Uniprot TFCat				
USP52	103135	NM_133992	Cytoplasmic - Uniprot	nuc	4.49	nucl	23
XAB2	67439	NM_026156	Cytoplasmic - Predicted	cytop	2.46	nucl	30
XRN1	24127	NM_011916	Cytoplasmic - Uniprot	nuc	4.42	nucl	32
ZFP143	20841	NM_009281	Cytoplasmic - Predicted	cytop	2.43	nucl	23
ZFP148	22661	NM_011749	Cytoplasmic - Predicted	cytop	2.35	cyto	11
ZFP259	22687	NM_011752	Cytoplasmic - Uniprot TFCat				
ZFP281	226442	NM_177643	Cytoplasmic - Locate	nuc	4.35	E.R.	8.5
ZFP346	26919	NM_012017	Cytoplasmic - Uniprot	nuc	4.80	nucl	32
ZFP36L1	12192	NM_007564	Cytoplasmic - Locate TFCat				
ZFP521	225207	NM_145492	Cytoplasmic - Locate TFCat				
ZFR	22763	NM_011767	Cytoplasmic - Uniprot	nuc	4.34	nucl	32

Table 3.5
Candidate Gene List: Cytoplasmic proteins (continued)

Gene	Uniprot Subcellular locations	Panther Biological Process	Panther Molecular Function
0610010I17RIK	Cytoplasm. Nucleus.	DNA repair	Damaged DNA-binding protei
1700051E09RIK	Cytoplasm perinuclear region. Membrane; Peripheral membrane protein.	Proteolysis	Other DNA-binding protein;Ub
2310005K03RIK	Endoplasmic reticulum.	DNA degradation	Endodeoxyribonuclease;Hydr
2410006F12RIK	Nucleus. Cytoplasm.	mRNA polyadenylation;mRNA end-processing and stability	Endoribonuclease;mRNA poly
2810028N01RIK	Cytoplasm. Nucleus nucleolus.	RNA catabolism;Mitosis	Exoribonuclease;Hydrolase
4930517K23RIK	Nucleus. Cytoplasm.	Biological process unclassified	Nuclease
4932442K20RIK	Cytoplasm.	mRNA transcription regulation	Exoribonuclease
4933406L09RIK		Biological process unclassified	Transcription factor;Nuclease
5830483C08RIK	Nucleus.	Biological process unclassified	Endodeoxyribonuclease
A230103N10RIK	Cytoplasm. Nucleus.	mRNA transcription regulation	Exoribonuclease
ANG4		RNA catabolism;Angiogenesis	Endoribonuclease;Other enzy
ANKRD3		Protein phosphorylation	Non-receptor serine/threonin
ARC		Biological process unclassified	Molecular function unclassifie
BAT2	Cytoplasm. Nucleus.	Nucleoside, nucleotide and nucleic acid metabolism;Other metabolism	Transcription factor;Nuclease
BC034753	Nucleus.	DNA repair;Other metabolism	Endodeoxyribonuclease;DNA
BC052360	Cytoplasm.	Biological process unclassified	Other nucleic acid binding;Tra
BIN1		Endocytosis;Transport;Neurotransmitter release	Membrane traffic regulatory p
BZW1		Protein biosynthesis;Translational regulation;Other protein metabolism	Nuclease;Translation initiatior
BZW2		Protein biosynthesis;Translational regulation;Other protein metabolism	Nuclease;Translation initiatior
CDC5L		mRNA transcription regulation;Cell cycle control	Other transcription factor;Nuc
CSDA		mRNA transcription regulation	Other transcription factor;Nuc
D11ERTD497E		Other mRNA transcription	Other DNA-binding protein
D630024B06RIK	Nucleus nucleoplasm.	mRNA transcription regulation	Other DNA-binding protein
D7WSU87E	Cytoplasm. Nucleus.	Proteolysis;Other protein metabolism	Double-stranded DNA binding
DFFB	Cytoplasm. Nucleus.	DNA degradation;Apoptotic processes	Nuclease
DPF2		mRNA transcription regulation;Protein acetylation;Induction of apoptosis;Develo	Zinc finger transcription facto
DSIP1		mRNA transcription regulation	Other transcription factor
EAR1	Cytoplasmic granule.	RNA catabolism	Endoribonuclease;Hydrolase
EAR2	Cytoplasmic granule.	RNA catabolism	Endoribonuclease;Hydrolase
EAR3		RNA catabolism	Endoribonuclease;Hydrolase
EAR4	Lysosome. Cytoplasmic granule.	RNA catabolism	Endoribonuclease;Hydrolase
ENDOG	Mitochondrion.	DNA replication;Apoptotic processes;DNA replication	Endodeoxyribonuclease;Endo
EPS15		Endocytosis;Neurotransmitter release	Other G-protein modulator;Se
ERCC1		DNA repair	Endodeoxyribonuclease
EXOSC2	Cytoplasm. Nucleus nucleolus.	Nucleoside, nucleotide and nucleic acid metabolism	Exoribonuclease
EXOSC3	Cytoplasm. Nucleus nucleolus.	Nucleoside, nucleotide and nucleic acid metabolism;Other metabolism	Exoribonuclease;Esterase
EXOSC4	Cytoplasm. Nucleus nucleolus.	tRNA metabolism;rRNA metabolism;RNA catabolism	Exoribonuclease;Nucleotidyltr
EXOSC5	Nucleus nucleolus.	tRNA metabolism;rRNA metabolism;RNA catabolism	Exoribonuclease;Nucleotidyltr
EXOSC6	Cytoplasm. Nucleus nucleolus.	tRNA metabolism;rRNA metabolism;RNA catabolism	Exoribonuclease;Nucleotidyltr
EXOSC8		rRNA metabolism	Exoribonuclease;Hydrolase
EXOSC9	Cytoplasm. Nucleus nucleolus.	rRNA metabolism	Exoribonuclease;Hydrolase
HIC2	Nucleus.	mRNA transcription regulation	Zinc finger transcription facto
JARID1D		mRNA transcription;Spermatogenesis and motility	Other zinc finger transcriptior
JTV1		Protein biosynthesis	Damaged DNA-binding protei
LIN28	Cytoplasm. Nucleus nucleolus.	mRNA transcription regulation	Other transcription factor;Oth
LSM8	Nucleus.	mRNA splicing;Other metabolism	Nuclease;mRNA splicing facto
MAPK14		Protein phosphorylation;MAPKKK cascade	Non-receptor serine/threonin
MAPK7		Protein phosphorylation;MAPKKK cascade	Non-receptor serine/threonin
MATR3	Nucleus matrix.	Miscellaneous	Other RNA-binding protein;Ot
MBD1	Nucleus speckle.	mRNA transcription	Other DNA-binding protein
MBD3L1	Nucleus.	mRNA transcription	Other DNA-binding protein
MBD4	Nucleus.	DNA repair;Developmental processes;Other metabolism	Nuclease;Methyltransferase
MBNL2	Nucleus. Cytoplasm.	Muscle development	Double-stranded DNA binding
MLH1		DNA repair;Meiosis;Oncogenesis	Other DNA-binding protein
MRE11A	Nucleus.	DNA repair;DNA recombination;Meiosis	Exodeoxyribonuclease;Endod
MRPS28	Mitochondrion.	Biological process unclassified	Nuclease
NANOS1	Cytoplasm perinuclear region.	Protein biosynthesis;Translational regulation;Oogenesis;Meiosis;Embryogenesis	Nuclease
NDEL1		Biological process unclassified	Molecular function unclassifie
Ndn	Cytoplasm. Nucleus nucleoplasm. Nucleus matrix.	Cell cycle control;Cell proliferation and differentiation	Double-stranded DNA binding
NFKBIA		mRNA transcription regulation;NF-kappaB cascade;Intracellular protein traffic;S	Select regulatory molecule
NFKBIE		Biological process unclassified	Molecular function unclassifie
Otud7a	Cytoplasm. Nucleus.	Proteolysis;Other protein metabolism	Double-stranded DNA binding
PA2G4	Cytoplasm. Nucleus nucleolus.	Protein biosynthesis;Cell proliferation and differentiation	Other transcription factor;Oth
PAWR		Apoptosis	Molecular function unclassifie
PCBP3		Protein metabolism and modification	Select regulatory molecule
PELO	Nucleus. Cytoplasm.	Meiosis;Mitosis	Nuclease;Translation release f
PER3		mRNA transcription regulation;Cell communication	Transcription cofactor
PLEKHA3	Cytoplasm. Membrane; Peripheral membrane protein.	Biological process unclassified	Other nucleic acid binding;Tra
POGK	Nucleus.	Other mRNA transcription;Developmental processes	Other nucleic acid binding
PPP1R13B		Apoptosis;Cell cycle control;Cell proliferation and differentiation	Select regulatory molecule
Ptrf		mRNA transcription termination;rRNA metabolism	Other transcription factor
PURG	Nucleus.	General mRNA transcription activities	Other transcription factor;Sin
RABGEF1		Endocytosis;Other intracellular protein traffic	Guanyl-nucleotide exchange f
RAD1	Nucleus.	DNA repair;DNA recombination;Cell cycle control	Exodeoxyribonuclease;Hydrol
REX3		Biological process unclassified	Molecular function unclassifie
ROCK2		Protein phosphorylation;Other intracellular signaling cascade;Cell adhesion;Oth	Non-receptor serine/threonin
RPP21	Nucleus nucleolus.	tRNA metabolism	Nuclease;Hydrolase
RXRG		Regulation of lipid, fatty acid and steroid metabolism;mRNA transcription regul	Nuclear hormone receptor;Tra
SART3	Cytoplasm. Nucleus. Nucleus speckle.	Other nucleoside, nucleotide and nucleic acid metabolism	Nuclease
SBDS	Cytoplasm.	Biological process unclassified	Transcription factor;Nuclease
SIRT2		mRNA transcription regulation;Chromatin packaging and remodeling	Chromatin/chromatin-binding
SNAPC3	Nucleus.	mRNA transcription;Other metabolism	Nuclease
SSBP2	Nucleus.	mRNA transcription initiation;mRNA transcription regulation	Single-stranded DNA-binding
STATIP1		Biological process unclassified	Molecular function unclassifie
SWAP70	Cytoplasm. Cell membrane. Nucleus. Cell projection lamellipodium.	DNA recombination;B-cell- and antibody-mediated immunity	Other nucleic acid binding
TCEB2		mRNA transcription elongation	Transcription cofactor
TDRD1	Cytoplasm.	Nucleoside, nucleotide and nucleic acid metabolism	Nuclease
THAP11	Nucleus. Cytoplasm.	Biological process unclassified	Other DNA-binding protein
TRERF1	Nucleus.	DNA metabolism;RNA catabolism	Other DNA-binding protein
TSN	Cytoplasm. Nucleus.	DNA recombination;Immunity and defense	Single-stranded DNA-binding
TTF1	Nucleus. Nucleus nucleolus.	mRNA transcription termination	Other DNA-binding protein
TXK		Protein phosphorylation;Intracellular signaling cascade;T-cell mediated immunit	Non-receptor tyrosine protein
USP52	Cytoplasm. Nucleus.	RNA catabolism;Proteolysis	Exoribonuclease;Esterase;Cys
XAB2	Nucleus.	DNA repair;mRNA transcription;Developmental processes	Other transcription factor;Dar
XRN1	Cytoplasm.	Nucleoside, nucleotide and nucleic acid metabolism	Exoribonuclease
ZFP143		mRNA transcription regulation	Zinc finger transcription facto
ZFP148	Nucleus.	mRNA transcription regulation	Zinc finger transcription facto
ZFP259		Biological process unclassified	Other miscellaneous function
ZFP281		mRNA transcription regulation	Zinc finger transcription facto
ZFP346	Nucleus nucleolus. Cytoplasm.	Electron transport;Apoptosis;Other metabolism	Zinc finger transcription facto
ZFP36L1		RNA catabolism;Intracellular signaling cascade	Other RNA-binding protein
ZFP521		Biological process unclassified	KRAB box transcription factor
ZFR	Nucleus. Cytoplasm. Cytoplasmic granule.	Apoptotic processes;Anterior/posterior patterning;Cell cycle	Other RNA-binding protein;Ot

ligase, interferon-inducible tripartite-motif (TRIM) 56, is a positive regulator of the ISD pathway that targets STING for ubiquitination and subsequent type I IFN induction[139]. Overexpression of TRIM56 increased IFN-β promoter activation while knockdown reduced type I IFN activation. Negative regulators active in other innate immune pathways, including CYLD, A-20, and DUBA, shut off activation signals by using enzymes that dephosphorylate or deubiquitinate their targets[165-168]. CYLD is a regulatory mechanism of the NF-κB pathway that mediates inhibitory activity by reversing the ubiquitination of tumor necrosis factor receptor (TRAF)-associated factors TRAF2 and TRAF6[165]. Additionally, it has been shown that ectopic expression of CYLD inhibits the IRF3 signaling pathway and IFN production triggered by RIG-I; conversely, CYLD knockdown enhances the response[166]. Another regulator of the NF-κB pathway, A-20, is a potent inhibitor of NF-κB signaling. A-20 deficient mice fail to regulate NF-κB, resulting in increased cell death and chronic inflammation[169]. A-20 exerts two opposing activities: sequential deubiquitination and ubiquitination of the TNF receptor-interacting protein (RIP) an essential mediator of the TNF receptor signaling complex, thereby targeting RIP to proteasomal degradation[167].

Deubiquitinating enzyme A (DUBA) targets TRAF3, an adapter protein critical to the type I IFN response. By selectively cleaving polyubiquitin chains of TRAF3, DUBA effectively dissociates TRAF3 from the downstream signaling complex containing TBK1[168]. Lastly, the importance of negative regulation of the ISD pathway is exemplified by 3'-5' exonuclease, TREX1[100, 104, 170]. Thought to prevent cell-intrinsic initiation of autoimmunity through clearance of endogenous retroelements, TREX1 is normally involved in clearance of ssDNA, but mice deficient in TREX1 have an accumulation of ~60bp ssDNA that drives the activation of DNA-damage associated signaling pathways[104, 171]. Loss-of-function mutations in the human gene *TREX1* cause

Aicardi- Goutières syndrome (AGS) and chilblain lupus, possibly driven by the accumulation of endogenous retroelements.

To this end, we identified 126 phosphatases (annotated from GO:0004721, phosphoprotein phosphatase activity) and 71 deubiquitinases (annotated from GO:0004221, ubiquitin thiolesterase activity, as well as a prior curation[172]) as a source of potential negative regulators of the ISD pathway (Table 3.6). We supplemented the phosphatases and deubiquitinases candidates to include 36 putative negative regulators of the RIG-I pathway identified in protein-protein interaction networks of influenza-host interactions that were also upregulated following stimulation with IFN-β in expression profiles of HBECs (included in Table 3.6)[134]. Finally, we added a collection of 38 known signaling molecules, and negative (no siRNA or negative control siRNA) and positive (siIRF3) controls (Table 3.7).

3.7 - *Conclusions*

In conclusion, we developed a comprehensive candidate list to identify novel components of the ISD-sensing pathway. Our list of 1003 genes represents a targeted list representative of both experimental and hypothesis-driven evidence. We selected candidates from gene expression experiments that were regulated in response to IFN-β or transfected DNA. In a study using ISD as bait in IFN-β stimulated cells, we identified candidates with SILAC-based mass spectrometry. We hypothesized that any annotated helicase could play a role in the ISD sensing pathway and thus, included all proteins with enzymatic helicase activity. Additionally, using cellular localization algorithms and annotated databases, we selected DNA-binding proteins with cytoplasmic localization. In a pilot screen to identify potential negative regulators of the ISD pathway, we included annotated phosphatases and deubiquitinases, supplemented with putative negative

Table 3.6
Candidate Gene List: MiniScreen / Phosphatases / Deubiquitinases

Gene	GeneID	Accession	Panther Biological Process	Panther Molecular Function
0710001B24RIK	67446	NM_175118	Protein modification;MAPKKK cascade	Kinase inhibitor;Protein phosphatase
0710007A14RIK	67444	NM_023343	Protein phosphorylation;MAPKKK cascade;Other intracellular signaling cascade	Protein phosphatase
0710007A14RIK	67444	NM_023343	Protein phosphorylation;MAPKKK cascade;Other intracellular signaling cascade	Protein phosphatase
1110007C05RIK	66124	NM_025368	Biological process unclassified	Molecular function unclassified
1110012L19RIK	68618	NM_026787	Biological process unclassified	Molecular function unclassified
1300006C06RIK	74158	NM_028792	Biological process unclassified	Molecular function unclassified
1500011L16RIK	68991	NM_026899	Biological process unclassified	Molecular function unclassified
1700095N21RIK	76630	NM_029682	JAK-STAT cascade;Cell proliferation and differentiation	Cytokine
1810034K20RIK	67881	NM_023397	Biological process unclassified	Molecular function unclassified
2310043K02RIK	66959	NM_025869	Protein modification	Protein phosphatase
2410018I08RIK	69727	NM_177561	Proteolysis	Cysteine protease
2600013N14RIK	72201	NM_152812	Sex determination;Chromosome segregation	Nucleic acid binding;Other proteases
2700002L06RIK	72344	XM_126772	Proteolysis	Cysteine protease
2810004N20RIK	66461	NM_025576	Phospholipid metabolism;Protein phosphorylation	Protein phosphatase;Other phosphatase
2810403L02RIK	66525	NM_025616	Biological process unclassified	Molecular function unclassified
2810423O19RIK	67905	NM_198931	Protein phosphorylation;MAPKKK cascade;Other intracellular signaling cascade	Protein phosphatase
2810439M11RIK	72749	NM_183091	Biological process unclassified	Molecular function unclassified
2810449C13RIK	170707	NM_028344	Proteolysis	Cysteine protease
4921523A10RIK	110332	NM_173449	Protein phosphorylation;MAPKKK cascade;Other intracellular signaling cascade	Protein phosphatase
4930511O11RIK	75083	NM_029163	Proteolysis	Cysteine protease
4930550B20RIK	77593	NM_152825	Proteolysis	Cysteine protease
4930553M18RIK	75316	NM_026541	Biological process unclassified	Molecular function unclassified
4930586I02RIK	68149	NM_026580	Protein metabolism and modification;Immunity and defense	Other hydrolase
4932415L06RIK	319651	NM_176972	Proteolysis	Cysteine protease
5730538E15RIK	70675	NM_173443	Biological process unclassified	Molecular function unclassified
6330567E21RIK	76179	XM_357781	Proteolysis	Cysteine protease
9130017A15RIK	320717	NM_177242	Protein phosphorylation	Protein phosphatase
9930028C20RIK	226418	NM_178691	Biological process unclassified	Protease
A230072I16RIK	217057	NM_175004	Protein biosynthesis;Other metabolism	Esterase
A430075L18RIK	319468	NM_176919	Protein phosphorylation;MAPKKK cascade;Other intracellular signaling cascade	Protein phosphatase
A630020C16RIK	74996	NM_177249	Proteolysis;Gametogenesis	Cysteine protease
AA939927	99526	NM_133857	Proteolysis	Ubiquitin-protein ligase
ACOX2	93732	NM_053115	Fatty acid beta-oxidation	Oxidase
ACP1	11431	NM_021330	Protein phosphorylation;Receptor protein tyrosine kinase signaling pathway;Cell proliferation and differentiation	Protein phosphatase
ACSL1	14081	NM_007981	Fatty acid metabolism	Other ligase
AMOTL2	56332	NM_019764	Angiogenesis;Cell motility	Other miscellaneous function protein
APOA5	66113	NM_080434	Lipid and fatty acid transport;Transport;Blood circulation and gas exchange	Transporter;Apolipoprotein
AW456874	218232	NM_207232	Protein phosphorylation	Protein phosphatase
B130021E18	329908	NM_183225	Proteolysis	Cysteine protease
BAP1	104416	NM_027088	Proteolysis	Cysteine protease
BC002236	79560	NM_024475	Biological process unclassified	Molecular function unclassified
C130067A03RIK	320713	NM_177239	mRNA transcription regulation	Transcription cofactor;Nucleic acid binding
C330046L10RIK	224836	NM_198421	Proteolysis	Cysteine protease
C6.1A	210766	NM_145956	Biological process unclassified	Other miscellaneous function protein
C6.1AL	368203	NM_203660	Biological process unclassified	Other miscellaneous function protein
C79127	232941	NM_177691	Protein phosphorylation;MAPKKK cascade;Other intracellular signaling cascade	Protein phosphatase
C920001D21RIK	320139	NM_177081	Protein phosphorylation;MAPKKK cascade	Protein phosphatase
CART1	216285	NM_172553	mRNA transcription regulation;Skeletal development	Homeobox transcription factor;Other DNA-binding protein
CDC14A	229776	XM_149387	Protein phosphorylation;Cell cycle control;Mitosis	Protein phosphatase
CDC14B	218294	NM_172587	Protein phosphorylation;Cell cycle control;Mitosis	Protein phosphatase
CDC25A	12530	NM_007658	Protein modification;Phosphate metabolism;Cell cycle control;Cell proliferation and differentiation;Other metabolism	Protein phosphatase
CDC25B	12531	NM_023117	Protein modification;Phosphate metabolism;Cell cycle control;Cell proliferation and differentiation;Other metabolism	Protein phosphatase
CDC25C	12532	NM_009860	Protein modification;Phosphate metabolism;Cell cycle control;Cell proliferation and differentiation;Other metabolism	Protein phosphatase
CDKN3	72391	XM_354809	Protein phosphorylation;Cell cycle	Protein phosphatase
CLK2	12748	NM_007712	Protein phosphorylation	Non-receptor serine/threonine protein kinase
COPS6	26893	NM_012002	Cell cycle control	Molecular function unclassified
CTBP1	13016	NM_013502	Amino acid biosynthesis	Transcription cofactor;Dehydrogenase
CTDP1	67655	NM_026295	mRNA transcription initiation	Protein phosphatase
CTDSP1	227292	NM_153088	Biological process unclassified	Molecular function unclassified
CTDSP2	52468	NM_146012	Biological process unclassified	Molecular function unclassified
CTDSPL	69274	NM_133710	Biological process unclassified	Molecular function unclassified
CYLD	74256	NM_173369	Protein biosynthesis	Ribosomal protein
D430025H09RIK	102747	NM_145616	Biological process unclassified	Molecular function unclassified
DHCR24	74754	NM_053272	Cholesterol metabolism	Reductase
DNAJB9	27362	NM_013760	Protein folding	Other chaperones
DNAJC6	72685	NM_198412	Protein phosphorylation;Other protein metabolism;Endocytosis	Other chaperones;Non-receptor serine/threonine protein kinase
DUB-1A	381944	NM_201409	Proteolysis	Cysteine protease
DUB1	13531	NM_007887	Proteolysis	Cysteine protease
DUB2	13532	NM_010089	Proteolysis	Cysteine protease
DULLARD	67181	NM_026017	Neurogenesis	Molecular function unclassified
Dusp1	19252	NM_013642	Protein phosphorylation;MAPKKK cascade;JNK cascade;Muscle development;Cell cycle control	Kinase inhibitor;Protein phosphatase
DUSP10	63953	NM_022019	Protein phosphorylation;JNK cascade;Stress response	Kinase inhibitor;Protein phosphatase
DUSP11	72102	NM_028099	mRNA capping	Phosphatase;Nucleotidyltransferase
DUSP12	80915	NM_023173	Glycolysis;Protein phosphorylation	Protein phosphatase
DUSP13	27389	NM_013849	Protein modification	Protein phosphatase
DUSP14	56405	NM_019819	Protein modification;MAPKKK cascade	Kinase inhibitor;Protein phosphatase
DUSP15	252864	NM_145744	Biological process unclassified	Molecular function unclassified
DUSP18	75219	NM_173745	Protein modification;MAPKKK cascade	Kinase inhibitor;Protein phosphatase
DUSP19	68082	NM_024438	Protein phosphorylation;JNK cascade	Protein phosphatase
DUSP2	13537	NM_010090	Protein phosphorylation;MAPKKK cascade;JNK cascade;Muscle development;Cell cycle control	Kinase inhibitor;Protein phosphatase
DUSP22	105352	NM_134068	Protein phosphorylation;JNK cascade;Stress response	Protein phosphatase
DUSP23	68440	XM_129566	Protein phosphorylation	Protein phosphatase
DUSP3	72349	NM_028207	Protein modification;Other intracellular signaling cascade	Protein phosphatase
DUSP4	319520	NM_176933	Protein phosphorylation;MAPKKK cascade;JNK cascade;Muscle development;Cell cycle control	Kinase inhibitor;Protein phosphatase
DUSP6	67603	NM_026268	Protein phosphorylation;MAPKKK cascade;Cell proliferation and differentiation	Kinase inhibitor;Protein phosphatase
DUSP7	235584	NM_153459	Protein phosphorylation;MAPKKK cascade;Cell proliferation and differentiation	Kinase inhibitor;Protein phosphatase
DUSP8	18218	NM_008748	Protein phosphorylation;JNK cascade;Stress response	Kinase inhibitor;Protein phosphatase
DUSP9	75590	NM_029352	Protein phosphorylation;MAPKKK cascade;Cell proliferation and differentiation	Kinase inhibitor;Protein phosphatase
DXIMX46E	54644	NM_138604	Sex determination;Chromosome segregation	Nucleic acid binding;Other proteases
E230037B21RIK	328417	XM_283217	Biological process unclassified	Molecular function unclassified
E430004F17	327799	NM_183199	Proteolysis	Cysteine protease
EIF3S3	68135	NM_080635	Developmental processes	Other transcription factor
EIF3S5	66085	NM_025344	Protein biosynthesis	Translation initiation factor
EPM2A	13853	NM_010146	Phospholipid metabolism;Protein phosphorylation	Protein phosphatase;Other phosphatase
EYA1	14048	NM_010164	Vision;Developmental processes	Hydrolase
EYA2	14049	NM_010165	Vision;Developmental processes	Hydrolase
EYA3	14050	NM_010166	Vision;Developmental processes	Hydrolase
Fosl1	14283	NM_010235	mRNA transcription regulation;T-cell mediated immunity;Stress response;Cell proliferation and differentiation;Cell motility	Other transcription factor;Nucleic acid binding
FOXJ1	15223	NM_008240	Carbohydrate metabolism;mRNA transcription regulation;Other receptor mediated signaling pathway;Cell communication;Vision;Embryogenesis;Anterior/posterior patterning	Other transcription factor;Nucleic acid binding
GM1395	231637	NM_198109	Protein phosphorylation;Other intracellular signaling cascade;Developmental processes;Cell structure and motility	Protein phosphatase
IGSF4A	54725	NM_018770	Biological process unclassified	Receptor
ITPKB	320404	XM_205854	Biological process unclassified	Molecular function unclassified
LOC240672	240672	XM_140740	Protein phosphorylation;MAPKKK cascade;JNK cascade	Kinase inhibitor;Protein phosphatase
MAP2K1	26395	NM_008927	Protein phosphorylation;Phosphate metabolism;Other metabolism	Protein kinase;Transferase
MJD	110616	NM_029705	Other mRNA transcription	Other transcription factor
MTM1	17772	NM_019926	Phospholipid metabolism;General vesicle transport;Other developmental process	Other phosphatase
MTMR1	53332	NM_016985	Phospholipid metabolism;General vesicle transport;Other developmental process	Other phosphatase
MTMR2	77116	NM_023858	Phospholipid metabolism;General vesicle transport;Other developmental process	Other phosphatase
MTMR3	74302	NM_028860	Phospholipid metabolism;General vesicle transport	Other phosphatase
MTMR4	170749	NM_133215	Phospholipid metabolism;General vesicle transport	Other phosphatase
MTMR6	219135	NM_144843	Phospholipid metabolism;General vesicle transport	Other phosphatase
MTMR7	54384	NM_019433	Phospholipid metabolism;General vesicle transport	Other phosphatase
NALP10	244202	NM_175532	Induction of apoptosis	Molecular function unclassified
NKX6-1	18096	NM_144955	mRNA transcription regulation;Skeletal development;Muscle development;Heart development	Homeobox transcription factor;Other DNA-binding protein
ORF21	224440	NM_145482	Biological process unclassified	Molecular function unclassified
OSR2	107587	NM_054049	mRNA transcription;Other developmental process;Segment specification	Other zinc finger transcription factor;Nucleic acid binding
OTUB1	107260	NM_134150	Protein metabolism and modification;Immunity and defense	Other hydrolase
PHPT1	75454	XM_203853	Biological process unclassified	Molecular function unclassified
PIP3AP	268783	NM_172958	Phospholipid metabolism;General vesicle transport	Other phosphatase
PLCE1	74055	NM_019588	Lipid metabolism;Intracellular signaling cascade	Other signaling molecule;Guanyl-nucleotide exchange factor;Hydrolase;Select calcium binding protein
PLEKHE1	98432	XM_129968	Other immune and defense;Miscellaneous	Other miscellaneous function protein
PPEF2	19023	NM_011148	Regulation of carbohydrate metabolism;Other polysaccharide metabolism;Glycogen metabolism;mRNA transcription;Protein phosphorylation;Translational regulation	Protein phosphatase;Select calcium binding protein
PPM1B	19043	NM_011151	Protein phosphorylation;MAPKKK cascade;Other intracellular signaling cascade	Protein phosphatase
PPM1D	53892	NM_016910	Protein phosphorylation;MAPKKK cascade;Other intracellular signaling cascade	Protein phosphatase
PPM1F	68606	NM_176833	Protein phosphorylation;MAPKKK cascade;Other intracellular signaling cascade	Protein phosphatase
PPM1G	14208	NM_008014	Protein phosphorylation;MAPKKK cascade;Other intracellular signaling cascade	Protein phosphatase
PPM1L	242083	NM_178726	Protein phosphorylation;MAPKKK cascade;Other intracellular signaling cascade	Protein phosphatase
PPP1CA	19045	NM_031868	Glycogen metabolism;Pre-mRNA processing;Protein phosphorylation;Apoptosis;Meiosis;Cell structure and motility	Protein phosphatase

Table 3.6
Candidate Gene List: MiniScreen / Phosphatases / Deubiquitinases (continued)

Gene	GeneID	Accession	Panther Biological Process	Panther Molecular Function
PPP1CB	19046	NM_172707	Glycogen metabolism;Pre-mRNA processing;Protein phosphorylation;Apoptosis;Meiosis;Cell structure and motility	Protein phosphatase
PPP1CC	19047	NM_013636	Glycogen metabolism;Pre-mRNA processing;Protein phosphorylation;Apoptosis;Meiosis;Cell structure and motility	Protein phosphatase
PPP1R12A	17931	XM_137239	Protein phosphorylation	Phosphatase modulator
PPP1R15B	108954	NM_133819	Biological process unclassified	Molecular function unclassified
PPP1R3A	140491	NM_080464	Glycogen metabolism	Phosphatase modulator
Ppp2ca	19052	NM_019411	Biological process unclassified	Protein phosphatase
PPP2CB	19053	NM_017374	Biological process unclassified	Protein phosphatase
PPP2CZ	71887	NM_027982	Protein phosphorylation;MAPKKK cascade;Other intracellular signaling cascade	Protein phosphatase
PPP3CA	19055	NM_008913	Regulation of carbohydrate metabolism;Other polysaccharide metabolism;Glycogen metabolism;mRNA transcription;Protein phosphorylation;Translational regulation;	Protein phosphatase;Select calcium binding protein
PPP3CB	19056	NM_008914	Regulation of carbohydrate metabolism;Other polysaccharide metabolism;Glycogen metabolism;mRNA transcription;Protein phosphorylation;Translational regulation	Protein phosphatase;Select calcium binding protein
PPP3CC	19057	NM_008915	Regulation of carbohydrate metabolism;Other polysaccharide metabolism;Glycogen metabolism;mRNA transcription;Protein phosphorylation;Translational regulation	Protein phosphatase;Select calcium binding protein
PPP3R2	19059	XM_143782	Protein modification;Calcium mediated signaling;Other homeostasis activities	Protein phosphatase;Calmodulin related protein
PPP4C	56420	NM_019674	Biological process unclassified	Protein phosphatase
PPP5C	19060	NM_011155	Regulation of carbohydrate metabolism;Other polysaccharide metabolism;Glycogen metabolism;mRNA transcription;Protein phosphorylation;Translational regulation	Protein phosphatase
PPP6C	67857	NM_024209	Biological process unclassified	Protein phosphatase
PRKACA	18747	NM_008854	Protein phosphorylation;Intracellular signaling cascade;Cell communication;Developmental processes;Mitosis;Cell proliferation and differentiation	Non-receptor serine/threonine protein kinase
PRPF8	192159	NM_138659	mRNA splicing	mRNA splicing factor
PSMD14	59029	NM_021526	Protein metabolism and modification	Other miscellaneous function protein
PSMD7	17463	NM_010817	Proteolysis;Developmental processes;Cell cycle control;Cell proliferation and differentiation	Other miscellaneous function protein
PTEN	19211	NM_008960	Phospholipid metabolism;Protein phosphorylation;Signal transduction;Cell adhesion;Immunity and defense;Induction of apoptosis;Cell cycle control;Cell proliferation and differentiation;Tumor suppressor	Protein phosphatase;Other phosphatase
PTP4A1	19243	NM_011200	Protein phosphorylation	Protein phosphatase
PTP4A2	19244	NM_008974	Protein phosphorylation	Protein phosphatase
PTP4A3	19245	NM_008975	Protein phosphorylation	Protein phosphatase
Ptpla	30963	NM_013935	Miscellaneous	Other miscellaneous function protein
PTPLB	70757	NM_023587	Biological process unclassified	Other miscellaneous function protein
Ptpn1	19246	NM_011201	Protein phosphorylation;Cytokine and chemokine mediated signaling pathway;Receptor protein tyrosine kinase signaling pathway;JAK-STAT cascade;Glucose homeostasis	Protein phosphatase
PTPN11	19247	NM_011202	Protein phosphorylation;Cytokine and chemokine mediated signaling pathway;Receptor protein tyrosine kinase signaling pathway;JAK-STAT cascade	Protein phosphatase
PTPN12	19248	NM_011203	Protein phosphorylation;Cell adhesion-mediated signaling;Cell adhesion;Protein targeting;Cell motility	Protein phosphatase
PTPN13	19249	NM_011204	Protein modification;Ectoderm development;Cell motility	Protein phosphatase
PTPN14	19250	NM_008976	Protein modification;Ectoderm development;Cell motility	Protein phosphatase
PTPN18	19253	NM_011206	Protein phosphorylation	Protein phosphatase
PTPN2	19255	NM_008977	Protein phosphorylation;Cytokine and chemokine mediated signaling pathway;Receptor protein tyrosine kinase signaling pathway;JAK-STAT cascade;Glucose homeostasis	Protein phosphatase
PTPN20	19256	NM_008978	Protein modification;Ectoderm development;Cell motility	Protein phosphatase
PTPN21	24000	NM_011877	Protein modification;Ectoderm development;Cell motility	Protein phosphatase
PTPN23	104831	XM_135197	Protein phosphorylation;Other receptor mediated signaling pathway	Protein phosphatase
PTPN4	19258	NM_019933	Biological process unclassified	Molecular function unclassified
PTPN5	19259	NM_013643	Protein phosphorylation	Protein phosphatase
PTPN8	19260	NM_008979	Protein phosphorylation;Other receptor mediated signaling pathway;Cell adhesion-mediated signaling;Extracellular matrix protein-mediated signaling;Cell adhesion;Other immune and defense	Protein phosphatase
PTPN9	56294	NM_019651	Protein phosphorylation;General vesicle transport	Protein phosphatase
Ptpra	19262	NM_008980	Protein phosphorylation;Receptor protein tyrosine kinase signaling pathway;Cell adhesion-mediated signaling;Cell cycle control;Cell proliferation and differentiation;Cell structure;Cell motility	Protein phosphatase
PTPRB	19263	NM_029928	Biological process unclassified	Molecular function unclassified
PTPRC	19264	NM_011210	Protein phosphorylation;Cell communication;Immunity and defense;Cell proliferation and differentiation	Other receptor;Protein phosphatase
PTPRE	19267	NM_011212	Protein phosphorylation;Receptor protein tyrosine kinase signaling pathway;Cell adhesion-mediated signaling;Cell cycle control;Cell proliferation and differentiation;Cell structure;Cell motility	Protein phosphatase
PTPRF	19268	NM_011213	Protein phosphorylation;Cell adhesion-mediated signaling;Cell adhesion;Other developmental process;Neurogenesis;Cell structure and motility	Other receptor;Protein phosphatase
PTPRG	19270	NM_008981	Protein phosphorylation;Cell surface receptor mediated signal transduction;Cell communication;Cell adhesion;Other developmental process	Other receptor;Protein phosphatase
PTPRJ	19271	NM_008982	Protein phosphorylation;Receptor protein tyrosine kinase signaling pathway;Cell communication	Other receptor;Protein phosphatase
PTPRK	19272	NM_008983	Protein phosphorylation;Cell surface receptor mediated signal transduction;Cell adhesion-mediated signaling;Cell adhesion	Other receptor;Protein phosphatase
PTPRM	19274	NM_008984	Protein phosphorylation;Cell surface receptor mediated signal transduction;Cell adhesion-mediated signaling;Cell adhesion	Other receptor;Protein phosphatase
PTPRN	19275	NM_008985	Biological process unclassified	Molecular function unclassified
PTPRN2	19276	NM_011215	Protein phosphorylation;Glucose homeostasis	Other receptor;Protein phosphatase
PTPRR	19279	NM_011217	Protein phosphorylation	Other receptor;Protein phosphatase
PTPRS	19280	NM_011218	Protein phosphorylation;Cell adhesion-mediated signaling;Cell adhesion;Other developmental process;Neurogenesis;Cell structure and motility	Other receptor;Protein phosphatase
PTPRT	19281	NM_021464	Protein phosphorylation;Cell surface receptor mediated signal transduction;Cell adhesion-mediated signaling;Cell adhesion	Other receptor;Protein phosphatase
PTPRU	19273	NM_011214	Protein phosphorylation;Cell surface receptor mediated signal transduction;Cell adhesion-mediated signaling;Cell adhesion	Other receptor;Protein phosphatase
PTPRV	13924	NM_007955	Protein modification;Cell surface receptor mediated signal transduction;Cell communication	Other receptor;Protein phosphatase
RASGRP1	19419	NM_011246	G-protein mediated signaling;Receptor protein tyrosine kinase signaling pathway;Intracellular signaling cascade	Guanyl-nucleotide exchange factor
RNASEL	24014	NM_011882	Biological process unclassified	Molecular function unclassified
RNGTT	24018	NM_011884	mRNA capping	Phosphatase;Nucleotidyltransferase
RPS6KA1	20111	NM_009097	Protein phosphorylation;MAPKKK cascade;Neurogenesis;Skeletal development;Cell cycle control	Non-receptor serine/threonine protein kinase
RPS6KA3	110651	NM_148945	Protein phosphorylation;MAPKKK cascade;Neurogenesis;Skeletal development;Cell cycle control	Non-receptor serine/threonine protein kinase
S100A3	20197	NM_011310	Oncogenesis	Calmodulin related protein
SLC27A1	26457	NM_011977	Fatty acid metabolism;Lipid and fatty acid transport;Transport;Other metabolism	Transporter
SSH2	237860	NM_177710	Protein phosphorylation;Other intracellular signaling cascade;Developmental processes;Cell structure and motility	Protein phosphatase
SSH3	245857	NM_198113	Protein phosphorylation;Other intracellular signaling cascade;Developmental processes;Cell structure and motility	Protein phosphatase
STAMBP	70527	NM_024239	JAK-STAT cascade;Cell proliferation and differentiation	Cytokine
STYX	56291	NM_019637	Biological process unclassified	Other miscellaneous function protein
Tenc1	209039	NM_153533	Signal transduction;Mesoderm development;Cell motility	Non-motor actin binding protein
TOB2	57259	NM_020507	Receptor protein serine/threonine kinase signaling pathway;MAPKKK cascade;JNK cascade;Cell cycle control;Cell proliferation and differentiation	Other miscellaneous function protein
UCHL1	22223	NM_011670	Proteolysis	Cysteine protease
UCHL3	50933	NM_016723	Proteolysis	Cysteine protease
UCHL4	93841	NM_033607	Proteolysis	Cysteine protease
UCHL5	56207	NM_019562	Proteolysis	Cysteine protease
USP1	230484	NM_146144	Proteolysis	Cysteine protease
USP10	22224	NM_009462	Proteolysis	Cysteine protease
USP11	236733	NM_145628	Proteolysis	Cysteine protease
USP12	22217	NM_011669	Proteolysis	Cysteine protease
USP13	72607	XM_130826	Proteolysis	Cysteine protease
USP14	59025	NM_021522	Proteolysis	Cysteine protease
USP15	14479	NM_027604	Proteolysis	Cysteine protease
USP16	74112	NM_024258	Proteolysis	Cysteine protease
USP2	53376	NM_016808	Proteolysis	Cysteine protease
USP20	74270	NM_028846	Proteolysis	Cysteine protease
USP21	30941	NM_013919	Proteolysis	Cysteine protease
USP22	216825	XM_109894	Proteolysis	Cysteine protease
USP26	83563	NM_031388	Proteolysis	Cysteine protease
USP27X	54651	NM_019461	Proteolysis	Cysteine protease
USP28	235323	NM_175482	Proteolysis	Cysteine protease
USP29	57775	NM_021323	Proteolysis	Cysteine protease
USP3	235441	NM_144937	Proteolysis	Cysteine protease
USP30	100756	XM_149655	Proteolysis	Cysteine protease
USP33	170822	NM_133247	Proteolysis	Cysteine protease
USP38	74841	NM_027554	Proteolysis	Cysteine protease
USP4	22258	NM_011678	Proteolysis	Cysteine protease
Usp42	76800	XM_132483	Biological process unclassified	Molecular function unclassified
USP43	216835	NM_173754	Proteolysis	Cysteine protease
USP47	74996	XM_133733	Proteolysis;Gametogenesis	Cysteine protease
USP5	22225	NM_013700	Proteolysis	Cysteine protease
USP54	78787	NM_030180	Biological process unclassified	Molecular function unclassified
USP7	252870	XM_148584	Proteolysis	Cysteine protease
USP8	84092	NM_019729	Proteolysis	Cysteine protease
USP9X	22284	NM_009481	Proteolysis;Gametogenesis	Cysteine protease
USP9Y	107868	NM_148943	Proteolysis;Gametogenesis	Cysteine protease
ZFPN1A3	22780	XM_283022	Biological process unclassified	KRAB box transcription factor

Table 3.7
Candidate Gene List: Signaling Molecules

Gene	GeneID	Accession	Panther Biological Process	Panther Molecular Function
ASC	66824	NM_023258	Proteolysis;Apoptosis	Protease inhibitor
AW046014	106759	NM_174989	Intracellular protein traffic;Transport	Transfer/carrier protein;Vesicle coat protein
BC034204	270151	NM_178420	Biological process unclassified	Molecular function unclassified
CARD15	257632	NM_145857	Other immune and defense;Induction of apoptosis	Molecular function unclassified
CASP8	12370	NM_009812	Proteolysis;Apoptosis	Cysteine protease
CHUK	12675	NM_007700	Protein phosphorylation;NF-kappaB cascade;Cytokine/chemokine mediated immunity	Non-receptor serine/threonine protein kinase
D430028G21RIK	228607	NM_144888	Biological process unclassified	Molecular function unclassified
EYA4	14051	NM_010167	Vision;Developmental processes	Hydrolase
FADD	14082	NM_010175	Apoptotic processes	Other miscellaneous function protein
IFI16	15951	NM_008329	mRNA transcription regulation;Interferon-mediated immunity;Cell proliferation and differentiation	Other transcription factor
Ikbkb	16150	NM_010546	Protein phosphorylation;NF-kappaB cascade;Cytokine/chemokine mediated immunity	Non-receptor serine/threonine protein kinase
IKBKE	56489	NM_019777	Protein phosphorylation;NF-kappaB cascade;Cytokine/chemokine mediated immunity	Non-receptor serine/threonine protein kinase
IKBKG	16151	NM_178590	Biological process unclassified	Molecular function unclassified
IRAK1	16179	NM_008363	Protein phosphorylation;Receptor protein serine/threonine kinase signaling pathway;Other immune and defense;Non-vertebrate process	Serine/threonine protein kinase receptor;Protein kinase
IRF3	54131	NM_016849	mRNA transcription regulation;Interferon-mediated immunity;Oncogenesis	Other transcription factor;Nucleic acid binding
ISGF3G	16391	NM_008394	mRNA transcription regulation;Interferon-mediated immunity;Oncogenesis	Other transcription factor;Nucleic acid binding
JAK1	16451	NM_146145	Protein phosphorylation;Receptor protein tyrosine kinase signaling pathway;JAK-STAT cascade;Apoptosis;Mesoderm development;Cell cycle control;Cell proliferation and differentiation;Cell motility	Non-receptor tyrosine protein kinase
JAK2	16452	NM_008413	Protein phosphorylation;Receptor protein tyrosine kinase signaling pathway;JAK-STAT cascade;Apoptosis;Mesoderm development;Cell cycle control;Cell proliferation and differentiation;Cell motility	Non-receptor tyrosine protein kinase
JAK3	16453	NM_010589	Protein phosphorylation;Receptor protein tyrosine kinase signaling pathway;JAK-STAT cascade;Apoptosis;Mesoderm development;Cell cycle control;Cell proliferation and differentiation;Cell motility	Non-receptor tyrosine protein kinase
MAPK1	26413	NM_011949	Protein phosphorylation;MAPKKK cascade	Non-receptor serine/threonine protein kinase
MYD88	17874	NM_010851	NF-kappaB cascade;Other immune and defense	Kinase activator
NFKBIA	18035	NM_010907	mRNA transcription regulation;NF-kappaB cascade;Intracellular protein traffic;Stress response;Apoptosis;Cell proliferation and differentiation	Select regulatory molecule
NFKBIB	18036	NM_010908	mRNA transcription regulation;NF-kappaB cascade;Intracellular protein traffic;Stress response;Apoptosis;Cell proliferation and differentiation	Select regulatory molecule
Pin1	23988	NM_023371	Protein folding;Cytokinesis	Other isomerase
RIPK1	19766	NM_009068	Protein phosphorylation;NF-kappaB cascade;Induction of apoptosis	Non-receptor serine/threonine protein kinase
RIPK3	56532	NM_019955	Protein phosphorylation;NF-kappaB cascade;Induction of apoptosis	Non-receptor serine/threonine protein kinase
SOCS1	12703	NM_009896	Cytokine and chemokine mediated signaling pathway	Kinase inhibitor
STAT3	20848	NM_011486	mRNA transcription regulation;JAK-STAT cascade;Stress response;Inhibition of apoptosis;Hematopoiesis;Other oncogenesis;Cell motility	Other transcription factor;Nucleic acid binding
TANK	21353	NM_011529	Biological process unclassified	Molecular function unclassified
TBK1	56480	NM_019786	Protein phosphorylation;NF-kappaB cascade;Cytokine/chemokine mediated immunity	Non-receptor serine/threonine protein kinase
TNFAIP3	21929	NM_009397	Proteolysis;Other protein metabolism	Double-stranded DNA binding protein;Hydrolase
TRAF1	22029	NM_009421	Cytokine and chemokine mediated signaling pathway;Induction of apoptosis	Other signaling molecule;Other miscellaneous function protein
TRAF2	22030	NM_009422	Cytokine and chemokine mediated signaling pathway;NF-kappaB cascade;JNK cascade;Induction of apoptosis;Cell proliferation and differentiation	Other signaling molecule;Other miscellaneous function protein
TRAF3	22031	NM_011632	Cytokine and chemokine mediated signaling pathway;NF-kappaB cascade;JNK cascade;Induction of apoptosis;Cell proliferation and differentiation	Other signaling molecule;Other miscellaneous function protein
TRAF4	22032	NM_009423	Cytokine and chemokine mediated signaling pathway;NF-kappaB cascade;JNK cascade;Induction of apoptosis;Skeletal development	Other signaling molecule;Other miscellaneous function protein
TRAF5	22033	NM_011633	Cytokine and chemokine mediated signaling pathway;NF-kappaB cascade;JNK cascade;Induction of apoptosis;Cell proliferation and differentiation	Other signaling molecule;Other miscellaneous function protein
TRAF6	22034	NM_009424	Cytokine and chemokine mediated signaling pathway;NF-kappaB cascade;JNK cascade;Induction of apoptosis;Skeletal development	Other signaling molecule;Other miscellaneous function protein
TRAM	225471	NM_173394	Intracellular protein traffic;Transport	Transfer/carrier protein;Vesicle coat protein

regulators of the RIG-I pathway identified in a dissection of host-influenza interactions. Finally, to complement our list of putative negative regulators, we included known signaling molecules.

Chapter 4:

A High-throughput Loss-of-Function RNAi Screen for of the ISD-Sending Pathway Reveals Identifies Known Components and Novel Regulators

4.1 – *Introduction*

Following the development of a high-throughput screening system (Figure 2.18) and selection of candidate list (Figure 3.1) to identify novel components of the ISD pathway, we screened 1003 putative factors using siGENOME SmartPools (Dharmacon; ICCB, Harvard Medical School). We conducted our screening in four stages. First, as test of our screening workflow and analysis, we conducted a small pilot screen of 235 potential negative regulators. Second, we selected the strongest hits from the pilot screen, added them to the full screening set and conducted a screen of 783 candidate ISD pathway components. Next, we developed a database to quickly access and interrogate the top hits. In an attempt to recover candidates lost to siRNA toxicity we screened devonvoluted siGENOME SmartPools to identify additional candidates in the ISD-pathway. Finally, we conducted a series of secondary screens, including screens of our top candidates with deconvoluted siGENOME SmartPools pools and independent siRNA pools (Dharmacon On-TargetPLUS SmartPools) to identify DNA-response specificity. Select candidates were nominated for further investigation.

4.2 – *Pilot Screen: phosphatases and deubiquitinases*

With the completion of an siRNA-based screening strategy and the development of a set of candidate genes from genomic, proteomic, and domain-based studies we set out to conduct a pilot screen of system testing workflow and analysis. Our pilot screen target 235 phosphatases and deubiquitinases (Table 3.6). The average of triplicate wells is shown in Figure 4.1. A best-fit line based on non-targeting siRNA, no siRNA and siDEATH controls is drawn through Cxcl10. We then calculated Z-scores normalized to the best-fit line. Data are presented as Log(2) Cxcl10 pg/mL and CellTiter-Glo relative luminesce units (Table 4.1 and Table 4.2).

Figure 4.1

Pilot Screen: Phosphatases and Deubiquitinases. Following knockdown of siRNAs targeting a total of 235 phosphatases and deubiquitinases, each well was stimulated with ISD and Cxcl10 was measured by ELISA. Log(2) Cxcl10 (pg/mL) is graphed on the y-axis. Cell viability after knockdown was measured by CellTiter-Glo; relative luminescence units are graphed on the x-axis. The averages of three replicate wells for each gene are represented as circles; red triangles at the bottom-right of each graph represent si*Irf3* positive controls; green triangles at the top-of each graph represent wells with non-targeting control siRNAs blue triangles represent siDeath controls and wells with no siRNA. Selected genes are marked with purple circles and bold text. A best-fit line is drawn through the Cxcl10 values of no siRNA, non-targeting and siDeath controls.

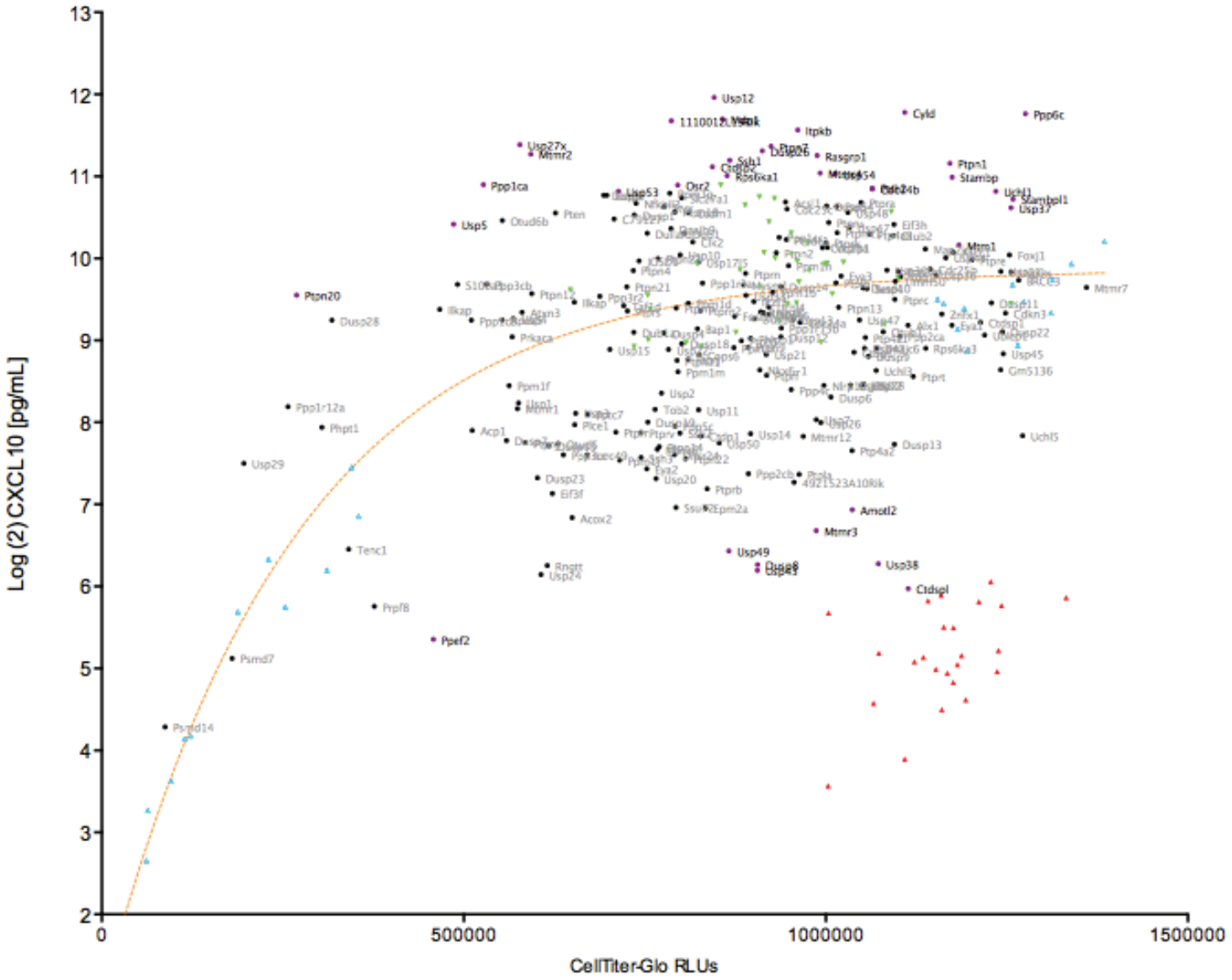


Figure 4.1

Pilot Screen: Phosphatases, Deubiquitinases and Signaling Molecules (continued)

Table 4.1

Negative Regulator Screen: Phosphatases

Gene Name	GeneID	CellTiter-Glo	log2(CXCL10, pg/mL)
No siRNA		1252393	12.47
Irf3	54131	1152920	8.69
4921523A10Rik	110332	956200	10.38
Acp1	11431	511613	10.99
C79127	232941	707507	13.48
Cdc14a	229776	964373	12.22
Cdc14b	218294	1064040	13.85
Cdc25a	12530	995200	13.13
Cdc25b	12531	1144240	12.87
Cdc25c	12532	946507	13.60
Cdkn3	72391	1248293	12.33
Ctdp1	67655	827373	10.87
Ctdsp1	227292	1213600	12.22
Ctdsp2	52468	843507	14.12
Ctdspl	69274	1113800	9.43
Dnajc6	72685	1070400	12.04
Dullard	67181	753453	13.31
Dusp1	19252	735547	13.53
Dusp10	63953	1051373	12.64
Dusp11	72102	1228907	12.46
Dusp12	80915	938053	12.29
Dusp13	27389	1094867	10.76
Dusp14	56405	938120	12.66
Dusp15	252864	617880	10.72
Dusp18	75219	801053	12.21
Dusp19	68082	754120	11.12
Dusp2	13537	901267	12.26
Dusp22	105352	1243907	12.29
Dusp23	68440	601680	10.35
Dusp26	66959	912147	14.31
Dusp28	67446	318053	12.25
Dusp3	72349	693360	13.77
Dusp4	319520	776320	12.56
Dusp5	240672	1000960	13.64
Dusp6	67603	1007187	11.72
Dusp7	235584	558893	11.25
Dusp8	18218	905947	9.86
Dusp9	75590	1058827	12.06
Epm2a	13853	833600	10.26
Eya1	14048	1174227	12.18
Eya2	14049	752653	10.62
Eya3	14050	1020947	12.79
Ilkap	67444	652627	12.47
Map2k1	26395	1137360	13.11
Mdp1	67881	857213	14.70
Mtm1	17772	1183827	13.16
Mtmr1	53332	574280	11.28
Mtmr12	268783	968920	10.97
Mtmr2	77116	592600	14.27
Mtmr3	74302	986827	9.96
Mtmr4	170749	992373	14.04
Mtmr6	219135	767400	10.77
Mtmr7	54384	1359973	12.64
Phlpp	98432	896187	11.90
Phpt1	75454	303947	10.98
Ppef2	19023	457947	8.49
Ppm1b	19043	926893	12.59
Ppm1d	53892	810093	12.45
Ppm1f	68606	562760	11.53
Ppm1g	14208	784640	13.80
Ppm1h	319468	948480	12.91
Ppm1j	71887	715200	10.52
Ppm1l	242083	873053	11.96
Ppm1m	67905	795653	11.70
Ppp1ca	19045	527160	13.90
Ppp1cb	19046	510400	12.25
Ppp1cc	19047	935707	13.26
Ppp1r12a	17931	257280	11.13
Ppp1r15b	108954	938533	12.15
Ppp1r3a	140491	829880	12.70
Ppp2ca	19052	1102253	12.30
Ppp2cb	19053	893013	10.57
Ppp3ca	19055	1099787	12.84
Ppp3cb	19056	531053	12.68
Ppp3cc	19057	637547	10.72
Ppp3r2	19059	687747	12.54
Ppp4c	56420	952093	11.67
Ppp5c	19060	791413	11.09
Ppp6c	67857	1275907	14.76
Pptc7	320717	670973	11.47
Pten	19211	626493	13.55
Ptp4a1	19243	1054733	12.24
Ptp4a2	19244	1036307	10.91
Ptp4a3	19245	1060640	13.30
Ptpdc1	218232	945440	13.23
Ptpla	30963	963213	10.50
Ptplb	70757	1014240	12.70
Ptpmt1	66461	1015600	13.31
Ptpn1	19246	1171400	14.16
Ptpn11	19247	794520	11.80
Ptpn12	19248	593733	12.57
Ptpn13	19249	1017880	12.40
Ptpn14	19250	770053	10.57
Ptpn18	19253	791467	13.57
Ptpn2	19255	931493	13.07
Ptpn20	19256	269093	12.55
Ptpn21	24000	725027	12.65
Ptpn22	19260	806120	10.59
Ptpn23	104831	768240	13.00
Ptpn4	19258	734347	12.85
Ptpn5	19259	883427	12.17
Ptpn7	320139	924387	14.37
Ptpn9	56294	892093	11.82
Ptpra	19262	1048747	13.68
Ptprb	19263	836013	10.10
Ptprc	19264	1095280	12.50
Ptpre	19267	1201827	12.98
Ptprf	19268	918360	11.71
Ptprg	19270	584987	10.97
Ptprj	19271	809973	11.82
Ptprk	19272	1000120	13.19
Ptprm	19274	793320	12.39
Ptprn	19275	889227	12.82
Ptprn2	19276	826760	12.37
Ptprr	19279	710000	11.32
Ptprs	19280	725907	12.36
Ptprt	19281	1120840	11.90
Ptpru	19273	1004160	13.44
Ptprv	13924	744733	10.77
Rngtt	24018	615387	9.54
Ssh1	231637	867280	14.20
Ssh2	237860	798600	10.97
Ssh3	245857	744827	10.60
Ssu72	68991	793067	10.23
Styx	56291	776827	13.63
Tenc1	209039	340813	9.76
Timm50	66525	1096107	12.72
Ublcp1	79560	1219333	12.27

Table 4.2

Negative Regulator Screen: Deubiquitinases

Gene Name	GeneID	CellTiter-Glo	log2(CXCL10, pg/mL)
No siRNA		1252393	12.47
Irf3	54131	1152920	8.69
Atxn3	110616	580653	12.34
Bap1	104416	822693	12.14
Brcc3	210766	1266227	12.73
Cops6	26893	825187	11.81
Cyld	74256	1109120	14.78
Dub1	13531	804600	13.31
Dub1a	381944	734987	12.16
Dub2	13532	823893	12.95
Eif3f	66085	622373	10.59
Eif3h	68135	1094093	13.41
Gm5136	368203	1241480	11.81
Josd1	74158	742213	12.97
Josd2	66124	1049867	11.46
Josd3	75316	721000	12.42
Mysm1	320713	886320	12.68
Otub1	107260	1079320	12.16
Otub2	68149	1092760	13.28
Otud5	54644	630227	10.94
Otud6b	72201	553160	13.46
Prpf8	192159	376427	8.85
Psmd14	59029	87347	6.71
Psmd7	17463	180093	7.78
Stambp	70527	1174547	13.99
Stambpl1	76630	1258400	13.72
Uchl1	22223	1234920	13.82
Uchl3	50933	1069520	11.68
Uchl4	93841	921360	12.41
Uchl5	56207	1271973	10.93
Usp1	230484	575520	11.12
Usp10	22224	798747	13.04
Usp11	236733	824427	11.23
Usp12	22217	845787	14.96
Usp13	72607	954680	12.26
Usp14	59025	896133	10.78
Usp15	14479	701933	11.68
Usp16	74112	914093	12.33
Usp2	53376	773147	11.25
Usp20	74270	765533	10.24
Usp21	30941	917680	11.75
Usp22	216825	782653	12.23
Usp24	329908	606640	9.49
Usp26	83563	993280	10.92
Usp27x	54651	577387	14.39
Usp28	235323	1052707	11.43
Usp29	57775	196280	10.29
Usp3	235441	654373	11.11
Usp30	100756	1084307	12.86
Usp31	76179	1241680	12.84
Usp33	170822	889720	12.55
Usp36	72344	1152133	12.80
Usp37	319651	1255640	13.62
Usp38	74841	1072520	9.46
Usp4	22258	568280	12.27
Usp42	76800	1053493	12.20
Usp43	216835	905573	9.28
Usp44	327799	1165787	13.01
Usp45	77593	1245093	11.95
Usp46	69727	921693	12.31
Usp47	74996	1053267	13.77
Usp48	170707	1030533	13.56
Usp49	224836	866147	9.80
Usp5	22225	485787	13.42
Usp50	75083	852840	10.87
Usp53	99526	713600	13.82
Usp54	78787	1012507	14.03
Usp7	252870	986733	10.74
Usp8	84092	697547	13.77
Usp9x	22284	1255720	12.83
Usp9y	107868	1033973	11.51
Vcpip1	70675	1002267	13.13
Yod1	226418	1103307	12.78

Knockdown of control gene *Irf3* reduced Cxcl10 expression by 19.1-fold (p-value <0.01, Mann Whitney U test). We identified seven potential negative regulators with 5.5 to 11.1-fold reduction in Cxcl10 response including the signaling protein *Amotl2*, phosphatases *Ctdspl*, *Dusp8, Mtmr3*, and *Ppef2* and ubiquitin specific peptidases *Usp38*, *Usp43*, and *Usp49*. Furthermore, we identified more than a dozen potential negative regulators with 5.09 to 8.8-fold increases in Cxcl10 response including the phosphatases *Itpkb*, *Mdp1*, *Mtmr2*, *Ppp1ca*, and *Ptpn20*, the deubiquitinase *Cyld* and the ubiquitin specific proteases *Usp5*, *Usp12*, and *Usp27x*. Most interesting among these candidates is *Cyld*, a negative regulator of RIG-I activation[166]. Following knockdown of *Cyld* we demonstrated an 8.8-fold increase in the CXCL10 response. Additionally, a strong candidate from the influenza screen, the hypothetical protein 1110012L19Rik, was replicated in our pilot screen. The strongest of the pilot screen hits were added to the 768 candidates selected for the complete screening set.

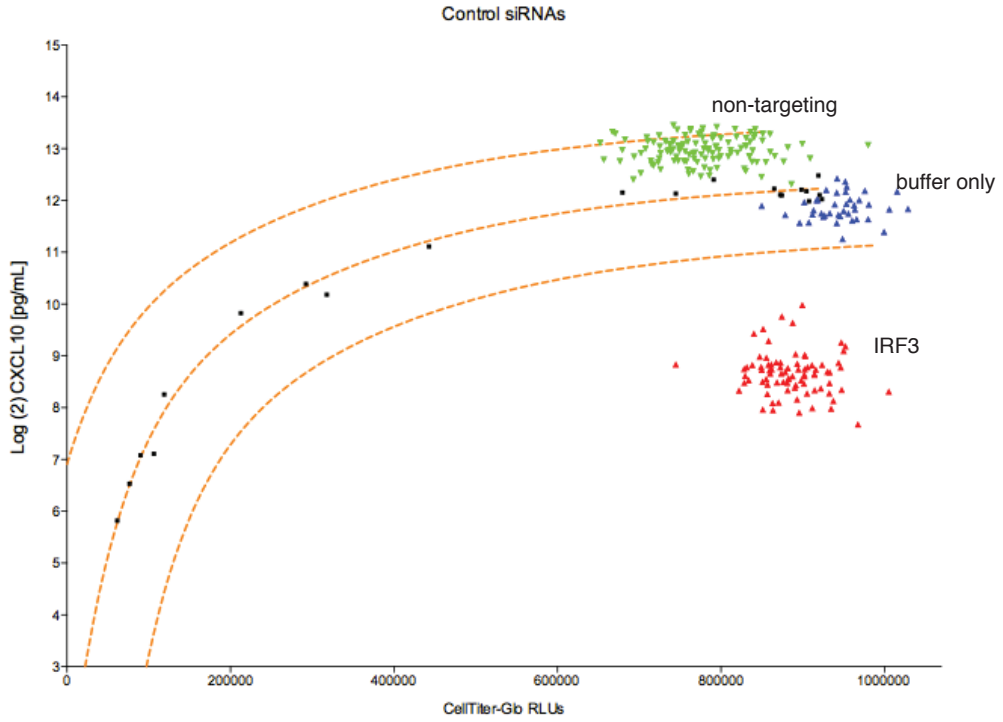
4.3 – *High-throughput loss-of-function RNAi screen*

Following the completion of the pilot screen, we conducted a screen of 783 siGENOME SmartPools, along with non-targeting, *Irf3* and no siRNA controls. A cell dilution curve was included to generate predicted Cxcl10 values for a given CellTiter-Glo value (Figure 4.2A). Screening fitness was assessed by Z-factor analysis (Figure 4.2B). Cxcl10 levels were detected following stimulation of non-targeting and *Irf3* controls samples at low and high supernatant volumes to determine assay detection limits in the context of the entire screen. Z-factor analysis (0.64 for low and 0.84 for high) indicated a robust screen. Furthermore, replicate samples (n=3 siGENOME SmartPools per gene on independent plates) correlated strongly for both Cxcl10 and CellTiter-Glo values (R^2 =

Figure 4.2

High-throughput Loss-of-Function RNAi Screen: Control siRNA and replicate assessment indicates a robust screen. A) Non-targeting and *Irf3* siRNA treated wells (green and red triangles, respectively) and buffer-only controls (blue triangles) were stimulated with ISD. Cxcl10 was measured by ELISA. Log(2) Cxcl10 (pg/mL) is graphed on the y-axis. Cell viability after knockdown was measured by CellTiter-Glo; relative luminescence units are graphed on the x-axis. A cell dilution curve (black squares) was used to derive a two-phase association best-fit line of expected values. Dashed orange lines represent cell dilution curve and 1.5 standard deviations above and below the predicted curve. B) Screening fitness was assessed by Z-factor. Non-targeting and *Irf3* control siRNA supernatants were tested for Cxcl10 induction at the upper (15ul) and lower (4.5ul) limits of detection. C) Top, ELISA of log2 Cxcl10 (pg/ml) production was plotted against ELISA of log2 Cxcl10 (pg/ml) production in replicate plates showing robustness of siRNA screening assay. Bottom, cell viability (RLU) after siRNA knockdown was plotted against cell viability (RLU) after siRNA knockdown in replicate plates showing robustness of siRNA treatment in screening assay. R2 values are shown.

A



B

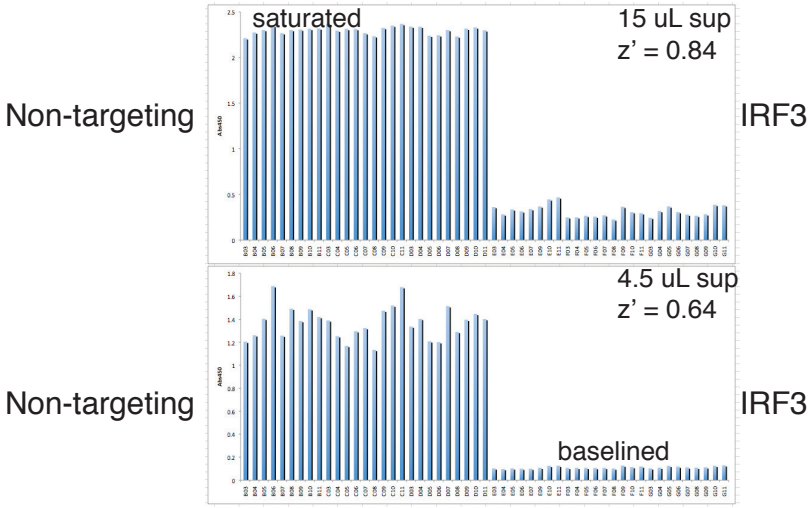
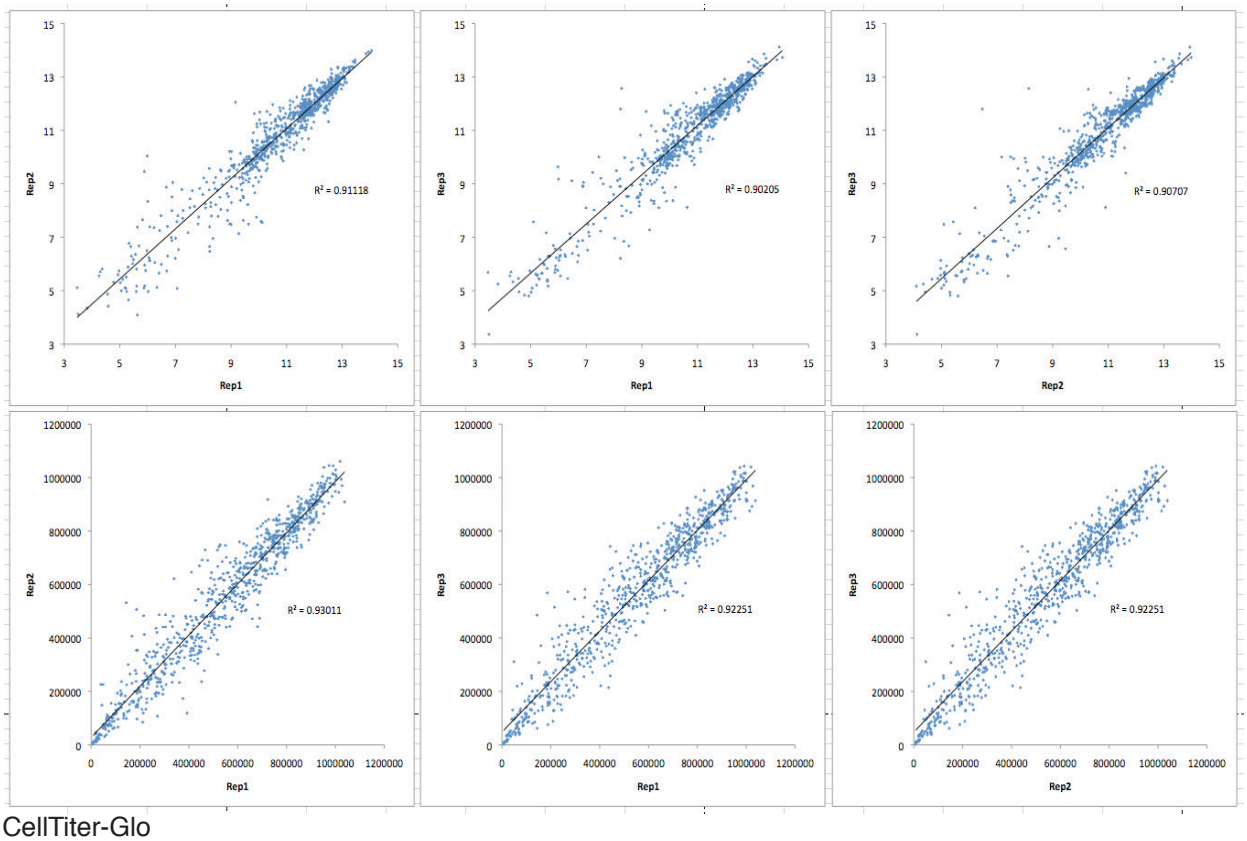


Figure 4.2 (continued)

C

Log(2) CXCL10 pg/mL



CellTiter-Glo

Figure 4.2 (continued)

0.902 – 0.911 between plates for Cxcl10 values and 0.922 – 0.930 between plates for CTG values) (Figure 4.2C).

The entire screen (Figure 4.3A) separated by each of the annotated candidates lists (Figure 4.3B-G and Table 4.3A-F) were plotted along with controls and the cell dilution curve. In addition to control gene *Irf3*, known components of the ISD-sensing pathway including *Sting* and *Tbk1* represented strong hits with over 90% reduction in Cxcl10 following stimulation with ISD. The siRNA screen identified 20 genes that, upon their knockdown, resulted in over 75% less Cxcl10 production in response to ISD stimulation, including ISD or IFN-β regulated candidates *Asb13* and *Ifitm1*, DNA interactors (for example, *Abcf1*, *Ifit1* and *Reep4*), helicases *Ddx46*, *Ddx59* and *Ercc3*, the signaling molecule *Chuk* (*Ikka*) and phosphatases and deubiquitinases (including *Ctdspl*, *Mtmr3* and ubiquitin specific proteases *Usp43* and *Usp49*). Furthermore there were eight genes for which Cxcl10 was upregulated more than three-fold after knockdown, including interferon-regulated genes (for example, *Tiparp*), the signaling molecule *Ripk1* and phosphatases including, *Ppp6c, Ptpn1* and *Mdp1* as well as deubiquitinases (for example, *Usp12* and *Cyld*).

4.4 – *Database development*

Screening results were curated in a database to provide convenient access to information regarding each candidate (Figure 4.4). We calculated a non-linear Z-score by subtracting average Cxcl10 values from an expected value derived from the cell dilution curve, divided by the standard deviation. The Z-scores were ranked and used as primary sorting field for the database. Screening data are presented for each candidate as CellTiter-Glo and Cxcl10 averages with fold-change relative to a negative control. A screenshot of the actual replicate data points is also included. Expression data from ISD

Figure 4.3

High-throughput Loss-of-Function RNAi Screen: **Full screening results.** A) Non-targeting and *Irf3* siRNA treated wells (green and red triangles, respectively) buffer-only controls (blue triangles) and candidate genes (black circles) were stimulated with ISD. Cxcl10 was measured by ELISA. Log(2) Cxcl10 (pg/mL) is graphed on the y-axis. Cell viability after knockdown was measured by CellTiter-Glo; relative luminescence units are graphed on the x-axis. A cell dilution curve (black squares with orange outlines) was used to derive a two-phase association best-fit line of expected values. Dashed orange lines represent cell dilution curve and 1.5 standard deviations above and below the predicted curve. B-G) Screening results separated by annotated candidate group: Microarry, DNA SILAC, helicases, cytoplasmic DNA-binding proteins, signaling molecules and selected phosphatases and deubiquitinases from pilot screen, respectively.

A

Figure 4.3 (continued)

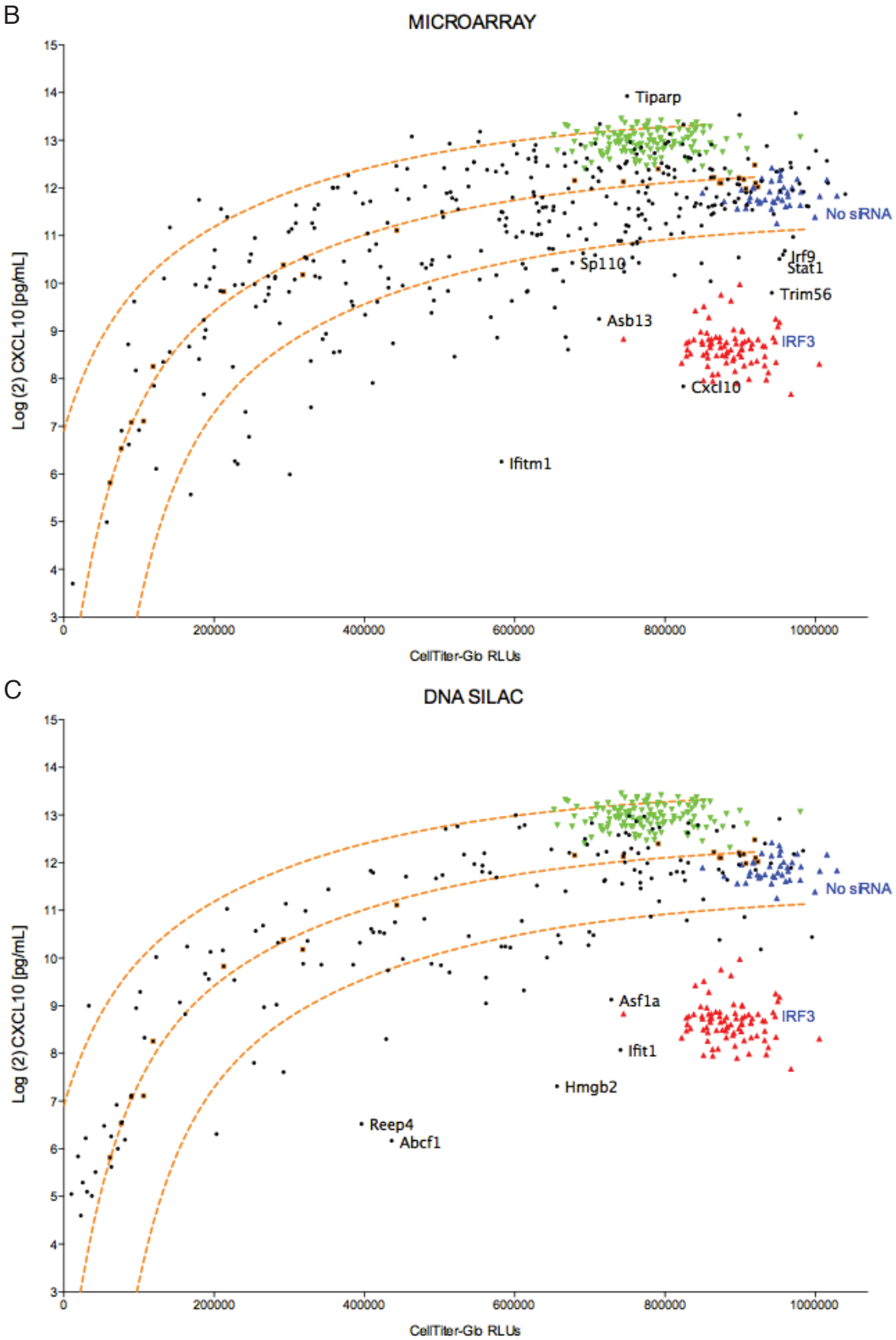


Figure 4.3 (continued)

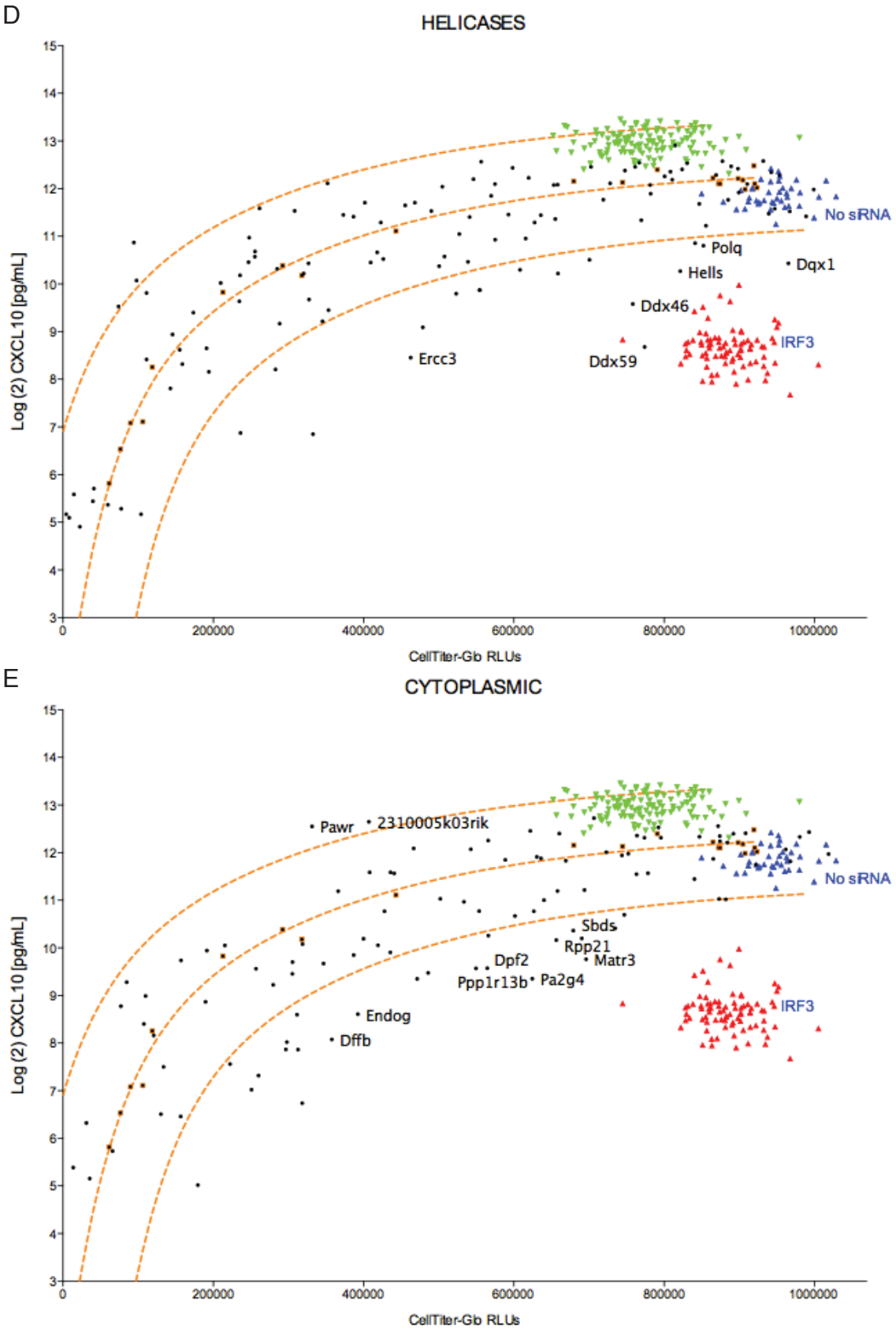
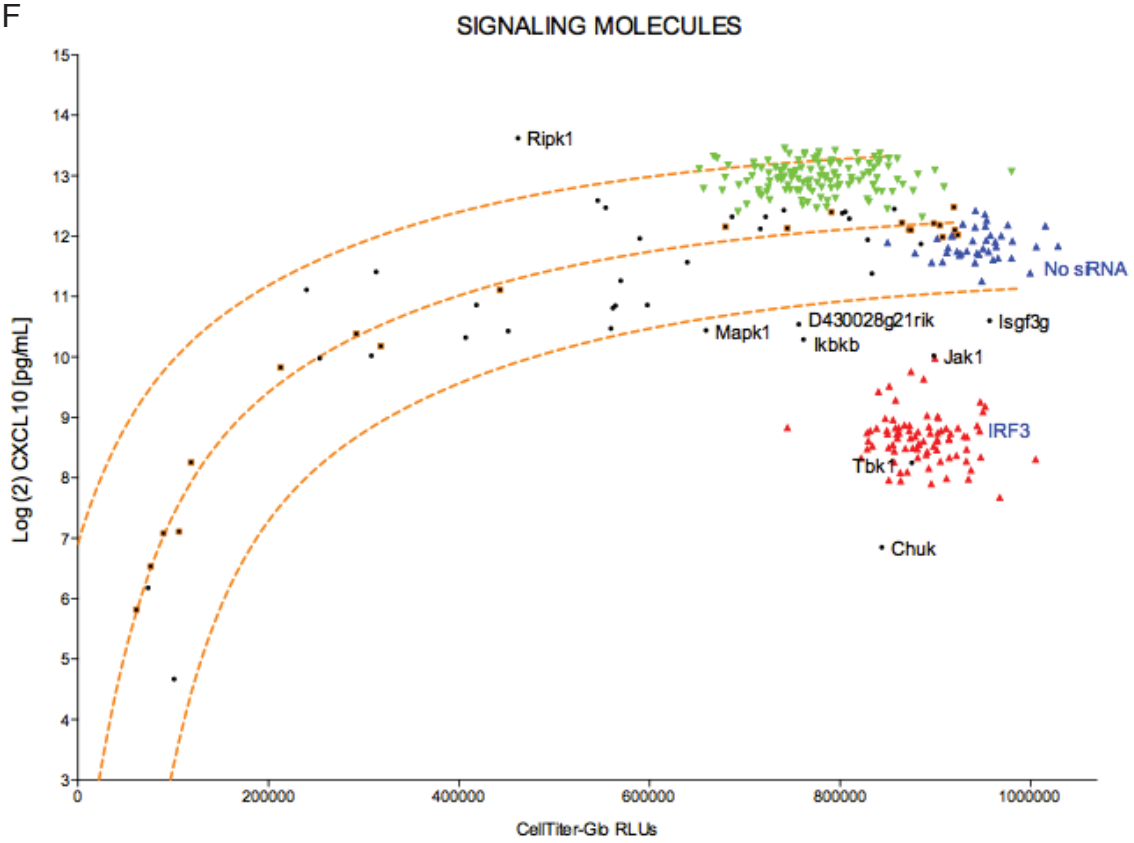


Figure 4.3 (continued)

F



G

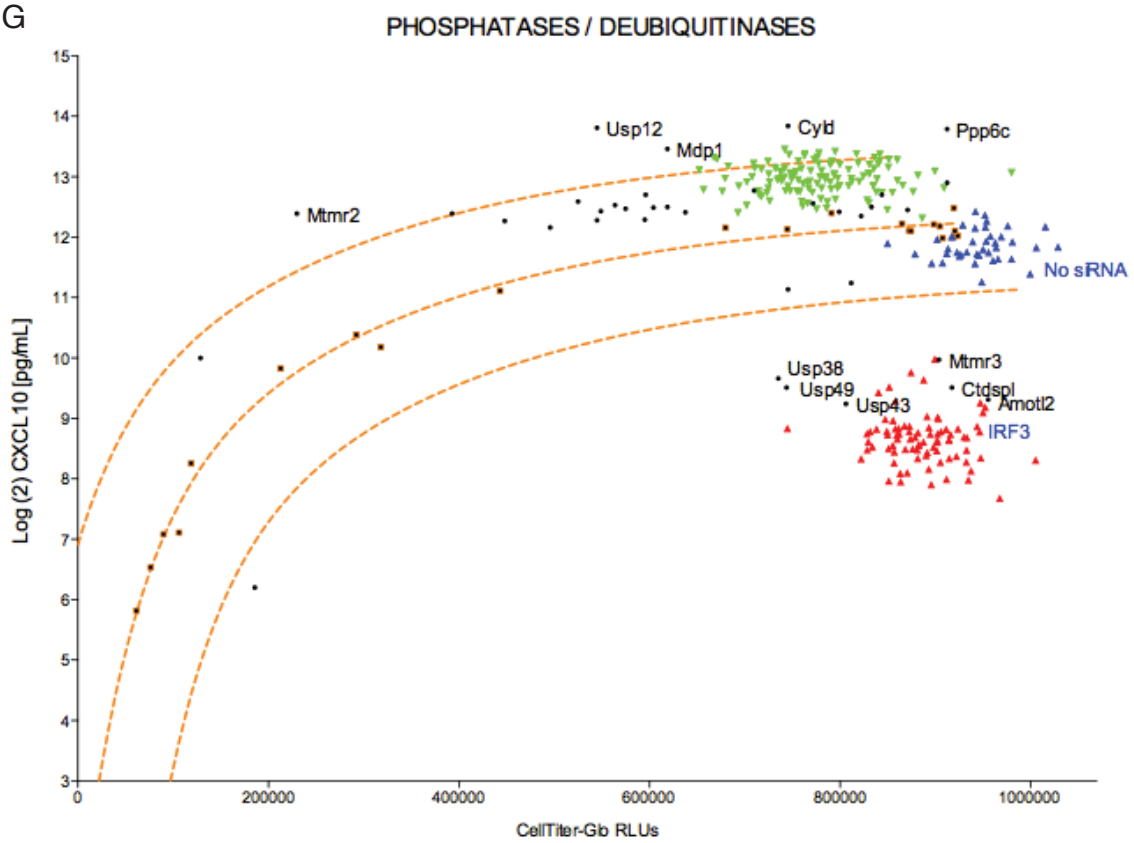


Figure 4.3 (continued)

Table 4.3 A
High-throughput Loss-of-Function RNAi Screen: Microarray

Gene Name	GeneID	CellTiter-Glo	CXCL10 (log_2, pg/mL)
No siRNA		945520	11.86
Irf3	54131	886771	8.66
1110003E01Rik	68552	622971	11.05
1110049F12Rik	66193	333550	11.56
1700010I14Rik	66931	254217	10.95
1700019G17Rik	75541	489101	10.39
4930547C10Rik	68274	629676	11.01
4933411K20Rik	66756	208430	9.84
9930111J21Rik	245240	695514	10.92
A230050P20Rik	319278	952951	11.94
Abhd11	68758	227566	6.27
Abhd3	106861	857853	11.60
Abtb2	99382	364495	12.01
Adar	56417	967911	12.43
Agrn	11603	762657	10.87
Agtrl1	23796	776544	12.02
AI451617	209387	860778	10.04
Aim1	11630	691219	10.63
Akt3	23797	651929	11.61
Alpk2	225638	734377	11.67
Apobec1	11810	666161	12.72
Apobec3	80287	575554	10.17
Arhgap9	216445	861944	11.44
Arid5a	214855	583754	12.48
Asb13	142688	712364	9.25
Asb14	142687	179894	8.41
Atp1b4	67821	655891	10.29
AY036118	170798	589310	11.08
B2m	12010	808280	11.03
Batf2	74481	358390	12.00
Baz2a	116848	238351	9.98
BC006779	229003	513478	12.93
Bptf	207165	779800	12.19
Brd4	57261	481951	10.32
Bst2	69550	618283	12.61
Btn1a1	12231	798103	12.31
Cabp1	29867	632222	11.51
Carhsp1	52502	340480	10.12
Casp7	12369	712238	12.42
Ccnb3	209091	812937	12.22
Cd274	60533	636840	11.80
Cd40	21939	596233	9.63
Cdadc1	71891	847505	11.66
Cenpj	219103	707149	10.59
Chek1	12649	589089	11.37
Chmp4c	66371	959464	10.68
Chmp5	76959	357976	12.01
Chn2	69993	909186	11.64
Clec2d	93694	860034	12.40
Cntn3	18488	480637	12.03
Copg2	54160	516824	11.79
Cops5	26754	246423	6.78
Cpeb3	208922	749199	12.56
Cphx	105594	687028	11.15
Csf1	12977	892342	11.71
Csprs	114564	175198	9.97
Cxcl10	15945	824505	7.84
D14Ertd668e	219132	372781	10.46
D1Pas1	110957	307301	10.08
Dab1	13131	800717	12.93
Daf2	13137	764737	11.42
Daxx	13163	711473	12.14
Dcpp1	13184	241912	7.30
Ddx3y	26900	249183	10.61
Ddx4	13206	513053	10.56
Ddx58	230073	166904	8.67
Dhx58	80861	378625	12.26
Ddx60	234311	862904	12.38
Dnase1l3	13421	330519	11.26
Ear5	54159	227534	9.96
Gm8995	668139	432088	10.10
Egfbp2	13647	861792	12.91
Eif2ak2	19106	188896	9.02
Elf1	13709	893660	11.70
Gm4836	22526	338428	11.65
Epm2aip1	77781	795583	12.52
Epsti1	108670	630676	11.02
Fam76b	72826	823445	12.92
Fas	14102	602820	12.72
Fbxl21	213311	538492	12.03
Foxa1	15375	770260	12.14
Galntl5	67909	209269	11.56
Gas7	14457	492498	9.64
Gbp1	14468	769162	11.45
Gbp2	14469	367108	8.57
Gbp3	55932	762919	10.84
Gbp4	17472	602249	10.97
Gbp5	229898	935426	12.88
Gbp6	229900	702127	11.99
Gbp9	236573	86616	6.62
Gca	227960	334541	10.15
Gch1	14528	511336	9.84
Gdap10	14546	681332	12.54
Gimap4	107526	255643	9.51
Glud1	14661	759506	12.81
Gngt1	14699	979035	12.09
Gpr137b	83924	713375	13.34
Gpr34	23890	671012	8.61
Gvin1	74558	774374	11.89
Hap1	15114	738386	12.45
Herc5	67138	509426	11.89
Higd1b	75689	737976	10.89
Hmg20a	66867	751720	11.20
Hmx3	15373	806384	13.13
Hopx	74318	272792	9.91
Hormad1	67981	194188	10.09
Hoxd13	15433	820280	11.57
Hsp90b1	22027	408000	11.72
Ifi203	15950	359386	8.55
Ifi204	15951	747272	12.51
Ifi205	226695	914170	11.91
Ifi27	76933	321026	10.55
Ifi35	70110	189528	9.93
Ifi44	99899	666608	11.47
Ifi47	15953	186496	7.67
Ifih1	71586	861028	12.22
Ifit2	15958	133398	10.10
Ifit3	15959	628079	12.64
Ifitm1	68713	582417	6.26
Ifitm2	80876	621922	12.22
Ifitm3	66141	200794	10.70
Igtp	16145	384972	9.53
Ihpk1	27399	916120	11.54
Iigp1	60440	563306	12.38
Iigp2	54396	846183	12.69
Il6	16193	756841	10.82
Insl6	27356	745690	10.24
Insm1	53626	323289	10.52
Irf2	16363	344555	8.83
Irf7	54123	677288	11.70
Irf9	16391	957040	10.60
Irgm	15944	645866	10.73
Isg20	57444	443497	11.96
Jam2	67374	786189	12.73
Jarid2	16468	804460	12.97
Kcne3	57442	753714	12.96
Klhl10	66720	907919	11.90
Klk1b16	16615	863107	11.63
Klk1b5	16622	519988	8.46
Klra16	27424	741652	10.89
Klra8	16639	401152	9.34
Lamp3	239739	991133	12.42
Lima1	65970	628132	12.55
Lin7b	22342	778454	11.73
Mcl1	17210	132281	8.35
Mef2c	17260	186496	9.23
Mela	17276	476356	10.41
Mov10	17454	506074	12.42
Mpa2l	100702	824552	13.33
Mpzl2	14012	855094	11.99
Mrps15	66407	555098	10.68
Ms4a4d	66607	984414	11.92
Mup4	17843	1039856	11.87
Mx1	17857	818464	11.40
Mx2	17858	299299	11.08
Mycl1	16918	456883	12.40
Myd88	17874	545668	12.54
Myef2	17876	1016931	12.10
Myt1	17932	766716	10.94
N4bp1	80750	451017	9.95
Nfxl1	100978	616138	9.82
Nlgn2	216856	514608	11.91

Table 4.3 A

High-throughput Loss-of-Function RNAi Screen: Microarray (continued)

Gene Name	GeneID	CellTiter-Glo	CXCL10 (log_2, pg/mL)
Nmi	64685	866355	11.64
Nphp3	74025	879664	12.41
Nr0b1	11614	908374	11.64
Nrtn	18188	553297	10.08
Ntng2	171171	778319	12.38
Nucks1	98415	824098	11.06
Nudt13	67725	883566	12.18
Numb	18222	826336	12.66
Oas1a	246730	841320	11.68
Oas1b	23961	610630	11.08
Oas1c	114643	643947	12.46
Oas1g	23960	875417	11.65
Oas2	246728	76748	6.91
Oas3	246727	199817	10.34
Oasl1	231655	629233	11.68
Oasl2	23962	287378	9.16
Ogfr	72075	667174	12.46
Olfm3	229759	489297	9.38
Oog1	193322	948235	12.34
Oplah	75475	967039	12.53
Pank2	74450	626055	11.84
Parp11	101187	415740	10.82
Parp12	243771	95965	8.17
Parp14	547253	898116	10.54
Parp9	80285	625718	12.79
Pax6	18508	667857	8.87
Pcbp2	18521	382768	9.62
Pbld2	67307	831332	12.64
Pdlim5	56376	610538	10.41
Pdx1	18609	406745	10.69
Peli1	67245	432051	9.91
Pex13	72129	707461	11.24
Phc3	241915	664454	10.04
Phex	18675	627895	12.90
Phf11	219131	421557	10.35
Phip	83946	764619	11.92
Phyhipl	70911	388043	11.90
Pigb	55981	1015601	12.57
Pik3c2g	18705	217498	11.39
Pla2g1b	18778	463113	13.08
Plac8	231507	879663	11.56
Plagl1	22634	300736	5.99
Plec1	18810	582278	11.50
Plekha4	69217	333320	11.16
Plk2	20620	11794	3.70
Plscr1	22038	536707	12.39
Plscr2	18828	598618	12.90
Pml	18854	442623	12.46
Pnp	18950	522037	11.75
Pnpla2	66853	587626	11.21
Pnpt1	71701	715127	12.36
Pnrc1	108767	751899	11.57
Podxl	27205	85720	8.72
Pou4f3	18998	280999	10.10
Ppm1a	19042	398060	11.67
Ppm1k	243382	704228	12.23
Prg2	19074	1000833	12.78
Psmb9	16912	783409	11.67
Psme1	19186	551584	12.97
Psrc1	56742	504843	11.39
Pvr	52118	762486	12.32
Pyhin1	236312	762328	12.71
Qk	19317	574878	12.34
Rab27a	11891	898937	12.07
Rab3c	67295	576793	8.86
Rai2	24004	93687	9.62
Rasgef1b	320292	848656	10.42
Rbck1	24105	410992	7.91
Rbm43	71684	898089	12.19
Rnf151	67504	486895	9.90
Rnf31	268749	524896	11.55
Rs1	20147	417205	9.61
Rsad2	58185	740467	12.42
Rshl1	83434	696178	12.62
Rsrc1	66880	568229	9.29
Rtp4	67775	637185	11.59
Samd9l	209086	839221	11.61
Sap30l	50724	225033	8.25
Scotin	66940	328723	8.37
Sec1	56546	804651	12.28
Serpina3m	20717	57360	4.99
Serpina6	12401	896298	12.08
Serpinb9	20723	707277	12.93
Serpini1	20713	848523	11.21
Shh	20423	503371	11.50
Slc1a3	20512	768230	12.85
Slc25a22	68267	271956	10.15
Slc25a28	246696	281271	10.14
Slc3a1	20532	635425	11.03
Slc6a14	56774	299739	9.82
Slc9a8	77031	970460	10.97
Slco1a6	28254	404338	11.05
Slfn10	237887	140896	8.56
Slfn2	20556	671520	11.99
Slfn3	20557	801240	11.25
Slfn4	20558	100099	6.92
Slfn5	327978	589349	12.32
Slfn8	276950	864810	11.47
Slfn9	237886	122982	6.11
Socs2	216233	812804	10.17
Sp100	20684	815062	11.87
Sp110	109032	676793	10.43
Sri	109552	988446	11.62
Stat1	20846	952376	10.51
Stat2	20847	267958	9.63
Stk31	77485	653389	9.49
Stk32c	57740	812730	12.38
Stk4	58231	430857	11.27
T2bp	211550	438109	8.74
Tap1	21354	597027	11.80
Tap2	21355	650915	10.72
Tapbp	21356	748280	11.22
Tbc1d10a	103724	641639	11.60
Tc2n	74413	773731	11.33
Tcirg1	27060	897895	11.78
Tcte3	21647	671322	11.00
Tdrd3	219249	308988	11.64
Tdrd7	100121	579683	11.34
Tgtp	21822	910768	12.61
Thrsp	21835	332374	10.47
Tiparp	99929	749564	13.93
Tlr3	142980	120136	7.85
Tmem106a	217203	231492	6.21
Tmem140	68487	928344	12.14
Tmem150c	231503	722485	11.82
Tmem229b	268567	651580	10.28
Tmem79	71913	838806	12.37
Tnc	21923	179956	11.75
Tnfaip3	21929	885004	11.87
Tnp1	21958	646111	11.30
Tor1aip1	208263	583996	11.49
Tor1aip2	240832	765239	10.43
Tor3a	30935	956014	12.38
Tra2a	101214	632183	11.24
Trafd1	231712	490427	10.49
Trim12c	319236	747450	11.08
Trim14	74735	991001	11.61
Trim21	20821	308798	11.26
Trim25	217069	458718	11.77
Trim26	22670	779900	11.15
Trim27	19720	804956	12.43
Trim30	20128	375426	10.00
Trim34	94094	834299	12.62
Trim56	384309	942177	9.80
Trim69	70928	617258	9.84
Twf1	19230	168922	5.57
Txndc10	67988	614666	10.31
Tyki	22169	973778	13.57
Ube1l	74153	687827	12.91
Ube2l6	56791	625045	11.91
Unc93b1	54445	969884	11.95
Usp18	24110	328952	7.40
Usp25	30940	348854	8.94
V1rd2	81016	944655	11.45
Vdac1	22333	651235	11.53
Vnn1	22361	187054	8.86
Vtcn1	242122	303170	11.14
Wdr78	242584	695988	12.39
Zbp1	58203	764439	11.56
Zc3hav1	78781	868569	11.81
Zfp313	81018	141227	11.17
Zkscan14	67235	899147	13.53

Table 4.3 B
High-throughput Loss-of-Function RNAi Screen: DNA SILAC

Gene Name	GeneID	CellTiter-Glo	CXCL10 (log_2, pg/mL)
No siRNA		945520	11.86
Irf3	54131	886771	8.66
1810029B16Rik	66282	898550	11.96
2810432D09Rik	69961	226980	9.54
8430406I07Rik	74528	324319	10.36
8430410A17Rik	232210	164207	10.24
Abcf1	224742	436318	6.17
Alkbh2	231642	781249	10.87
Anp32b	67628	891206	12.76
Anxa4	11746	594207	10.22
Anxa5	11747	642998	10.01
Apex1	11792	648863	12.28
Aplf	72103	612191	9.32
Aptx	66408	843595	12.79
Ascc1	69090	899815	12.15
Ascc2	75452	933005	12.40
Ascc3	77987	920689	11.85
Asf1a	66403	728637	9.13
Asf1b	66929	561790	9.59
Cggbp1	106143	292498	7.61
Cirbp	12696	411695	11.79
Cnn2	12798	161480	8.82
Crcp	12909	704860	10.48
Creb1	12912	561847	9.05
Creb3l1	26427	411838	10.54
Creb3l2	208647	995551	10.44
Dazap1	70248	479852	10.82
Ddb1	13194	188136	9.67
Ddb2	107986	764434	12.20
Dek	110052	820898	12.05
Dhx36	72162	255296	10.57
Dnajc9	108671	792280	11.19
Dr1	13486	807828	11.77
Drap1	66556	699328	10.55
Drg1	13494	96289	8.95
Eef1d	66656	601563	13.00
Eif2s1	13665	37353	5.01
Eif2s2	67204	22633	4.60
Eif2s3x	26905	63032	6.26
Eif5b	226982	89914	7.12
Fbxo18	50755	546232	11.97
Fen1	14156	557238	11.94
Frg1	14300	657963	10.48
Gata4	14463	850969	11.23
Get4	67604	752239	12.58
Gnb2l1	14694	122989	10.02
H1fx	243529	30898	5.10
Hist1h1a	80838	217293	11.03
Hist1h1b	56702	772476	12.97
Hist1h1c	50708	508069	12.71
Hist1h1e	50709	755248	12.30
Hmga1	15361	501887	9.85
Hmga2	15364	815652	11.81
Hmgb1	15289	767225	11.58
Hmgb2	97165	656372	7.31
Hmgb3	15354	830666	12.63
Hmgn1	15312	575799	11.69
Hnrnpu	51810	42304	5.51
Hnrpa3	229279	983795	12.25
Hnrpab	15384	532545	12.17
Hnrpd	11991	203355	6.31
Ifi202b	26388	101660	9.29
Ifit1	15957	813719	12.12
Kif22	110033	629769	11.52
Kif2a	16563	872593	10.38
Mdh1	17449	408462	11.83
Mdh2	17448	792701	12.74
Morf4l2	56397	868922	12.67
Mpg	268395	385051	9.86
Msh2	17685	830878	12.00
Msh3	17686	777511	12.31
Msh6	17688	710868	12.17
Msi2	76626	831157	12.83
Mtap1b	17755	342664	9.86
Mtap4	17758	567121	11.79
Mybbp1a	18432	107391	8.33
Nagk	56174	722064	12.23
Neil1	72774	698294	10.27
Nfib	18028	871568	11.86
Nthl1	18207	759335	11.85
Nuak2	74137	606236	12.74
Obfc1	108689	882905	11.64
Ogg1	18294	440781	10.75
Parp1	11545	696445	12.38
Parp2	11546	495252	11.67
Parp3	235587	420347	10.55
Pcna	18538	265036	10.68
Pdia3	14827	195208	10.13
Phf6	70998	693461	12.50
Pkm2	18746	351983	11.44
Pnkp	59047	829014	10.79
Polb	18970	905836	10.86
Pold2	18972	33411	9.00
Pole3	59001	766904	11.80
Polr3a	218832	63248	5.62
Polr3b	70428	266540	8.97
Polr3f	70408	53710	6.48
Prkar2b	19088	77542	6.56
Prrx1	18933	967833	11.89
Psip1	101739	431376	9.74
Pycr2	69051	747930	12.63
Rabggtb	19352	743104	11.56
Rbm28	68272	513202	9.70
Rbm39	170791	25132	5.29
Rbms1	56878	667531	11.40
Rbms2	56516	503962	11.74
Rbms3	207181	824174	11.80
Rbpj	19664	776597	11.99
Recql	19691	782063	11.67
Reep3	28193	752344	12.98
Reep4	72549	396200	6.52
Rfc1	19687	19021	5.84
Rfc4	106344	941082	11.99
Rpa1	68275	10061	5.05
Rpa2	19891	29230	6.22
Rpa3	68240	70375	6.92
Rrbp1	81910	661163	10.32
Rsl1d1	66409	283078	9.02
Samhd1	56045	253250	7.80
Serbp1	66870	971441	12.18
Sfrs2	20382	81501	6.19
Sfrs3	20383	71897	6.00
Sfrs7	225027	409102	10.61
Smarcal1	54380	285000	10.32
Snrpa	53607	489258	9.88
Snx9	66616	952324	12.92
Sub1	20024	927505	10.18
Tbp	21374	154417	9.07
Tcfe3	209446	826303	11.63
Tcfeb	21425	745699	12.20
Tdg	21665	455305	11.71
Tdp1	104884	429013	8.30
Tead1	21676	738063	12.58
Tfam	21780	761936	12.87
Themis	210757	672618	12.21
Thex1	67276	318535	9.88
Thoc4	21681	295545	11.14
Thyn1	77862	192873	9.56
Top1	21969	582200	10.24
Top2a	21973	556680	11.81
Trex1	22040	559668	12.20
Trip4	56404	609616	10.78
Trmt6	66926	767673	11.46
Trmt61a	328162	613226	12.79
Tsnax	53424	523860	12.76
Ttf2	74044	702655	12.83
Ubtf	21429	211731	10.16
Usp39	28035	321735	10.99
Vrk1	22367	746806	12.01
Wdr76	241627	451104	9.98
Wrn	22427	426305	10.52
Xbp1	22433	393908	10.32
Xpa	22590	720436	11.64
Xrcc1	22594	587468	10.24
Xrcc5	22596	718991	11.61
Xrcc6	14375	538894	10.46
Zc3h15	69082	783636	12.71
Znfx1	98999	876984	11.79

Table 4.3 C
High-throughput Loss-of-Function RNAi Screen: Helicases

GeneName	GeneID	CellTiter-Glo	CXCL10 (log_2, pg/mL)
No siRNA		945520	11.86
Irf3	54131	886771	8.66
0610007p08rik	76251	910729	12.10
1810014j18rik	109151	658248	12.08
2310061o04rik	69663	155434	8.62
2410004f06rik	71981	808497	12.36
2610007k22rik	67040	800715	12.26
2610528a15rik	72198	326792	10.43
2610528e23rik	66497	761436	12.46
2810457m08rik	234733	8358	5.10
4930422g04rik	71643	627764	11.29
5430439g14rik	71389	423056	11.29
6030422m02	240697	386575	11.41
6330505f04rik	236790	767067	12.55
A330009g12rik	330149	504878	12.04
A330064g03rik	320632	77415	5.29
A930037j23rik	269254	700587	10.51
Ai449441	208084	556436	12.57
Ascc3	77987	899329	11.92
Atrx	22589	846812	11.68
Aw494914	106794	967105	11.52
Aw540478	269400	814865	12.91
B830009d23rik	96957	308350	11.53
Bat1a	53817	59701	5.37
Bc004701	236930	830709	12.53
Bc019206	216161	877633	12.58
Blm	12144	728104	12.11
Brip1	237911	353386	9.45
C130058g22rik	319955	190950	8.65
Chd1	12648	575249	10.93
Chd1l	68058	810963	12.19
Chd2	244059	615637	10.95
Chd3	216848	142930	7.81
Chd4	107932	282879	8.21
Chd8	67772	898194	12.42
Ddx1	104721	635994	11.44
Ddx11	320209	574877	12.10
Ddx18	66942	22418	4.91
Ddx19	13680	14503	5.58
Ddx20	53975	619683	12.23
Ddx21	56200	401697	11.70
Ddx24	27225	41034	5.71
Ddx25	30959	569914	11.85
Ddx26	18130	173614	9.40
Ddx28	71986	351987	12.11
Ddx39	68278	953505	12.28
Ddx3x	13205	554494	9.87
Ddx41	72935	247041	10.46
Ddx42	72047	939339	11.47
Ddx46	212880	758214	9.58
Ddx47	67755	409583	10.45
Ddx48	192170	4241	5.17
Ddx5	13207	869154	12.29
Ddx50	94213	855763	11.22
Ddx52	78394	523205	9.80
Ddx54	71990	111236	9.81
Ddx55	67848	769520	11.34
Ddx56	52513	158920	8.32
Ddx59	67997	774107	8.68
Ddx6	13209	478935	9.09
Dhx15	13204	103906	5.17
Dhx16	69192	236181	6.88
Dhx29	218629	823996	12.41
Dhx30	72831	247718	10.97
Dhx32	101437	841199	10.86
Dhx33	216877	658754	10.22
Dhx34	71723	889125	12.47
Dhx35	71715	653287	12.08
Dhx36	72162	255296	10.57
Dhx37	208144	327515	9.68
Dhx38	64340	345361	9.22
Dhx40	67487	988873	11.42
Dhx8	217207	39772	5.44
Dhx9	13211	999108	11.98
Dqx1	93838	965443	10.43
E130016e03rik	623474	931886	12.58
E130315b21rik	327762	468222	11.70
E430027o22rik	107182	111212	8.42
Eif4a1	13681	97821	10.07
Eif4a2	13682	500603	10.37
Ep400	75560	94413	10.87
Ercc2	13871	332805	6.85
Ercc3	13872	462652	8.46
Fbxo18	50755	546232	12.20
Fin14	13205	554494	9.87
G22p1	14375	538894	10.46
Hells	15201	821512	10.27
Ighmbp2	20589	209991	10.02
Mcm2	17216	255622	10.68
Mcm3	17215	541156	11.41
Mcm4	17217	608123	10.30
Mcm5	17218	234675	9.64
Mcm6	17219	655205	11.36
Mcm7	17220	527445	11.05
Mcm8	66634	747850	12.38
Mov10l1	83456	942592	12.34
Peo1	226153	194107	8.16
Polq	77782	852246	10.80
Rad54l	19366	261634	11.59
Recql	19691	782063	11.89
Recql4	79456	885122	11.77
Recql5	170472	288420	9.17
Rent1	19704	507727	10.58
Ruvbl1	56505	145733	8.94
Ruvbl2	20174	235655	10.18
Shprh	268281	373224	11.45
Skiv2l	108077	598387	12.44
Smarca2	67155	592819	11.46
Smarca3	20585	320297	10.22
Smarca4	20586	418080	10.67
Smarca5	93762	947355	11.58
Smarcad1	13990	781267	12.08
Smarcal1	54380	285000	10.32
Srisnf2l	81000	490230	11.53
Supv3l1	338359	74001	9.53
Ttf2	74044	702655	12.46
Wrn	22427	426305	10.52
Xrcc5	22596	718991	11.77

Table 4.3 D

High-throughput Loss-of-Function RNAi Screen: Cytoplasmic DNA-binding proteins

GeneName	GeneID	CellTiter-Glo	CXCL10 (log_2, pg/mL)
No siRNA		945520	11.86
Irf3	54131	886771	8.66
0610010i17rik	114774	723071	12.01
1700051e09rik	83561	466794	12.09
2310005k03rik	227715	406971	12.65
2410006f12rik	69537	133884	7.50
2810028n01rik	67338	882112	11.02
4930517k23rik	19878	735138	10.41
4932442k20rik	170711	553902	10.77
4933406l09rik	58523	408088	11.59
5830483c08rik	27998	107690	8.40
A230103n10rik	17984	179534	5.02
Ang4	225207	636093	11.88
Ankrd3	74430	312734	7.86
Arc	234258	847010	12.33
Bat2	231999	533618	10.97
Bc034753	77634	157015	9.74
Bc052360	105559	435837	11.59
Bin1	12192	640224	11.00
Bzw1	13870	305440	9.70
Bzw2	19355	961846	12.07
Cdc5l	69639	297998	8.02
D11ertd497e	18628	588650	11.85
D630024b06rik	52626	865331	11.87
D7wsu87e	18037	630572	11.91
Dffb	360216	357858	8.08
Dpf2	66362	564805	9.57
Dsip1	11838	746985	10.70
Ear1	219033	873171	11.03
Ear2	71592	156744	6.46
Ear3	209334	743865	11.94
Ear4	53877	874085	12.35
Endog	26919	392433	8.61
Eps15	66847	778472	11.57
Ercc1	105083	763935	12.36
Exosc2	58180	77068	8.77
Exosc3	66230	215466	10.05
Exosc4	64383	296691	7.87
Exosc5	20183	109780	8.99
Exosc6	53876	65918	5.73
Exosc8	231872	191556	9.94
Exosc9	72544	346831	9.67
Hic2	66912	874394	12.24
Jarid1d	83435	189825	8.87
Jtv1	224829	980551	12.33
Lin28	218973	311436	8.60
Lsm8	20947	250836	7.02
Mapk14	83431	256902	9.56
Mapk7	13858	922082	11.75
Matr3	53761	696312	9.76
Mbd1	332397	386907	9.85
Mbd3l1	75029	679256	10.36
Mbd4	23939	840279	11.45
Mbnl2	22099	440851	11.57

GeneName	GeneID	CellTiter-Glo	CXCL10 (log_2, pg/mL)
Mlh1	17535	305315	9.45
Mre11a	26416	658186	11.20
Mrps28	19285	992884	12.43
Nanos1	17193	1018895	11.97
Ndel1	30948	763157	11.55
Ndn	231464	660440	12.40
Nfkbie	20592	795838	12.32
Otud7a	53890	621924	12.46
Pa2g4	66882	624573	9.35
Pawr	22687	331456	12.55
Pcbp3	22661	35516	5.16
Pelo	103135	669164	11.83
Per3	17350	773806	12.31
Plekha3	50911	502199	11.03
Pogk	17190	892251	12.40
Ppp1r13b	67673	549684	9.57
Ptrf	13586	693948	11.22
Purg	226442	318775	10.08
Rabgef1	72662	222419	7.56
Rad1	14605	871943	12.56
Rex3	22763	471462	9.35
Rock2	71957	706624	12.73
Rpp21	73503	656565	10.16
Rxrg	403171	884161	12.21
Sart3	67676	399914	10.19
Sbds	66711	690462	10.20
Sirt2	22165	752333	11.98
Snapc3	19716	792545	12.53
Ssbp2	21981	260277	7.32
Statip1	83557	565716	12.25
Swap70	19708	120784	8.16
Tceb2	17184	279749	9.22
Tdrd1	59016	84879	9.28
Thap11	13804	130266	6.51
Trerf1	18813	427848	10.77
Tsn	56715	968201	11.82
Ttf1	71702	31260	6.32
Txk	109075	486041	9.47
Usp52	72388	419169	10.05
Xab2	13368	13538	5.39
Zfp143	13587	318365	6.74
Zfp148	66970	601606	10.67
Zfp259	76522	626579	10.77
Zfp281	20841	435634	9.91
Zfp346	104625	908517	12.41
Zfp36l1	67439	366294	11.19
Zfp521	59093	542690	12.07
Zfr	22130	565955	10.26

Table 4.3 E

High-throughput Loss-of-Function RNAi Screen: Signaling Molecules

GeneName	GeneID	CellTiter-Glo	CXCL10 (log_2, pg/mL)
No siRNA		945520	11.86
IRF3	54131	886771	8.66
ASC	66824	716324	12.12
AW046014	106759	253827	9.98
BC034204	270151	418292	10.86
CARD15	257632	828969	11.94
CASP8	12370	597705	10.86
CHUK	12675	844002	6.85
D430028G21RIK	228607	756700	10.54
EYA4	14051	802501	12.38
FADD	14082	240107	11.11
Ikbkb	16150	761550	10.29
IKBKE	56489	308032	10.02
IKBKG	16151	73781	6.18
IRAK1	16179	809719	12.29
IRF3	54131	100987	4.67
ISGF3G	16391	957040	10.60
JAK1	16451	898386	10.02
JAK2	16452	451559	10.43
JAK3	16453	564355	10.85
MAPK1	26413	659396	10.44
MYD88	17874	545668	12.59
NFKBIA	18035	639737	11.57
NFKBIB	18036	721954	12.32
Pin1	23988	559680	10.47
RIPK1	19766	462036	12.98
RIPK3	56532	833591	11.38
SOCS1	12703	686699	12.32
STAT3	20848	589802	11.96
TANK	21353	553965	12.47
TBK1	56480	875356	8.25
TNFAIP3	21929	885004	11.87
TRAF1	22029	312947	11.41
TRAF2	22030	856914	12.45
TRAF3	22031	741107	12.43
TRAF4	22032	561503	10.81
TRAF5	22033	406982	10.32
TRAF6	22034	805470	12.40
TRAM	225471	569769	11.26

Table 4.3 F

High-throughput Loss-of-Function RNAi Screen: Selected phosphatases and deubiquitinases

GeneName	GeneID	CellTiter-Glo	CXCL10 (log_2, pg/mL)
No siRNA		945520	11.86
IRF3	54131	886771	8.66
1110012l19rik	68618	548870	12.43
2310043k02rik	66959	448140	12.27
A230072i16rik	217057	799065	12.42
Aa939927	99526	525051	12.59
Amotl2	56332	955543	9.31
C920001d21rik	320139	604289	12.49
Cdc14b	218294	771614	12.56
Ctdspl	69274	917571	9.51
Cyld	74256	745522	13.84
Dusp8	18218	811744	11.24
Gm1395	231637	709838	12.77
Itpkb	320404	596076	12.70
Mdp1	67881	618859	13.46
Mtmr2	77116	229776	12.39
Mtmr3	74302	903894	9.97
Mtmr4	170749	575008	12.47
Osr2	107587	595052	12.29
Ppm1g	14208	495791	12.16
Ppp1ca	19045	128748	10.00
Ppp6c	67857	912333	13.79
Ptpn1	19246	843810	12.70
Rasgrp1	19419	822119	12.35
Rngtt	24018	185651	6.20
Rps6ka1	20111	563747	12.53
Stambp	70527	870862	12.45
Uchl1	22223	833060	12.50
Usp12	22217	544907	13.81
Usp27x	54651	392922	12.39
Usp38	74841	735211	9.66
Usp43	216835	806159	9.24
Usp49	224836	743933	9.51
Usp54	78787	637640	12.41

Figure 4.4

High-throughput Loss-of-Function RNAi Screen: **Database development.** Screening results are catalogued in a searchable database. Basic screening results and gene characteristics including molecular function and biological process are displayed in the upper left hand corner. Replicate siRNAs are displayed in the context of the entire screen, upper right-hand side. Expression data are displayed in the top middle. On the top row in light blue, genes with human homology are displayed and are linked to genome-wide association study (GWAS) data based on chromosomal position. Gene expression across multiple mouse and human tissues can be accessed through the BioGPS Gene Portal radio button. Candidate genes are link directly to the specified NCBI Gene. Protein architecture, putative protein-protein interactions and post-translational modification data for each candidate are linked through the SMART database. Bottom table details Lentiviral shRNAs for each candidate, including glycerol stock location and validation data.

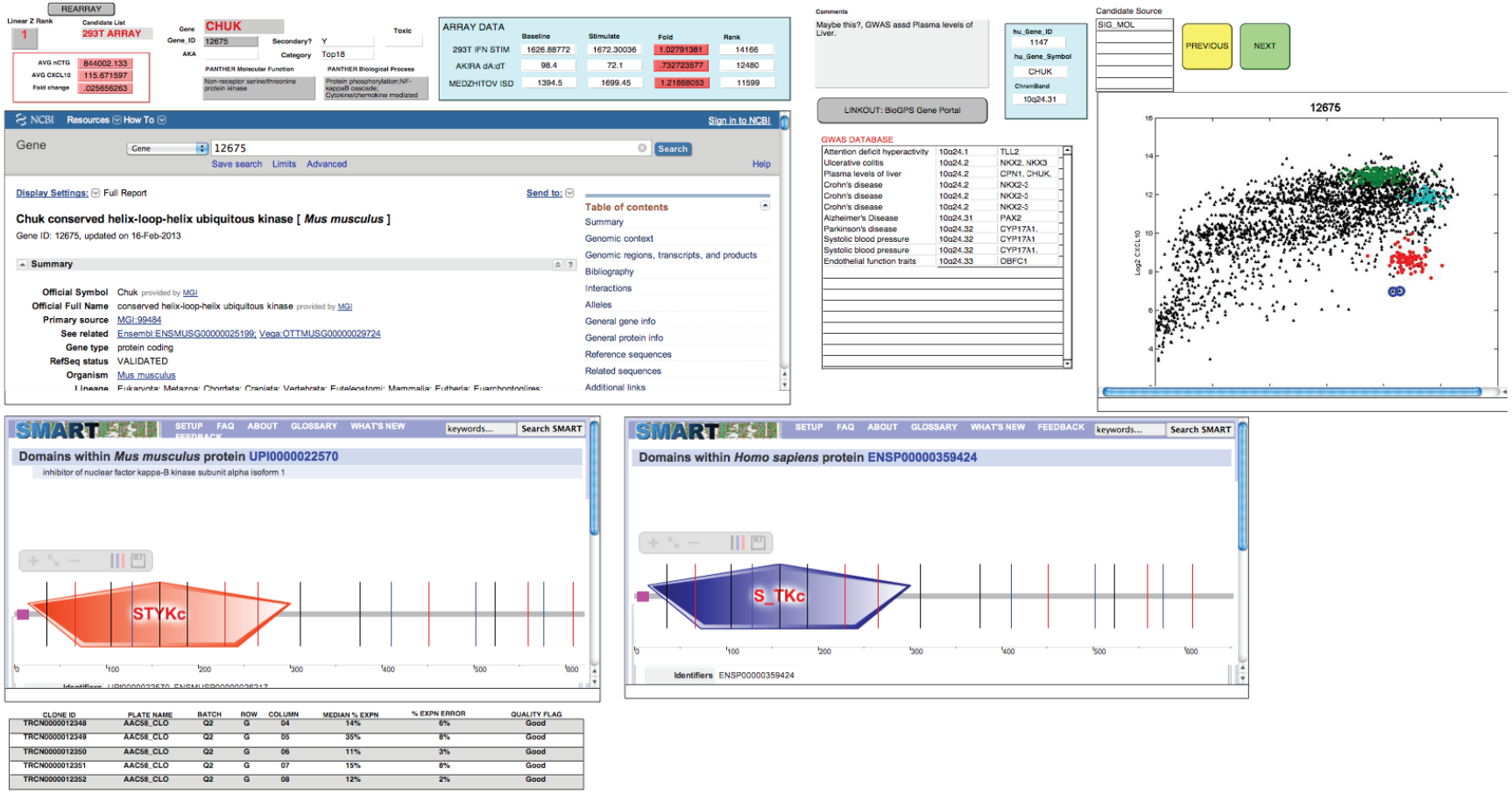


Figure 4.4 (continued)

and IFN-β stimulated microarrays are linked directly to each candidate[43, 44, 48]. To identify potential connections to human diseases, candidate genes were linked by homology to human genes. Drawing on a curated catalog of genome-wide association studies based on SNP-trait associations with p-values < 1.0 x 10-5, homologous genes were linked by nearest chromosomal band proximity[173]. Also, tissue-specific patterns of mRNA expression can be accessed for 79 human and 61 mouse tissues using the BioGPS portal[174]. Domain architecture, predicted functional partners and post-translation modification can be viewed for each candidate gene and its human homologue via the SMART (Simple Modular Architecture Research Tool) browser[175]. Additionally, location and knockdown validation of lentiviral shRNA clones available through the RNAi Consortium (Broad Institute) are provided.

4.5 – *Secondary screening*

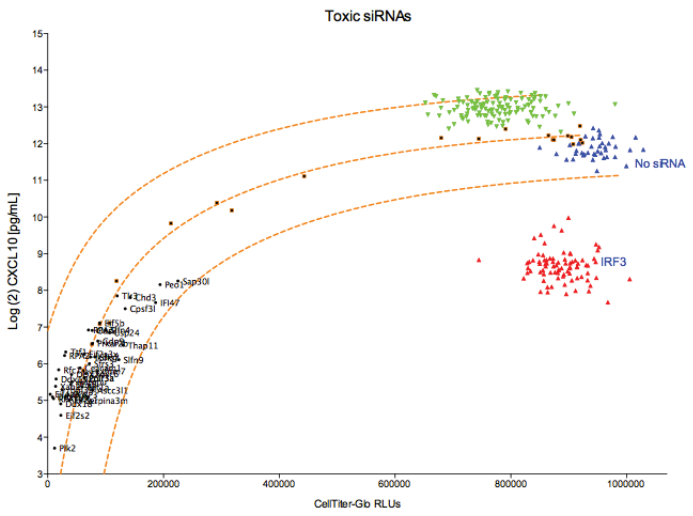
Secondary screening involved three distinct phases. First, we recovered siRNA candidates lost to toxicity. Next, in attempt to identify ISD-specific responses, we compared top candidates stimulated ISD or IVT-RNA. Third, we further interrogated candidates for ISD-specific responses with independent siRNA pools stimulated with ISD, Poly I:C, IVT-RNA or recombinant mouse IFN-β. The roles of the strongest candidates as regulators of the ISD pathway are dissected in the following chapter.

First we identified siRNAs from the full screen that had toxic effects. Toxicity could be caused by off-target effects of siRNAs, knockdown of an essential cellular component, or a reduction of a cell-intrinsic component required for the response to cytosolic DNA. Candidate siRNAs with CellTiter-Glo values below the 25th percentile were considered toxic (Figure 4.5A). To recover toxic candidates, siRNA pools were deconvoluted and knocked down separately. Following siRNA knockdown, samples

Figure 4.5

Deconvolution of toxic siRNA pools reveals additional ISD pathway candidates. A) Selection of toxic siRNA pools and controls from high-throughput screen for secondary screening with individual Dharmacon siGENOME siRNAs and stimulated with B) 1ug/mL ISD for 26 hours. CellTiter-Glo relative luminescent unites are on the x-axis, Log(2) Cxcl10 protein pg/mL, as detected by ELISA, are on the y-axis. Cxcl10 production was compared between stimuli. Control siRNAs (non-targeting controls, green triangles, *Irf3*, red triangles) are labeled, circles represent 1 of 4 individual siRNA candidates (red, orange, blue and green circles represent 1 of 4 deconvoluted siRNAs), strongest individual siRNAs are labeled.

A



B

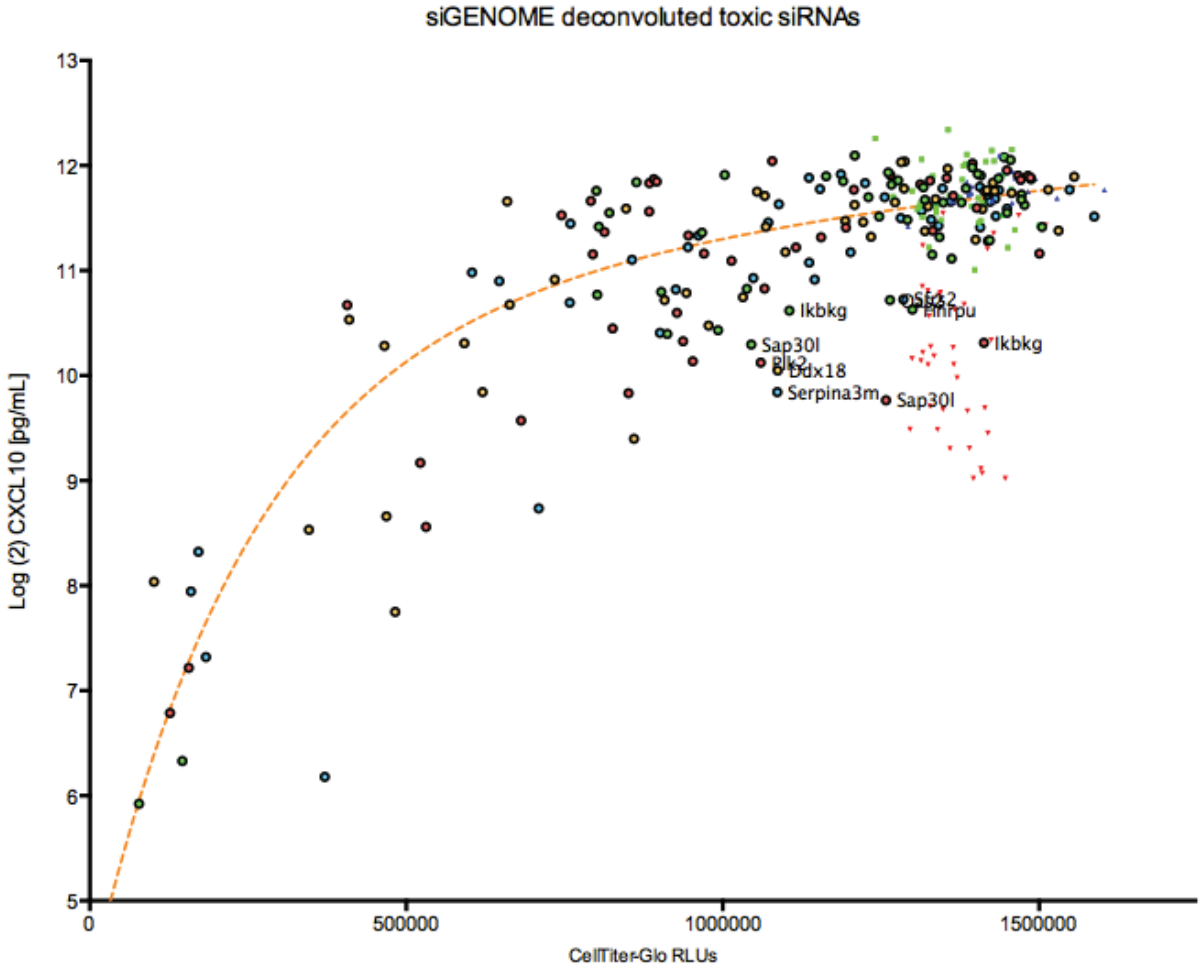


Figure 4.5 (continued)

were stimulated with ISD and Cxcl10 protein levels were subsequently detected by ELISA. Seven candidates, including the helicase *Ddx18* and DNA-SILAC candidate *Sfrs2*, and two with multiple siRNA targets (DNA-pathway regulated genes *Ikbkg* and *Sap30l*), had fourfold or greater reduction in Cxcl10 (Figure 4.5B). We included these recovered candidates in the following secondary screen designed to distinguish DNA-specific responses.

In an effort to identify ISD pathway-exclusive factors, we selected the top 200 potential positive and negative regulators and performed additional screens stimulated with both ISD and IVT-RNA, respectively (Figure 4.6A). We predicted that DNA-specific positive regulators of the ISD-sensing pathway provided the best chance at identifying a cytosolic DNA-sensor. Furthermore, focusing on DNA-specific response may simplify any subsequent validation of ISD-sensing pathway candidates. As expected, *Irf3* knockdown reduced expression of Cxcl10 in both ISD and IVT-RNA stimulated samples (greater than 90% reduction in Cxcl10), while knockdown of *Mavs* and *Rig-i* resulted in reduced Cxcl10 for RNA samples only and *Tbk1* knockdown reduced CXCL10 for ISD-stimulated samples (Figure 4.6B).

Twenty-five of the tested candidates produced 25% less Cxcl10 when stimulated with ISD than those stimulated with IVT-RNA including DNA-SILAC candidates *Abcf1*, *Hmgb1*, *Reep4* and *Skp1a* (greater than 40-fold reduction in Cxcl10 compared to non-targeting control siRNA treated wells), microarray candidates *Ifit1*, *Ifitm1*, *Ikka* and *Plagl* (18-fold or greater reduction in Cxcl10 levels), cytoplasmic DNA-binding candidate *Csda*, and helicases *Dhx15* and *Dhx16* (greater than 9-fold reduction of Cxcl10).

Conversely, a number of candidates produced 25% less Cxcl10 when stimulated with IVT-RNA than those stimulated with ISD including regulated genes *Hmx3*, *Stat1* and

Figure 4.6

Secondary screening of top 200 candidates identifies DNA-specific responses. A) Selection of top screening candidates and controls from high-throughput screen for secondary screening with Dharmacon siGENOME SmartPools and stimulated with B) 1ug/mL ISD, and 0.1 ug/mL in vitro transcribed RNA (IVT-RNA) for 26 hours. Data points are Cxcl10 protein pg/mL, as detected by ELISA. Cxcl10 production was compared between stimuli. Control siRNAs (red triangles) are labeled, black circles represent assayed candidates, strongest hits are labeled.

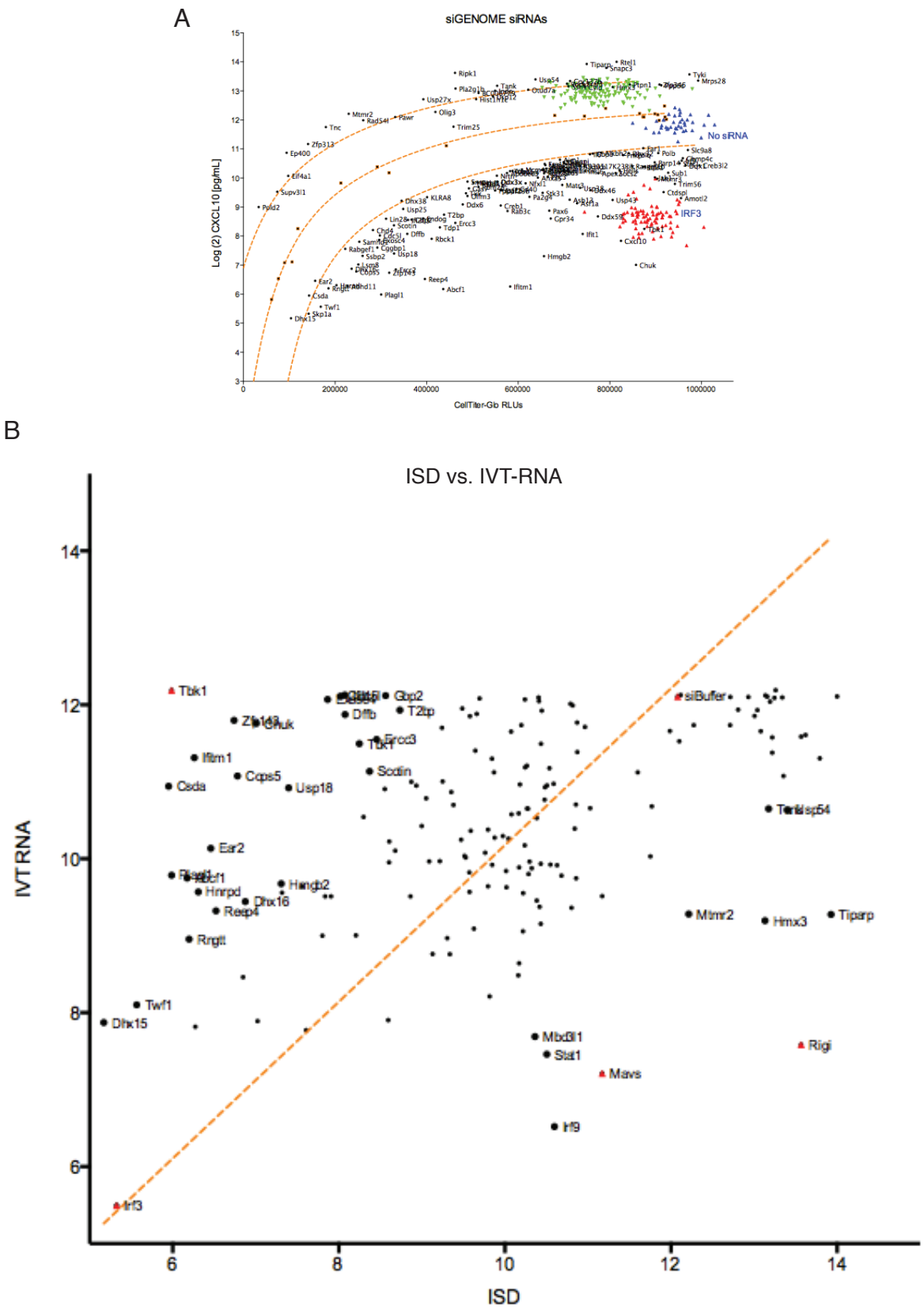


Figure 4.6 (continued)

putative negative regulators of the ISD-pathway *Mtmr2* and *Tiparp* (greater than 12-fold reduction in Cxcl10), and DNA-binding protein *Mbd3l1* (21.38 fold-reduction in Cxcl10).

Finally, we sought independent confirmation of our top hits using independent siRNA pools. To help reduce the potential of off-target effects of siRNA, we screened independent pools of siRNAs (Dharmacon On-Target PLUS SmartPools), corresponding to the top 40 candidates (Figure 4.7A). To further distinguish the ISD-pathway from RNA-responses, the top 40 candidates genes were knocked down and then stimulated with either ISD, Poly I:C or IVT-RNA for 26 hours (Figure 4.7B-C). Control siRNAs responded as expected; Cxcl10 responses were reduced for each nucleic acid stimulus in *Irf3*-siRNA treated samples (6.9, 6.8 and 6.0-fold reduction following stimulation with ISD, Poly I:C and IVT-RNA, respectively). Both Poly I:C and IVT-RNA induced Cxcl10 stimulation was reduced in *Mavs*-siRNA treated samples but only IVT-RNA responses were reduced in *Rig-i*-siRNA treated samples, consistent with the finding that RIG-I requires a 5'-triphosphate group on RNA to be detected[31]. Consistent with our previous findings, amongst others, signaling molecule *Ikka* and DNA-SILAC candidates *Abcf1*, *Ifit1* and *Reep4* appear to be DNA-specific positive regulators.

To further clarify the ISD-sensing capacity of the candidate genes, samples were stimulated with recombinant mouse IFN-β. Candidate genes may directly effect Cxcl10 production by an off-target effect or by targeting a component crucial in the secondary signaling cascade of events downstream of Irf3. In addition to cytoplasmic candidates including *Cnot6*, signaling molecule *Ikka* and DNA-SILAC candidates (*Abcf1*, *Ifit1* and *Reep4*) stimulation with recombinant mouse IFN-β demonstrated DNA-specific reduction in Cxcl10 production.

While chemical modifications effectively reduce off-target effects, a limitation of the On-Target PLUS siRNAs is reduced knockdown efficiency[176]. While we identified a

Figure 4.7

Secondary screening of top 40 candidates identifies DNA-specific responses. A) Selection of top screening candidates and controls from high-throughput screen for secondary screening with independent siRNA pools (Dharmacon On-Target PLUS SmartPools) and stimulated with B) ISD, Poly (I:C), in vitro transcribed RNA (IVT-RNA) or recombinant mouse IFNβ (mIFNβ) for 26 hours. Data points are Cxcl10 protein pg/mL, as detected by ELISA. CXCL10 production was compared between each stimulus. Control siRNAs (red triangles) are labeled, black circles represent assayed candidates, strongest hits are labeled. C) Absorbance values (450nm) of Cxcl10 ELISA in descending order for each stimulus, ISD, Poly (I:D), IVT-RNA and mIFNβ, respectively. Red bars represent putative negative regulators. Pink bars represent putative positive regulators. Screening controls (non-targeting, no-siRNA and IRF3) are indicated as light blue bars. Known ISD and RNA-sensing pathway components are represented as dark blue bars.

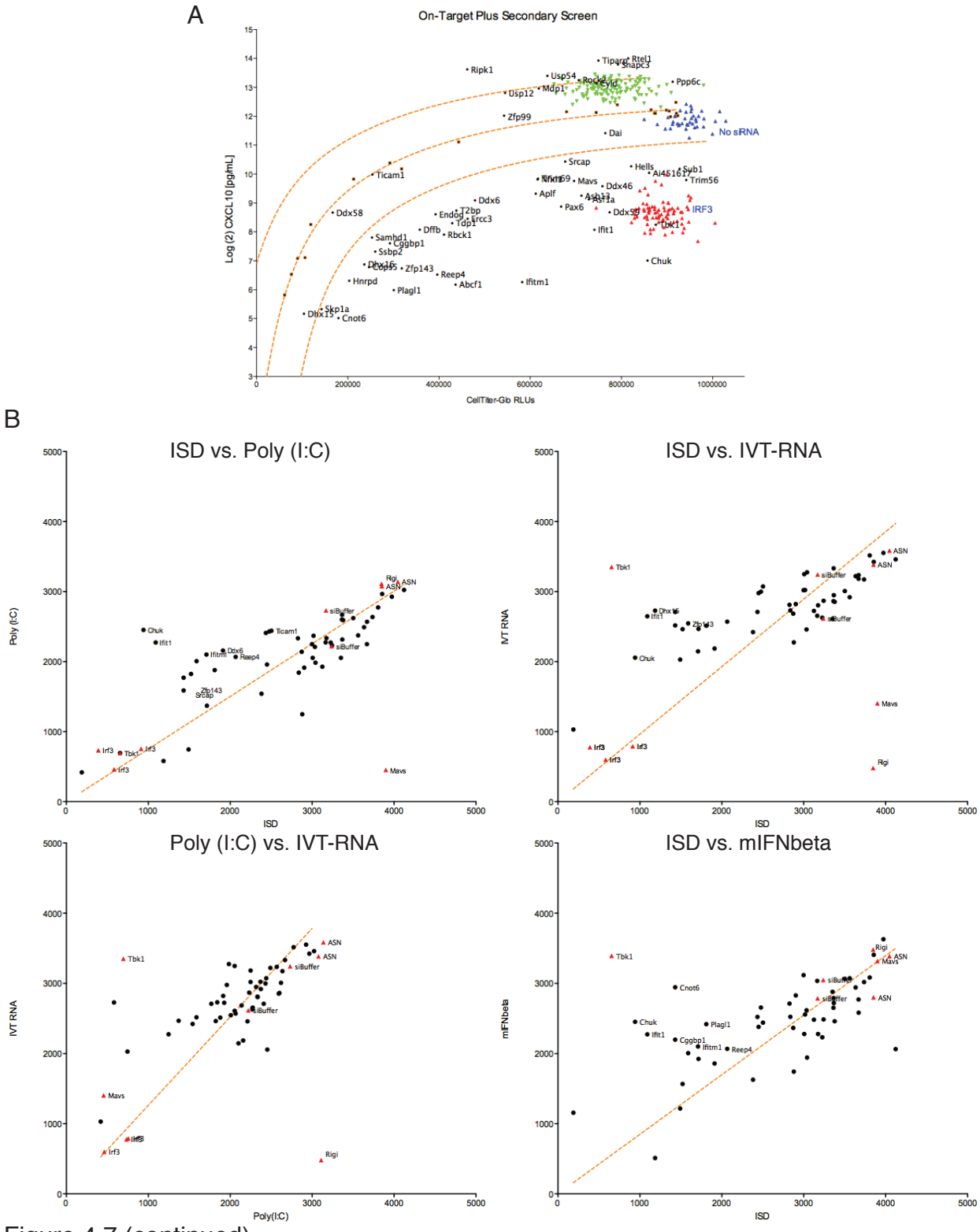


Figure 4.7 (continued)

C

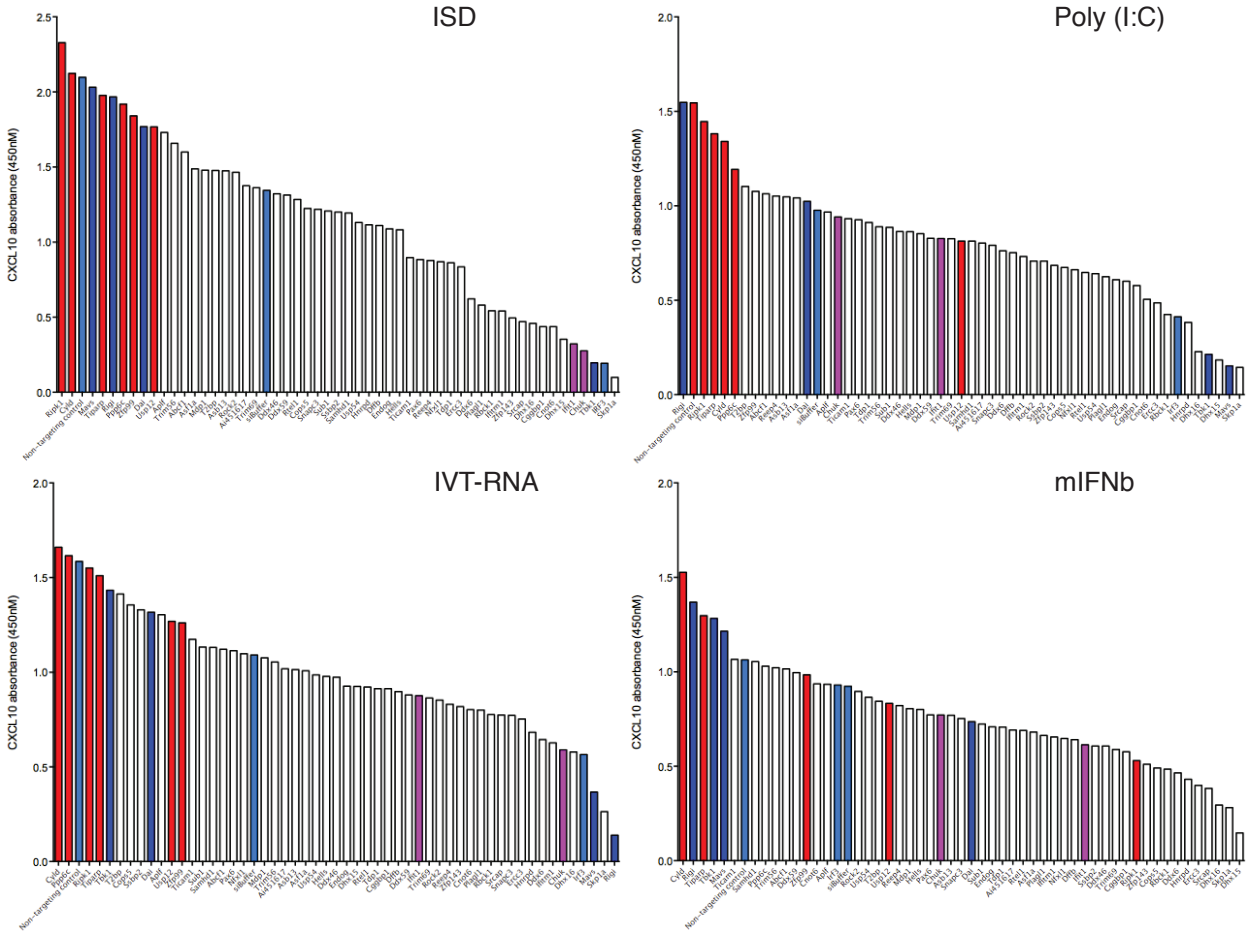


Figure 4.7 (continued)

number of candidates shared with siGENOME candidates, reduced knockdown efficiency may have obscured the role of potential candidate in the ISD pathway (Figure 4.7B-C and data not shown).

4.6 – *Conclusion*

Our high-throughput RNAi screen of 1003 putative ISD-pathway components revealed a number of novel factors representing candidates from our curated gene set of genomic, proteomic and domain-based candidates. By deconvoluting toxic siRNA pools, we recovered additional candidates. Top candidates were knocked down with independent siRNAs and stimulated with ISD, RNA or recombinant Ifn-β to identify putative ISD-specific components. We identified putative positive regulators including expression-induced genes (including *Ifitm1*, *Sp110*, *Trim56*, and *Tifa*), DNA-SILAC candidates (including *Abcf1*, *Ascc3*, *Hmgb2, Ifit1*, *Reep4*, and *Skp1a*), helicases (*Ddx18*, *Ddx46*, *Ddx59*, *Dhx15*, *Dhx16*, *Hells*, *Srcap*, and *Znfx1*), cytoplasmically located DNA-binding proteins (*Cnot6*, *Dffb*, *Endog* and *Zfp143*), phosphatases and deubiquitinases previously identified in our pilot screen including *Amotl2*, *Ctdspl*, *Mtmr3*, *Rasgrp1*, and deubiquitinases *Usp38*, *Usp43* and *Usp49*, and the signaling molecule *Ikka*. Furthermore, a number of putative negative regulators were identified including microarray candidate *Tiparp*, signaling molecule *Ripk1*, phosphatases (*Itpkb*, *Mdp1*, *Mtmr2*, *Ppp6c*, and *Ptpn1*), the deubiquitinase *Cyld* and ubiquitin specific proteases *Usp5*, *Usp12*, and *Usp27x*.

We identified putative ISD-pathway candidates through a high-throughput screen and secondary screens that will be further validated through siRNA-resistant cDNA rescue or targeted knockout. On their own, however, these candidates provide an enriched dataset of genes with likely roles in the ISD pathway.

Chapter 5:

Validation and Characterization of Novel Regulators of the DNA Sensing Pathway

5.1 – *Introduction*

Following the identification of several potential components of the ISD-sensing pathway, we selected several of the strongest candidates to investigate in more detail. We pursued candidates that were putative DNA sensors identified in our DNA-SILAC and RNAi screens, as well as novel components of the ISD-pathway signaling cascade, and screening hits with no known ISD-pathway interaction partners. We used three parallel modes of validation; first, to reduce the risk of off-target siRNA effects, we investigated our strongest hits through deconvolution of siRNA pools, testing with additional sh- and siRNAs and targeted cDNA rescue, secondly, we validated candidate genes through the use of chemical inhibitors. Furthermore, we investigated several candidates with targeted knockouts of putative ISD-sensing components.

5.2 – *Validation of putative DNA-sensors*

Consistent with the reported function of HMGB proteins as sentinels for nucleic acid responses, we identified Hmbg2 in our DNA-SILAC screen and in our siRNA screen as a regulator of the ISD-sensing pathway[143]. We demonstrated that MEFs deficient in Hmgb2 abrogate the IFN-response to dsDNA ligands but not to IVT-RNA, consistent with an established binding preference for dsDNA by HMGB2 (Figure 5.1). While HMGB family proteins act as promiscuous sensors of immunogenic DNA and RNA, DNA-SILAC hits that demonstrated strong, DNA-specific reduction in Cxcl10 following siRNA knockdown were selected for further investigation (Figure 5.2A). As previously described, our DNA-SILAC screen successfully identified known ISD-pathway components including the HMGB family proteins[143], AIM2 inflammasome components[49, 143] and the cytosolic RNA polymerase III complex as well as members of the SET complex[104, 106, 143, 144] and Aicardi-Goutières syndrome-associated proteins SAMHD1

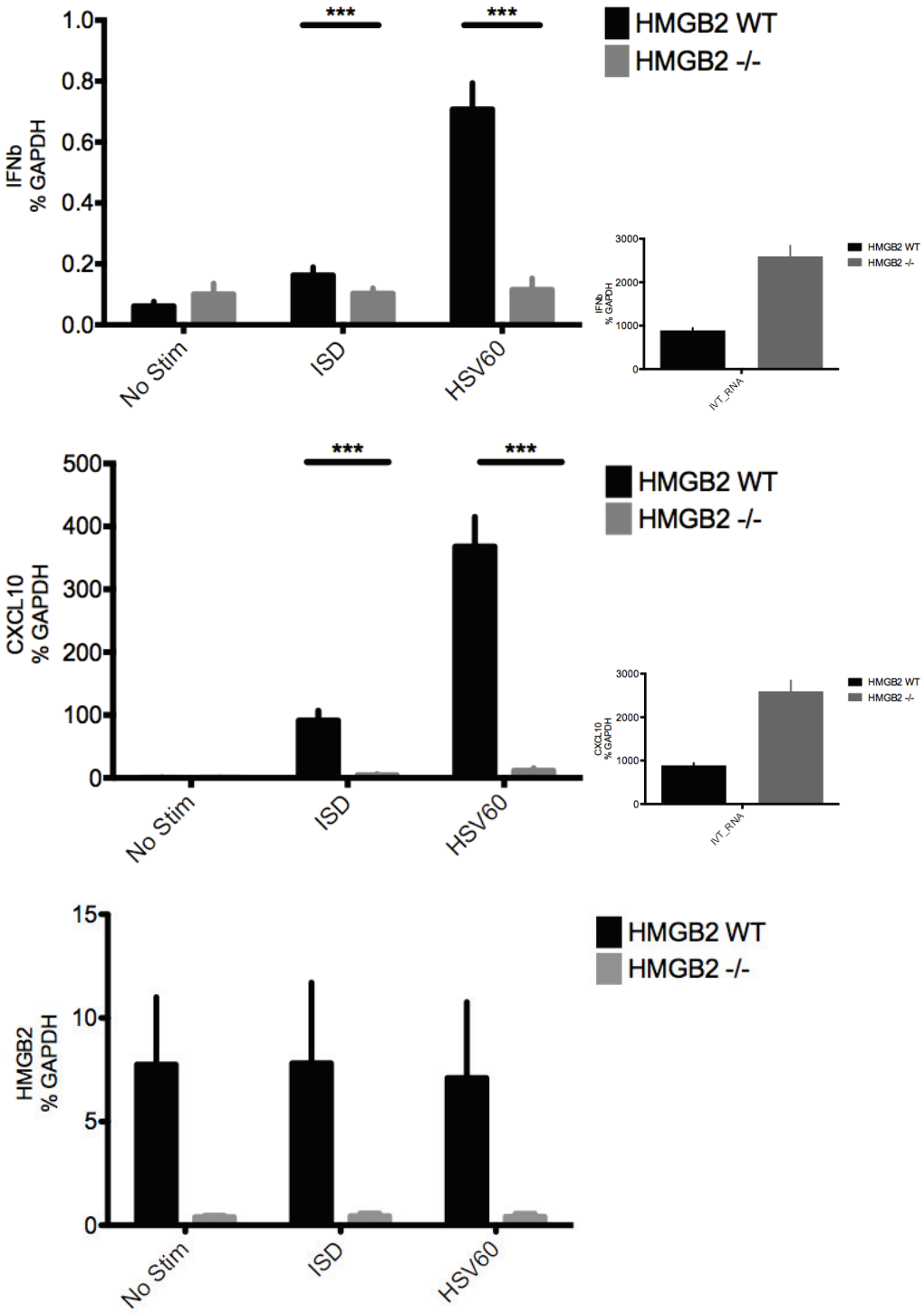


Figure 5.1

***Hmgb2* is a putative regulator of the ISD pathway.** *Hmgb2* wild-type and deficient MEFs were stimulated with the indicated dsDNA or IVT-RNA (inset) ligands for 6 hours. From top to bottom, *Ifnb*, *Cxcl10 and Hmgb2* mRNA expression is measured by quantitative RT-PCR. P-value <0.0001 (***), Student's t-test.

Figure 5.2

Putative DNA sensors *Abcf1*, *Ifit1* and *Reep4* have DNA-specific responses; independent of sequence or source of DNA ligand and partially replicate in an independent cell line. A) Selection of top candidates (red text) and controls (black text) from high-throughput siRNA screen. B) Selected candidates and control genes are knocked down and then stimulated with 1ug/mL ISD or 0.1ug.mL IVT-RNA for 26 hours. *Cxcl10* expression is measured by quantitative RT-PCR. C) Candidate genes are stimulated with the indicated sources of DNA following knockdown. Cxcl10 production is measured by ELISA and normalized to CellTiter-Glo. D) Mouse lung fibroblast were treated with the indicated siRNAs and then stimulated with ISD for 26 hours. *Cxcl10* production was measured by quantitative RT-PCR (left) and ELISA (right). P-value <0.01, 0.001, 0.0001 (*, **, ***, respectively), Student's t-test.

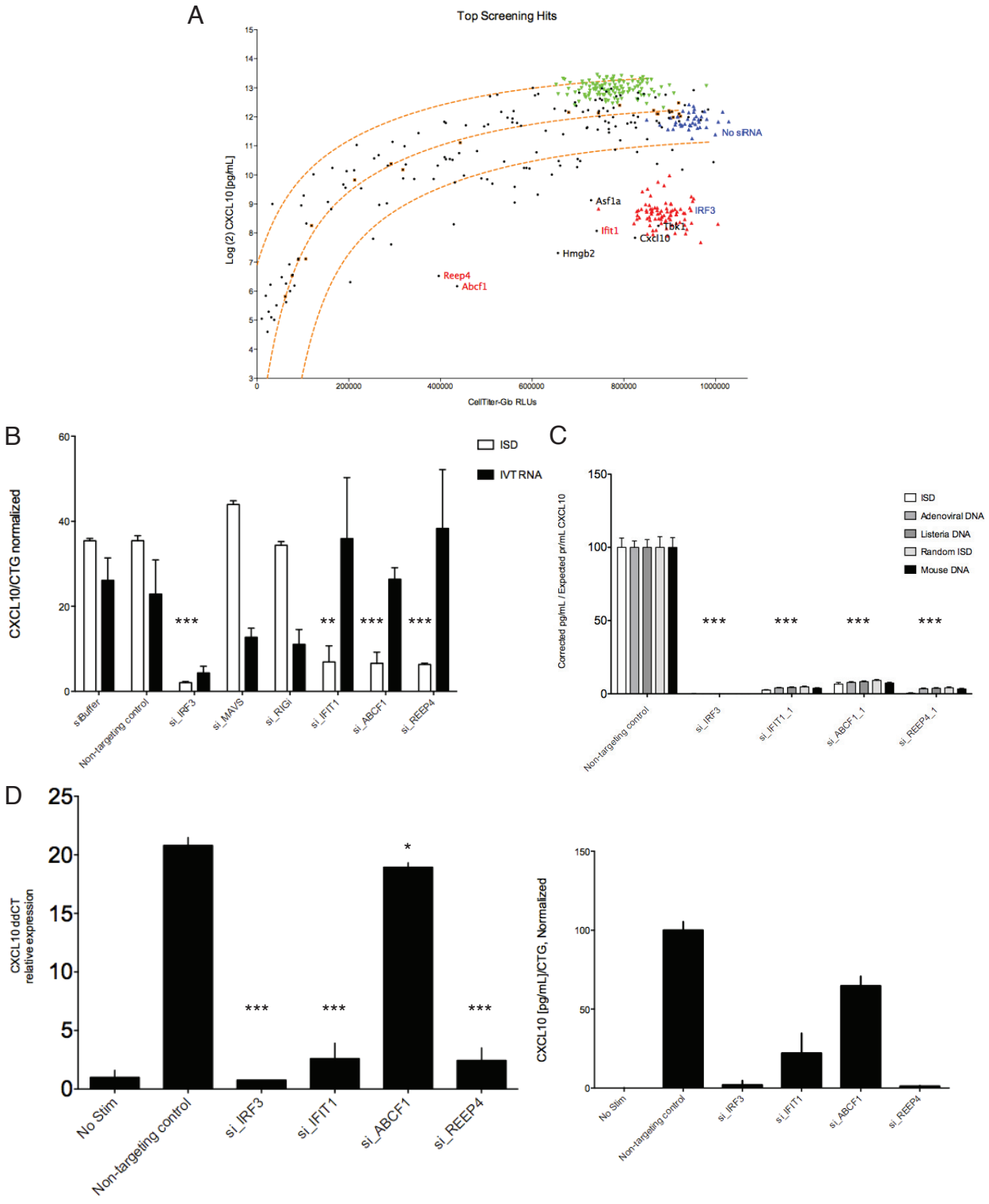


Figure 5.2 (continued)

and TREX1[100, 145, 170]. The combination of quantitative proteomic evidence, microarray-derived expression data and our siRNA-screen results place *Abcf1*, *Ifit1* and *Reep4* at the top of our candidate list.

We validated our screening results by repeating knockdown with independent siRNAs targeting each gene. Following knockdown, cells were stimulated with ISD or IVT-RNA. Cxcl10 induction and cell viability were measured by ELISA and CellTiter-Glo, respectively (Figure 5.2B). We demonstrated 5.1-, 5.4- and 5.6-fold reduction in Cxcl10 production for *Abcf1*, *Ifit1* and *Reep4*, respectively. Stimulation of siRNA-treated cells with multiple dsDNA ligands demonstrated that the phenotype was not limited to ISD (Figure 5.2C). DNA isolated from bacteria (*Listeria monocytogenes*), adenovirus, and mice, as well as a random ISD ODN, was transfected following knockdown. Viability-normalized Cxcl10 response was significantly reduced following knockdown with *Irf3* as well as for each of the candidate genes. Next, we demonstrated significant reduction in Cxcl10 levels following knockdown of each candidate gene in an independent cell line by investigating Cxcl10 responses in siRNA-treated mouse lung fibroblast cells (Figure 5.2D). Expression levels of candidate genes correlated with reduction in *Cxcl10* expression (data not shown). Knockdown of *Ifit1* and *Reep4* resulted in 8.0- and 8.5-fold reduction in *Cxcl10* response but only a 10% reduction following *Abcf1* knockdown, indicative of a cell-specific role of *Abcf1* in embryonic fibroblasts.

Following initial validation of these putative DNA sensors, we set out to further investigate the role of the *Abcf1*, *Ifit1* and *Reep4* in the DNA-sensing pathway.

5.3 – *Abcf1*

Abcf1 is member of the ATP Binding Cassette protein family, which, unlike most ABC proteins, lacks a transmembrane domain. Localized in the cytoplasm and ER,

Abcf1 interacts with eukaryotic initiation factor 2 (eIF2) to promote translation initiation[177]. Although a role in cytosolic DNA sensing has not been previously described for Abcf1, there is evidence that Abcf1 functions as a negative regulator of IL-6 and TNFα from Abcf1 heterozygote mice stimulated with CpG[178]. Additionally, recent evidence demonstrates that ABCF1 interacts with human polyomavirus 6 and 7[179].

We used 17 siRNAs targeting *Abcf1* in MEFs and measured *Cxcl10* expression by RT-PCR in response to stimulation of the ISD-sensing pathway. Knockdown of *Abcf1* correlated with reduced *Cxcl10* expression in ISD stimulated cells (R^2=0.615) (Figure 5.3A). A panel of *Abcf1*-mRNA-targeting siRNAs demonstrated a DNA-specific reduction when stimulated with either DNA or IVT-RNA, the strongest of which reduced Cxcl10 expression more than 94% compared to non-targeting control siRNAs as detected by ELISA (Figure 5.3B). We subsequently confirmed Abcf1 knockdown by immunoblot detection (Figure 5.3C). We also demonstrated that the phenotype could be significantly reversed in a doxycycline-dependent manner used to titrate the expression of an siRNA-resistant cDNA (*Abcf1* rescue gene) but not a *Renilla* luciferase cDNA control[140].

Abcf1 deficient mice are embryonic lethal at day 3.5 days post coitus, indicative of a crucial role in translation initiation[178]. *Abcf1* heterozygote (*Abcf1*$^{+/-}$) mice, however, appear to be developmentally normal, are fertile and show no significant differences in their gross anatomy when compared to their wild-type littermates. We investigated ISD responses in heterozygote *Abcf1* MEFs from independent littermates and observed conflicting results. Passage immortalized *Abcf1*$^{+/-}$ MEFs (passage number >8) were stimulated with dsDNA (ISD or HSV60, a 60 base-pair ODN derived from HSV-1). *Ifn-β* and *Cxcl10* expression are reduced over 95% following stimulation with either ISD or HSV60 dsDNA ligands (Figure 5.3D). In contrast, neither low passage MEFs from either *Abcf1* wild-type or heterozygous mice responded to ISD (Figure 5.3E). We also infected

Figure 5.3

***Abcf1* is a putative regulator of the ISD pathway.** A) *Cxcl10* mRNA expression in MEFs treated with 17 different siRNAs targeting *Abcf1* and stimulated with DNA, plotted against *Abcf1* mRNA expression for corresponding siRNA treated MEFs, measured by quantitative RT-PCR. B) ELISA of CXCL10 in MEFs treated with non-targeting and *Irf3* siRNAs and the indicated panel of *Abcf1*-targeting siRNAs were stimulated with ISD or IVT-RNA. Values were normalized to cell viability detected by CellTiter-Glo (right panel). Orange line represents p-value <0.05 cutoff for ISD stimulated samples compared to non-targeting control C) Immunoblot assay showing knockdown efficiency of non-targeting control siRNA and a representative siRNA targeting *Abcf1* mRNA; B-actin serves as a loading control. D) Passage immortalized *Abcf1* wild-type and heterozygous MEFs were stimulated with dsDNA ligands (ISD and HSV60) for 6 hours. Quantitative RT-PCR was used to detect *Abcf1*, *Ifnb* and *Cxcl10* expression from lysates. E) Low and high passage MEFs were stimulated with HSV60 or infected with HSVd109 (MOI=10) for 6 hours. Expression was determined by quantitative RT-PCR. F) MEFs were treated with the indicated siRNAs and stimulated with HSV60 or infected with HSVd109 for 6 hours. Expression for the indicated ISGs was determined by quantitative RT-PCR. P-value <0.01, 0.001, 0.0001 (*, **, ***, respectively), Student's t-test.

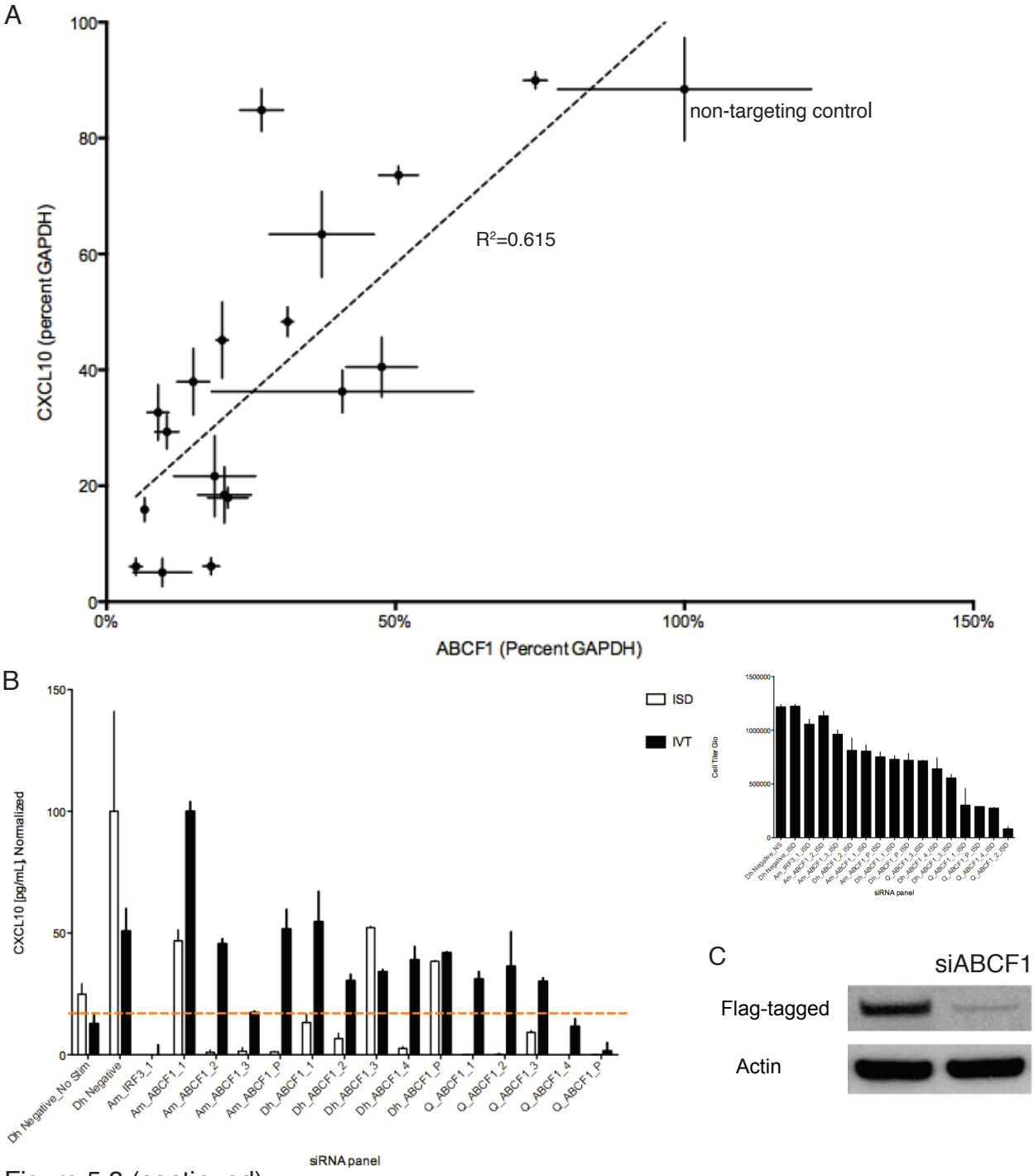


Figure 5.3 (continued)

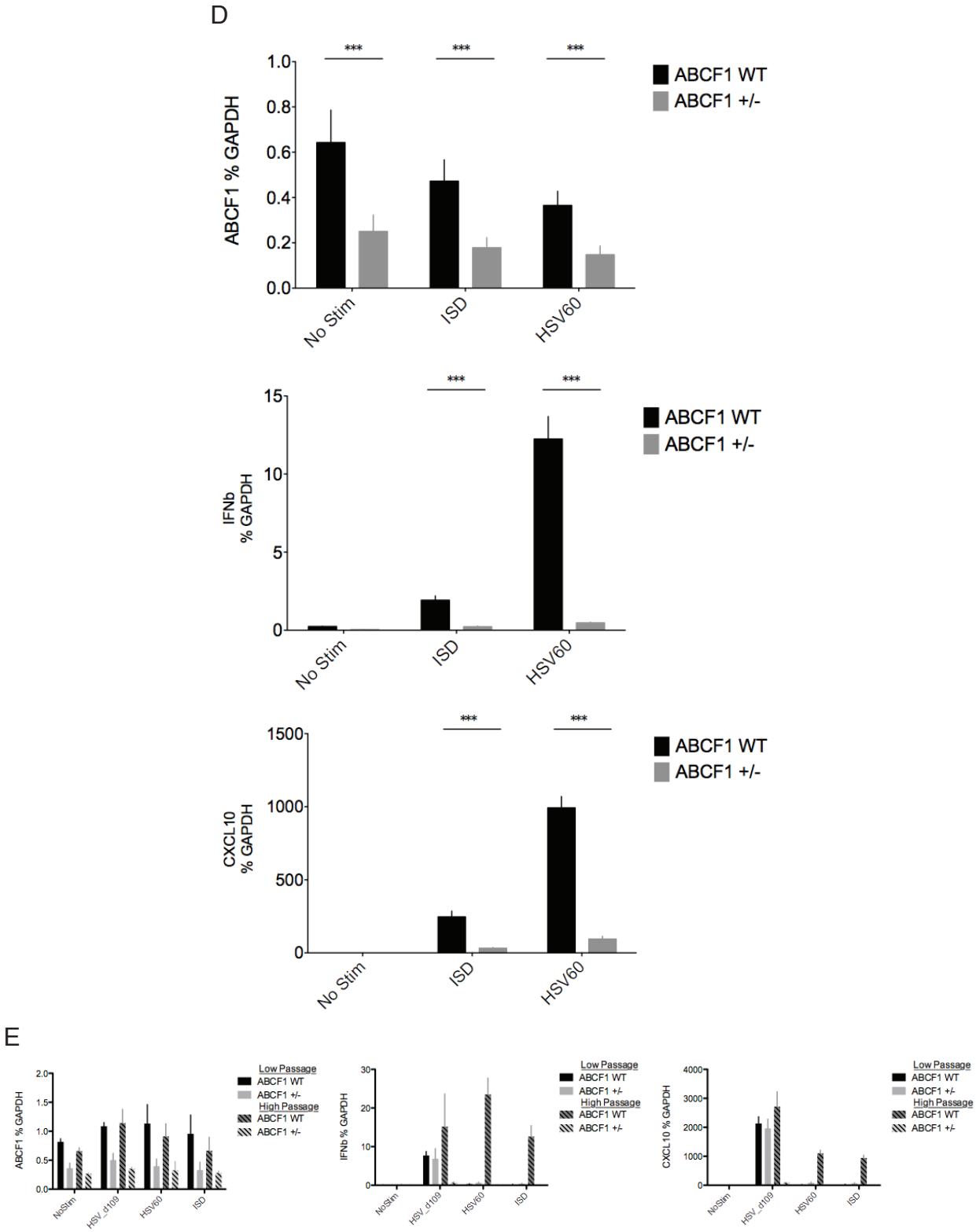


Figure 5.3 (continued)

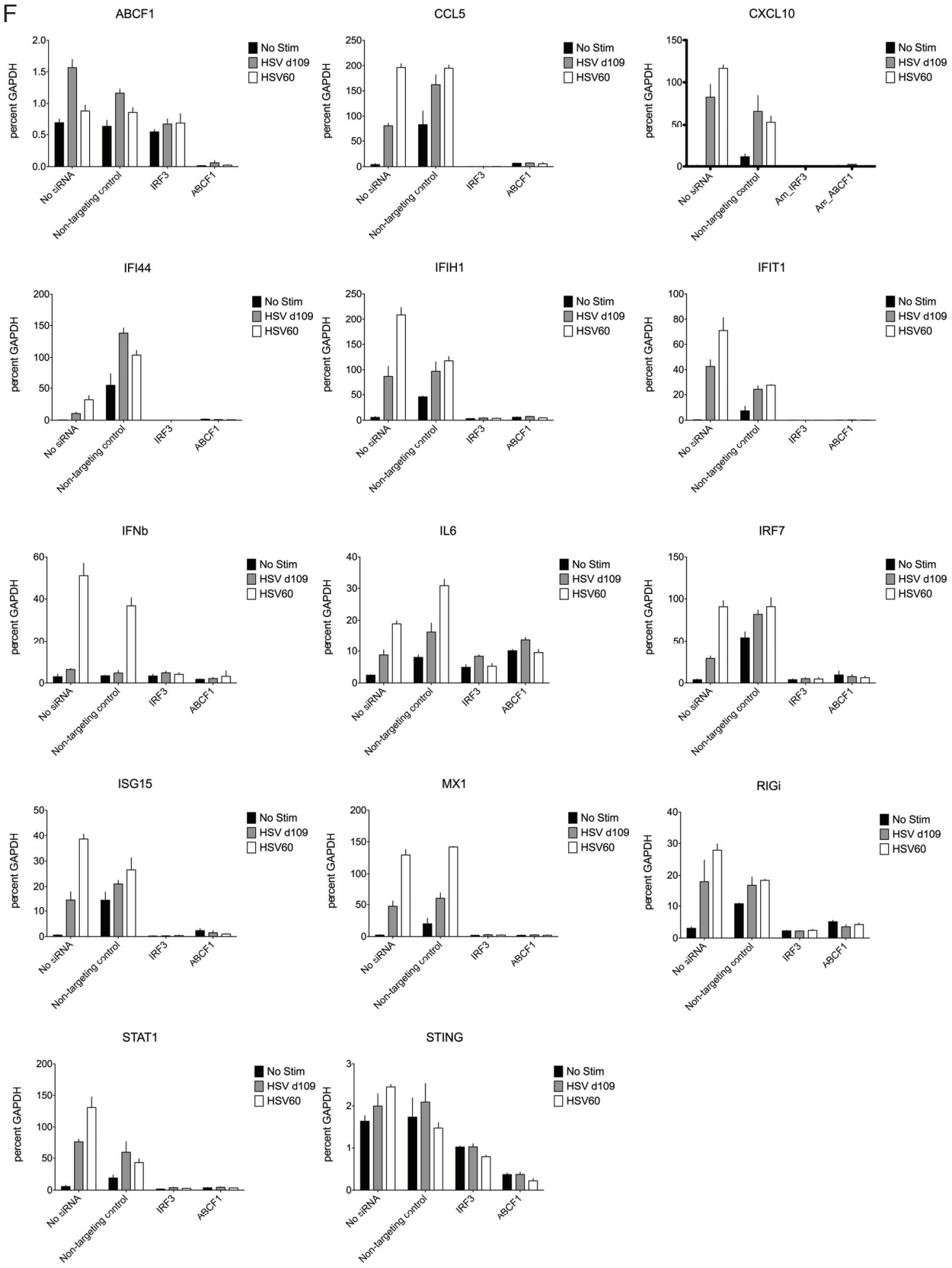


Figure 5.3 (continued)

low and high passage MEFs with replication-defective herpes simplex virus-1 (HSV-1) d109, a strain with mutations in the five immediate-early genes (ICP0, ICP4, ICP22, ICP27, and ICP47) effectively blocking all viral gene expression during infection, subsequently driving robust type I IFN expression[66, 180]. HSVd109 infection stimulated the expression of *Ifn-β* and *Cxcl10* in low passage *Abcf1* wild-type and heterozygous mice as well as high passage wild-type MEFs but not high passage *Abcf1*$^{+/-}$ MEFs. *Abcf1* mRNA expression was similar between heterozygotes (~50%) and wild-type MEFs, regardless of passage number. The phenotype described in the high-passage heterozygote *Abcf1* MEFs is indeed striking. How the two cell lines diverged is difficult to unravel. It is possible that passage immortalization drove these phenotypic differences. It appeared that with each passage wild-type cell ISD response increased, inversely correlated to a decrease in type I IFN responses in *Abcf1*$^{+/-}$. We have recently acquired embryonic stem (ES) cells from *Abcf1*-deficient mice. It is possible to convert ES cells to fibroblast-like cells by anti-MEF-antibody-mediated purification of embryoid bodies[181]. However, embryonic lethality of *Abcf1*-deficient mice suggests a key role in early development that may make the differentiation from MEFs from ES cells impossible.

In our recent publication, we further elucidated the role of *Abcf1* in the ISD-sensing pathway[140]. To further our understanding of Abcf1, we performed unbiased quantitative mass-spectrometry to identify 53 proteins that significantly (p-value <0.01) precipitated with hemagglutinin epitope-tagged (HA) Abcf1, three of which are members of the ER-associated SET complex (SET, Hmgb2 and Anp32a), and includes proteins identified in our DNA-SILAC experiments, Trex1 and Apex1. None of these proteins were present in a parallel pull-down experiment of HA-tagged Sting. Consistent with previous reports that Abcf1 localizes to both ER and cytosolic compartments, we found that a subset of Abcf1 localized with SET and the ER marker calreticulin by

immunofluorescence staining. In addition to the SET complex, we found that Abcf1 interacts with Hmgb2 and the putative DNA sensor Ifi204[65]. Additionally, we demonstrated that siRNA mediated knockdown of *Abcf1* suppressed Tbk1 and Irf3 phosphorylation following stimulation with ISD.

We also considered the role of Abcf1 in the response to viral infection. We demonstrated significant reduction of *Ifn-β* and ISG induction in $p53^{-/-}$ MEFs following siRNA-targeted knockdown of *Abcf1* and stimulation with the dsDNA ligand HSV60 or infection with HSVd109 (Figure 5.3F). Furthermore, expression of *Ccl5*, *Cxcl10*, *Ifi44*, *Ifih1*, *Ifit1*, *Ifn-β*, Irf7, *Isg1*, *Mx1*, *Rig-i* and *Stat1* was reduced more than 90% following stimulation with HSV60. There was no significant effect on Ifn-β and ISG induction by Sendai virus (which stimulates the Rig-i pathway) or by recombinant IFN-β[140]. These data indicate a potential role for Abcf1 in the response to cytosolic viral DNA but not RNA.

Lastly, we examined the role of Abcf1 in regulating host responses to retroviral infection[140]. Upon infection with an HIV-based retrovirus, *Trex1*-/- MEFs but not wild-type MEFs, produce *Ifn-β* and many ISGs[106]. It is thought that HIV-1 allocates Trex1 to degrade HIV retroelements and thus avoid detection by cytosolic DNA sensors. Knockdown of *Abcf1* in *Trex1*-/- MEFs and subsequent infection with an HIV-based retrovirus significantly reduced *Ifn-β* and *Cxcl10* expression implicating an Abcf1-associated mechanism for detecting retroelements.

Taken together, Abcf1 appears to be a critical factor in the DNA-sensing network. Abcf1 interacts with the SET complex, Hmgb2 and Ifi204, affects *Ifn-β* and ISG responses following infection with HSV-1 and may play a role in the innate immune response to retroviral infection as demonstrated by genetic perturbation and retroviral infection in *Trex1*-/- MEFs. The precise mechanism of Abcf1 remains to be elucidated.

Abcf1-SILAC data strongly suggest, however, that Abcf1 is part of a greater complex, the SET complex, which may be involved in early detection of DNA-based pathogens. The generation of an inducible Abcf1-knockout mouse may aid in the elucidation of its role in the ISD-sensing pathway.

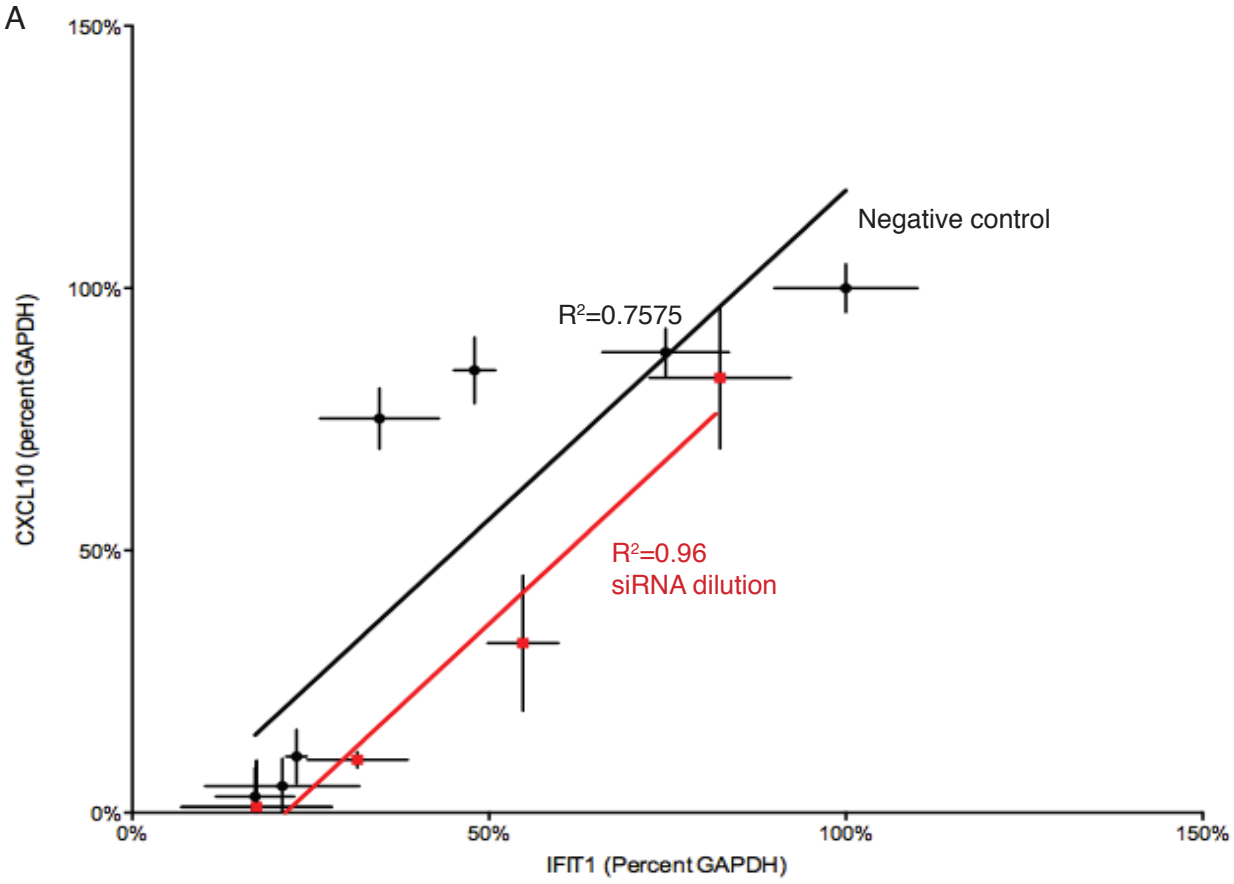
5.4 – *Ifit1*

We identified the interferon-inducible protein Ifit1 among our DNA-SILAC candidates as a strong DNA-binding partner. Ifit1 is a member of the IFIT family of cytoplasmic proteins consisting mainly of tetracopeptide repeats, a structural motif thought to mediate the assembly of multiprotein complexes, but contains no annotated nucleic-acid binding domain[182, 183]. Ifit1 is strongly induced following stimulation with nucleic acids and has recently been implicated as an antiviral protein that recognizes the 5' triphosphate group on viral RNAs[184]. Identified in a similar manner to our mass spectrometry approach, IFIT1, along with IFIT5, an IFIT protein found in humans but not mice, were found to directly associate with 5' triphosphate RNAs in human embryonic kidney cells (HEK293). The authors propose that the tetracopeptide repeats behave with an inherent binding plasticity that may facilitate promiscuous binding to RNA and potentially other nucleic acid ligands[185]. Prior to these recent discoveries, we pursued Ifit1 as a putative regulator of the ISD-pathway.

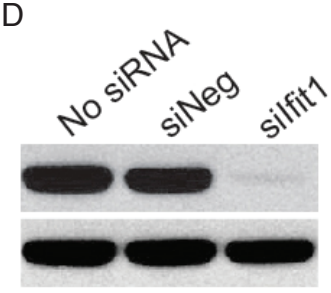
We tested seven siRNAs targeting *Ifit1* in MEFs and measured *Cxcl10* expression by RT-PCR in response to stimulation of the ISD-sensing pathway. Knockdown of *Ifit1* correlated with reduced *Cxcl10* expression in ISD stimulated cells (R^2=0.748) (Figure 5.4A). Deconvolution of the siGENOME SmartPool used in our siRNA screen revealed one of four *Ifit1*-targeting siRNAs reduced *Cxcl10* expression following stimulation with ISD. The reduction in *Cxcl10* strongly correlated with the amount of

Figure 5.4

Validation of *Ifit1* and its homologues as putative regulators of the ISD pathway A) *Cxcl10* mRNA expression in MEFs treated with 7 different siRNAs targeting *Ifit1* and stimulated with DNA, plotted against *Ifit1* mRNA expression for corresponding siRNA treated MEFs, measured by quantitative RT-PCR. Dilution curve (red line) of single *Ifit*-targeting siRNA from 5nM to 50nM siRNA. B) Deconvolution of *Ifit1*-targeting siGENOME SmartPools. Cxcl10 protein expression was detected by ELISA in MEFs treated with non-targeting and *Irf3* siRNAs and the indicated panel of *Ifit1*-targeting siRNAs were stimulated with ISD or IVT-RNA. Values were normalized to cell viability detected by CellTiter-Glo. C) MEFs were infected with lentiviral shRNAs targeting *Ifit1*. Three days following puromycin selection, cells were stimulated with ISD for 6 hours. *Cxcl10* and *Ifit1* expression was determined by quantitative RT-PCR. D) Immunoblot assay showing knockdown efficiency of non-targeting control siRNA and a representative siRNA targeting *Ifit1* mRNA; B-actin serves as a loading control. E) ProteinBLAST alignment of Ifit1 antibody epitope to homologous proteins. F) Schematic of the Ifit family, chromosome 19C1. G) Left panel, expression of Ifit family genes following stimulation with 1ug/mL ISD or 1000 units recombinant mouse IFNb. Right panel, expression of *Ifit1* family genes after siRNA transfection with the indicated siRNAs and stimulation with ISD. Ifit family expression detected by quantitative RT-PCR. H) Screen of *Ifit1* and homologous genes following siRNA treatment and stimulation with ISD. I) CXCL10 expression is normalized to cell viability of *Ifit1* and homologous genes. (Dotted orange line, p-value, <0.001, Student's t-test). MEFs knocked down with *Ifit1* super-family siRNAs are stimulated with 1.0 ug/mL of the indicated dsDNA ligands. Cxcl10 expression is detected by ELISA and normalized to cell viability. J) Wild-type and *Ifit1*$^{-/-}$ MEFs were stimulated with ISD for the indicated times. Expression of was determined by quantitative RT-PCR. K) Wild-type and *Ifit1*$^{-/-}$ MEFs were transfected with the indicated siRNAs and stimulated with ISD. Expression of was determined by quantitative RT-PCR.



D



E

IFIT1 protein alignment – SC Antibody 12-20AA, between 50-100 AA

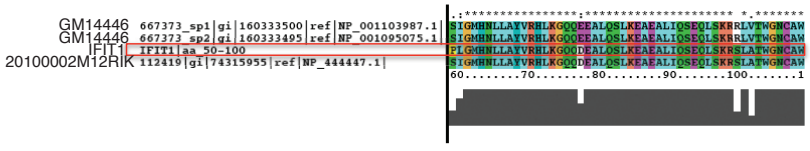
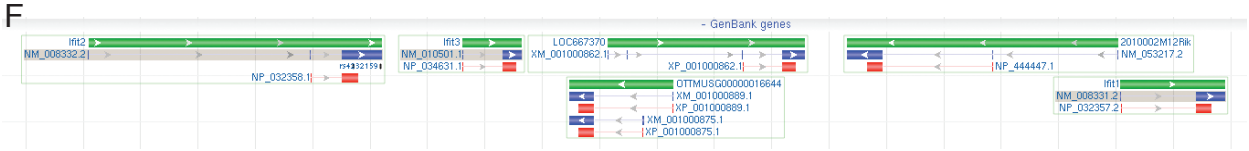
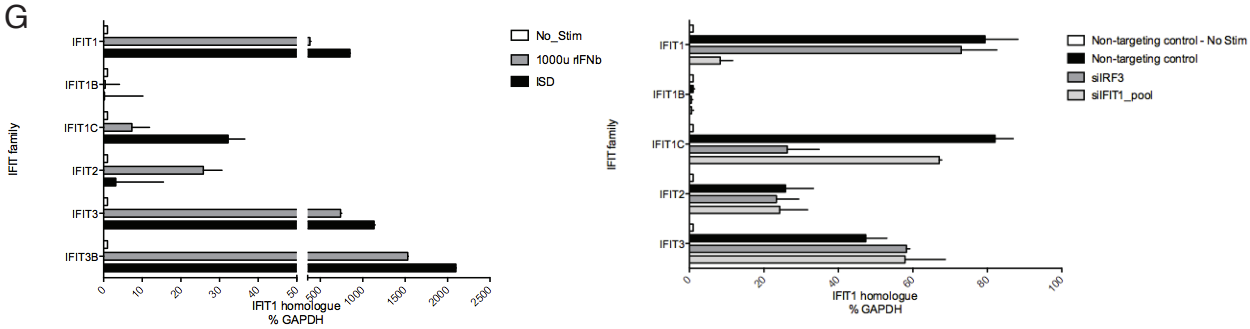


Figure 5.4 (continued)

F



G



H

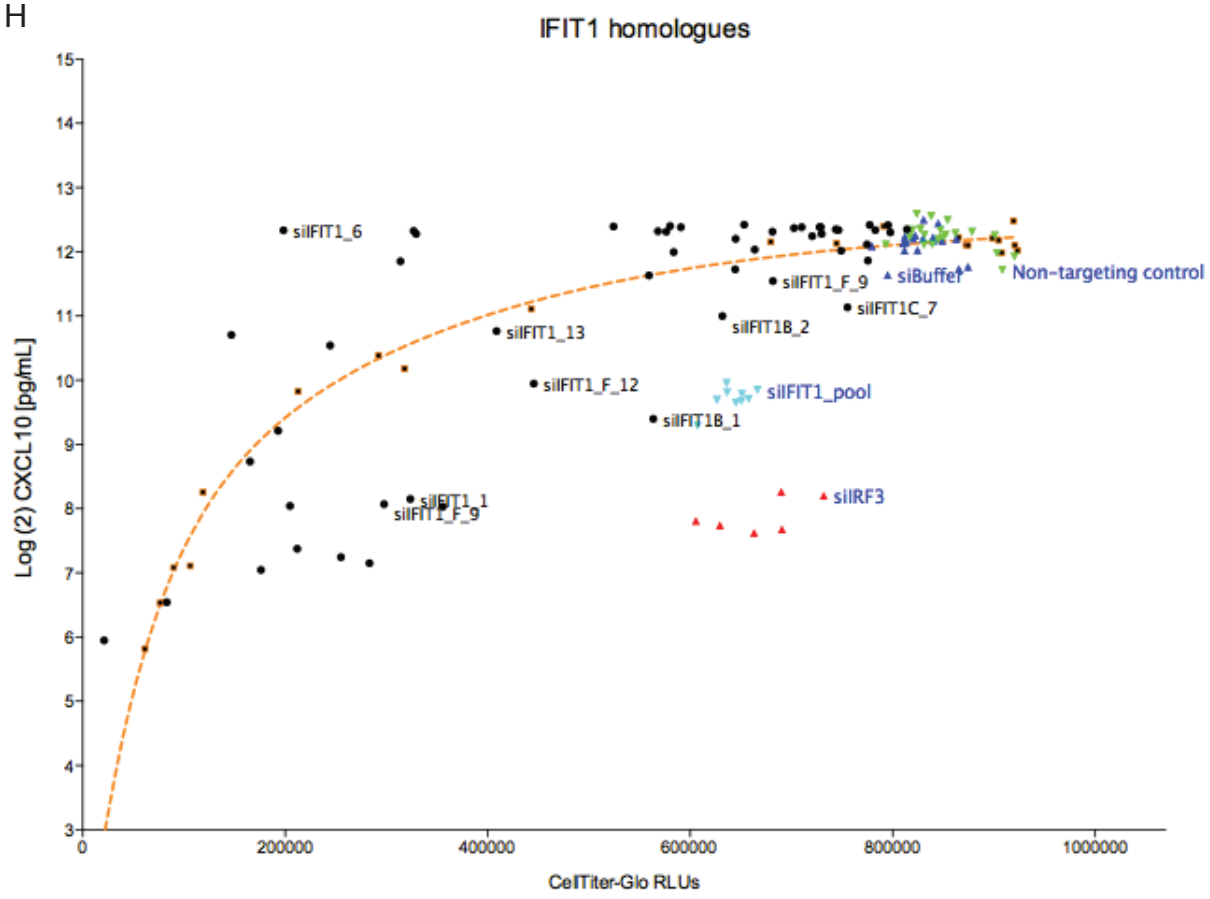


Figure 5.4 (continued)

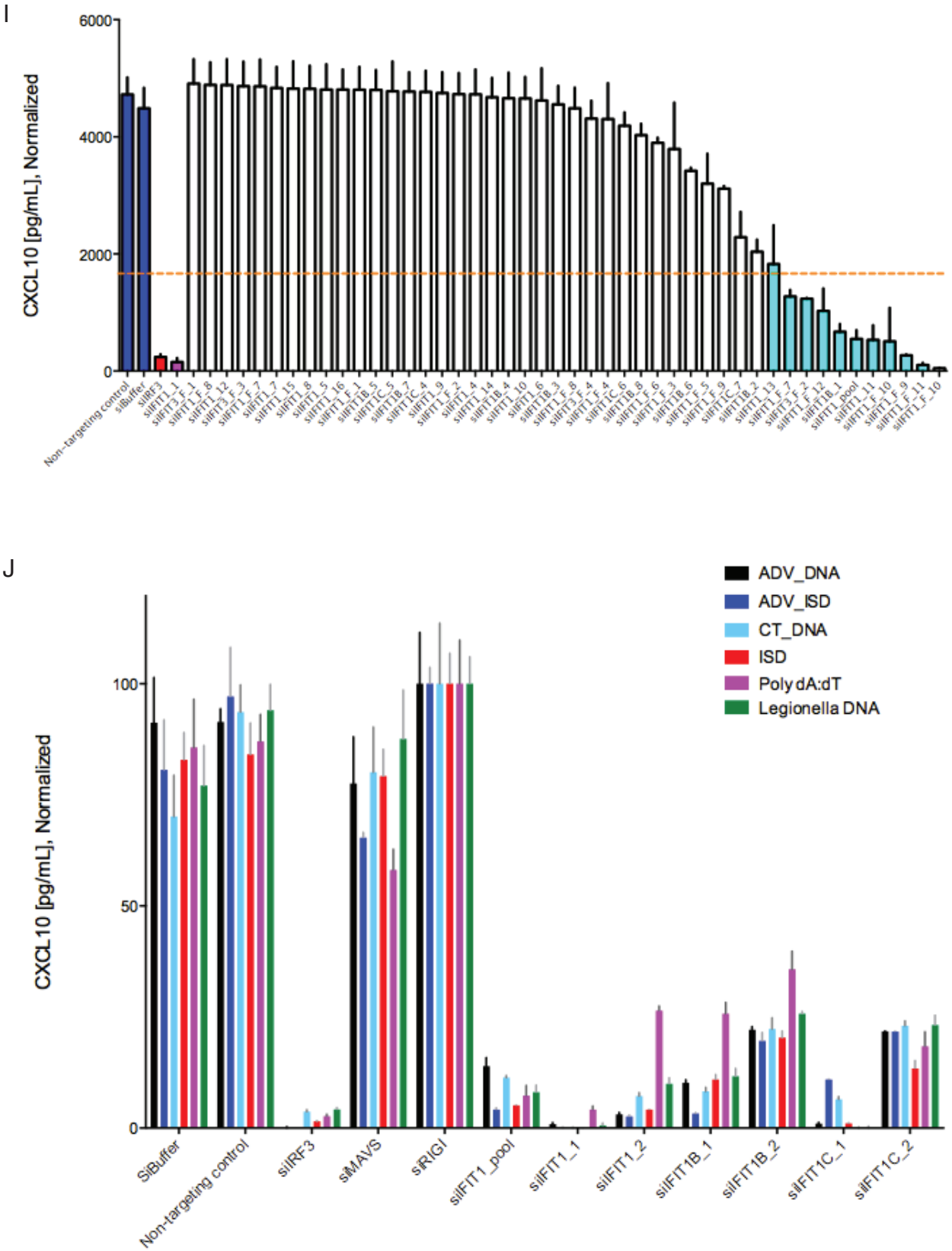


Figure 5.4 (continued)

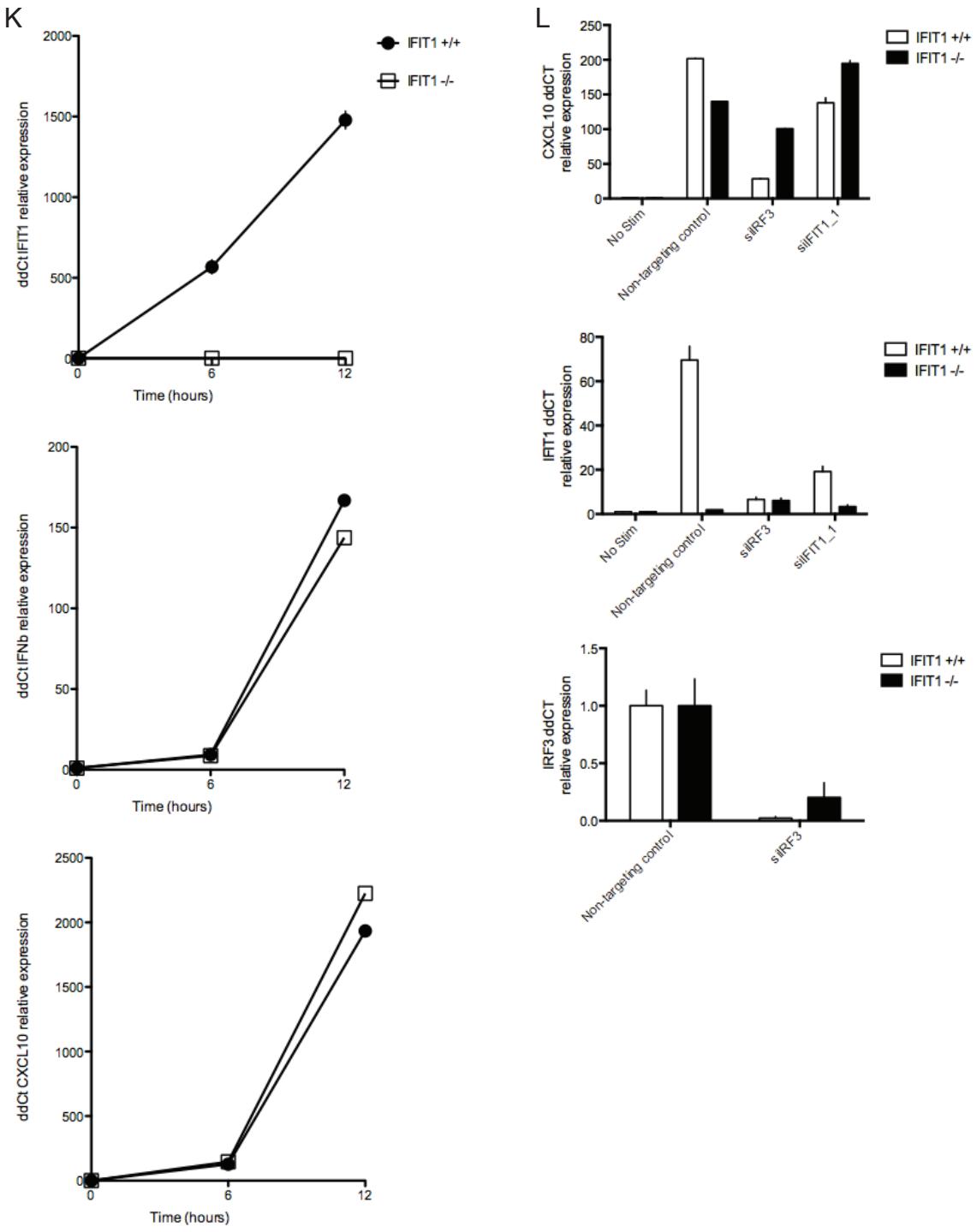


Figure 5.4 (continued)

transfected siRNA (R^2=0.955) (Figure 5.4A, B). The reduction in *Cxcl10* was DNA-specific, as response to transfected IVT-RNA remained unchanged. Additionally, $p53^{-/-}$ MEFs were infected with lentiviral shRNAs targeting *Ifit1* and, following puromycin selection, were stimulated with ISD. Cxcl10 protein levels, detected by ELISA, correlated with *Ifit1* expression (R^2=0.58) (Figure 5.4C). Specificity of *Ifit1* siRNA knockdown was further demonstrated by immunoblot of Ifit1 protein. Expression was reduced in MEFs treated with the strongest siRNA targeting *Ifit1* (Figure 5.3D). However, sequence alignment of the IFIT1 antibody epitope revealed homology with two proteins, GM14446 and 20100002M12RIK clustered within the IFIT family of proteins in a 110kb region on chromosome 19C1 (Figure 5.4E, F). We hypothesized that these proteins may contribute to the putative role of Ifit1 as a regulator of the ISD-pathway.

The murine IFIT family includes *Ifit2*, and *Ifit3*, as well as three interspersed IFIT-like genes, whose mRNAs may not have been fully sequenced. Little is known about their expression but sequence analysis reveals one or two ISREs in proximity to the transcriptional start site of each gene, tentatively named *Ifit3b* (official gene symbol, I830012O16Rik), *Ifit1b* (2010002M12Rik), and *Ifit1c* (Gm14446) [183]. Sequence analysis revealed that Ifit3b is 96% identical to *Ifit3*, and Ifit1b and Ifit1c are closely related to *Ifit1*; the encoded proteins share 78% identical amino acids with each other and 60% with murine Ifit1, both of which have reverse orientations with respect to *Ifit1* within the locus.

We explored the role of these putative Ifit1 homologues by measuring gene expression following activation of the ISD-sensing pathway. MEFs stimulated with ISD strongly induced the expression of *Ifit1*, homolog *Ifit1c*, *Ifit3* and its homolog *Ifit3b* but not *Ifit2* or the *Ifit1* homolog *Ifit1b* (Figure 5.4G). Recombinant Ifn-β, however, induced *Ifit1*, *Ifit2*, *Ifit3*, *Ifit3b*, and to a lesser extent, *Ifit1c*. Additionally, we assessed whether siRNAs

targeting *Ifit1* affected the expression of Ifit family. Knockdown of *Ifit1* by siRNA reduced expression of *Ifit1* but no other Ifit family members.

Following the discovery that ISD induced the expression of Ifit family members, we investigated their potential role in the ISD pathway by designing siRNAs that targeted each gene and their respective homology clusters (e.g. the Ifit super-family; *Ifit1*, *Ifit1b*, *Ifit1c*). In a screen of Ifit1 family members, Cxcl10 protein expression was measured following knockdown and ISD stimulation of 43 Ifit family-targeting siRNAs (Figure 5.4H). Five siRNAs targeting *Ifit1* and its homologous genes reduced Cxcl10 expression four-fold or more. An siRNA targeting the *Ifit1* homologue *Ifit1b* showed more than 7-fold reduction in Cxcl10 (compared to 28 and 30-fold reduction of Cxcl10 following *Irf3* and *Ifit1* knockdown, respectively) (Figure 5.4I). We confirmed knockdown of *Ifit1* super-family members by quantitative RT-PCR (data not shown) and selected the strongest siRNAs for a screen of various dsDNA sources. In addition to *Ifit1*, siRNAs targeting *Ifit1b* and *Ifit1c* reduced Cxcl10 protein expression by 60% or more following stimulation with viral, bacterial, mammalian and synthetic DNA (Figure 5.3J). Furthermore, knockdown Ifit2, Ifit3 and Ifit3b appeared to have no effect on type I IFN production in ISD-stimulated MEFs (data not shown).

The discovery that siRNAs targeting multiple IFIT1 super-family members reduced ISD-directed IFN responses provided a compelling argument for a pan-IFIT1 role in the ISD-pathway, but it complicated the potential role of off-target siRNA effects in obscuring the role of each gene. Incomplete knockdown of one gene, or partial knockdown of multiple genes could further obscure the function of *Ifit1* in the ISD-pathway, as compensatory mechanisms of homologous family members will be difficult to distinguish. To rule out off-target or partial knockdown effects of siRNA, we stimulated MEFs from wild-type and *Ifit1* deficient mice and measured IFN production (Figure 5.4K).

There was no difference between wild-type and *Ifit1*$^{-/-}$ MEFs following stimulation with ISD. Next we treated wild-type and knockout MEFs with siRNAs targeting *Ifit1* and then stimulated with ISD (Figure 5.4L). Knockdown with siRNAs targeting *Ifit1* had no effect on *Cxcl10* expression of *Ifit1*$^{-/-}$ MEFs and only a mild reduction in wild-type MEFs. Knockdown appeared to be efficient for *Irf3* as demonstrated by quantitative RT-PCR suggesting that our transfection of siRNAs was effective. However, reduction of *Ifit1* in wild-type MEFs was uncharacteristically low (~30% reduction). The unusual siRNA data make the experiment difficult to interpret. However, we can speculate that in the absence of *Ifit1*, homologous Ifit proteins may play a compensatory role, masking the effect of *Ifit1* deficiency in the ISD-sensing pathway. Early attempts at overexpression of Ifit1 did not affect ISD-sensing pathway responses in MEFs treated with non-targeting control, *Irf3*, *Tbk1* or *Ifit1* siRNAs (data not shown). Furthermore, we have not measured the expression of IFIT family members in *Ifit1* deficient mice following stimulation of the ISD-sensing pathway. If our compensation hypothesis is correct, we may see increased IFIT family expression in *Ifit1* knockout MEFs compared to wild-type.

While the role for Ifit1 remains unclear with regards to the ISD-pathway, more details as to its potential role in innate immunity have begun to take shape. It has previously shown that IFIT1 binds to the eukaryotic initiation factor eIF3 and may limit the translation of viral mRNA by blocking the interaction of eIF3 with the ternary complex eIF2[186-188]. Subsequent knockout of *Ifit1* revealed susceptibility to a West Nile Virus mutant defective in its mRNA 2'-0 methylation, but had no increased sensitivity to intranasal infection of vesicular stomatitis virus (VSV) or encephalomyocarditis virus (EMCV) infection[189]. More recently, it has been demonstrated that Ifit1 (and human IFIT5) promote antiviral immunity by sensing 5'-triphosporylated RNA, similar to Rig-i[184]. Knockdown of *IFIT1* in HeLa cells lead to increased replication of VSV or influenza A but

not EMCV, which does not generate 5' triphosphate RNAs during replication. Furthermore, in support of our findings, *IFIT1* deficiency did not affect the phosphorylation of IRF3 following transfection with IVT-RNA, Poly I:C, ISD and Poly (dA:dT). Wild-type and *IFIT1* deficient MEFs, bone marrow-derived macrophages and bone marrow-derived dendritic cells produced similar levels of IFN and interleukin-6 following transfection of DNA and RNA nucleic acid ligands. Conversely, *Ifit1*-deficient mice succumbed to VSV infection more readily than wild-type mice while EMCV infected MEFs had equal viral loads, regardless of their genotype. The emerging evidence does not provide a clear role for Ifit1 or its homologues in the ISD-sensing pathway. It is possible that Ifit1 and its homologous genes form a complex mediated by the protein-protein complex-forming activity commonly associated with the repetitive tetracopeptide domains. These promiscuous domains, that behave with a protein binding plasticity not dissimilar from the leucine-rich repeats regions of TLRs, could bind DNA and promote antiviral sensing much in the same way it recognizes 5'-triphosphorylated RNA. This concept is further supported by a recent report demonstrating reduced ISRE activity in *Ifit1*-deficient mice following LPS or CpG treatment[184]. Furthermore, a recent report dissected transcriptional data of TLR-mediated response and identified IFIT1 as a critical bottleneck in regulating the expression downstream immune genes[190].

Though we have provided evidence supporting a role for Ifit1 in the ISD-sensing pathway, we could not rule out the possibility of off-target effects. Were we to pursue Ifit1 further, we would generate targeted knockouts of the entire Ifit1 super-family locus. The precise genomic editing function of transcription activator-like effectors (TALE)-based zinc-fingers may be a viable means to ask this question on a gene-by-gene basis[191, 192]. More immediately, we could investigate binding partners of Ifit1 in an Ifit1-SILAC screen.

5.5 – *Reep4*

Receptor expression-enhancing protein 4 (Reep4) is a member of REEP family of genes involved in intracellular trafficking and secretion. REEP proteins appear to partner with G-protein couple receptors (GPCR) to promote cell-surface expression of mammalian odorant and taste receptors[193]. Little is known about the function of Reep4. In a *Xenopus tropicalis* model, REEP4 deficiency causes paralysis in embryos as a result of defects in both muscle and neural development[194]. Mutations in human REEP1 have been linked to hereditary spastic paraplegia (OMIM 610250) and distal hereditary motor neuronopathy type VB (OMIM 614751) though the molecular basis of these phenotypes is unclear [195, 196]. Although no known role in the immunity has been ascribed for Reep4, it was strong candidate in our DNA-SILAC screen and knockdown of *Reep4* resulted in a 24-fold reduction in Cxcl10 expression.

To help rule out off target effects, we used 14 different siRNAs targeting *Reep4* in MEFs and measured *Cxcl10* expression by RT-PCR in response to stimulation of the ISD-sensing pathway. Knockdown of *Reep4* correlated with reduced *Cxcl10* expression in ISD stimulated cells (R^2=0.524) (Figure 5.5A). Seven of 14 siRNAs targeting *Reep4* reduced Cxcl10 protein expression 10-fold or more, while response to transfected IVT-RNA remained largely unchanged when corrected for cell viability (Figure 5.5B). To illustrate siRNA specificity, immunoblot of Reep4 demonstrated reduced expression in MEFs treated with the strongest siRNA targeting *Reep4* (Figure 5.5C). We generated an siRNA resistant cDNA (*Reep4* rescue gene) but failed to demonstrate a reversal of the phenotype following administration of doxycycline (data not shown).

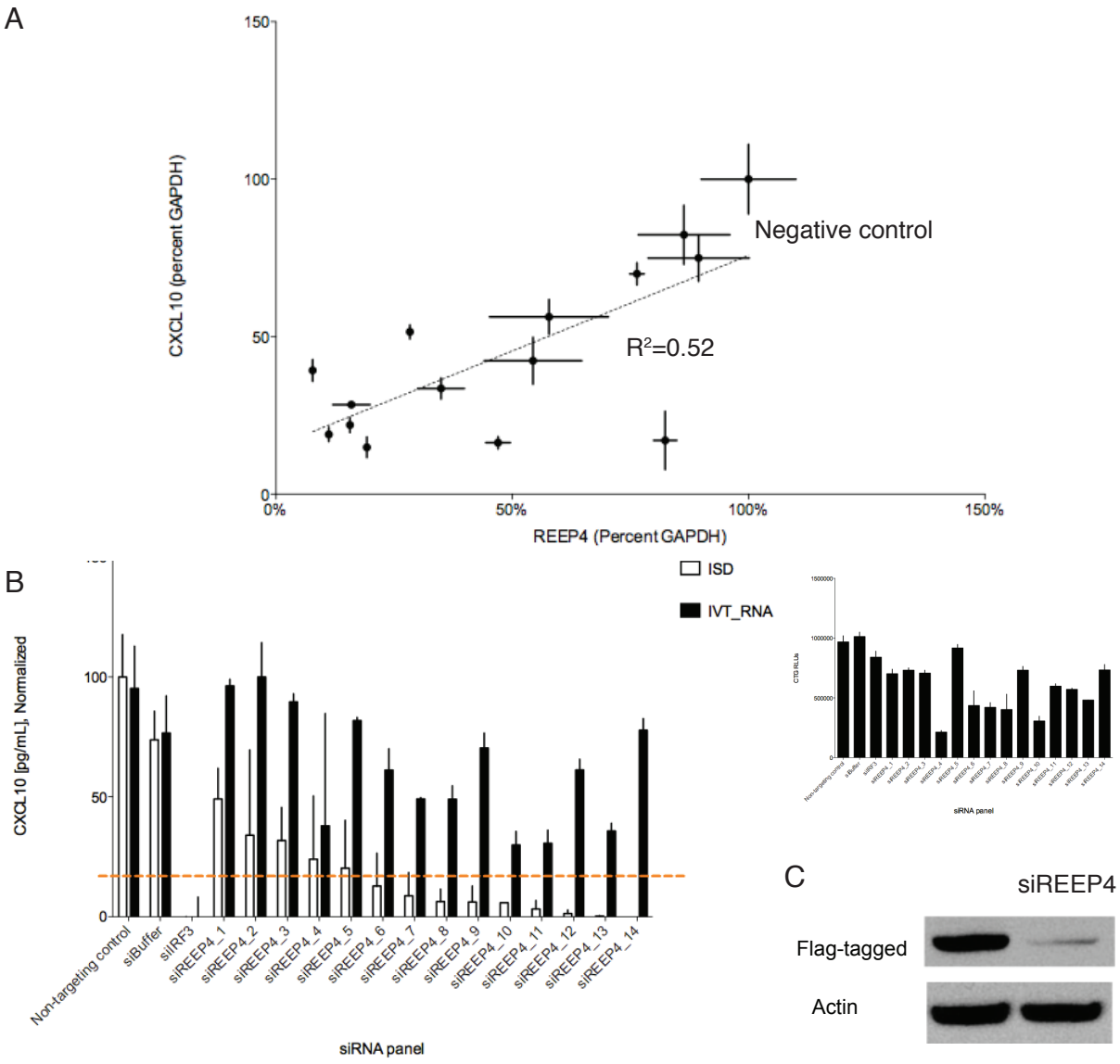


Figure 5.5

***Reep4* is a putative regulator of the ISD pathway.** A) *Cxcl10* mRNA expression in MEFs treated with 14 different siRNAs targeting *Reep4* and stimulated with DNA, plotted against *Reep4* mRNA expression for corresponding siRNA treated MEFs, measured by quantitative RT-PCR. B) ELISA of Cxcl10 in MEFs treated with non-targeting and *Irf3* siRNAs and the indicated panel of *Reep4*-targeting siRNAs were stimulated with ISD or IVT-RNA. Values were normalized to cell viability detected by CellTiter-Glo (right panel). Orange line represents p-value <0.05 cutoff for ISD stimulated samples compared to non-targeting control C) Immunoblot assay showing knockdown efficiency of non-targeting control siRNA and a representative siRNA targeting *Reep4* mRNA; B-actin serves as a loading control.

We further examined whether Reep4 also regulated the innate immune response to retroviral infection[140]. Knockdown of *Reep4* in *Trex1*-/- MEFs, followed by infection with an HIV-based retrovirus, resulted in significantly reduced *Cxcl10* expression. The human antimicrobial peptide LL-37, implicated in the pathogenesis of IFN-driven autoimmunity in psoriasis and lupus erythematosus, has recently been shown to facilitate the transport of self-DNA into monocytes via lipid rafts[197, 198]. LL-37 mediated transfer of dsDNA ligands induced the production of type I IFNs in a STING and TBK1-dependent manner. Additionally, recent evidence suggests that REEP family member REEP2 recruits sensory receptors into lipid-raft microdomains, improving GPCR receptor signaling and receptor access[199]. It is possible that Reep4 acts similarly to REEP2 but as an LL37-DNA complex recruiting mechanism for lipid-raft mediated endocytosis. This poses in interesting hypothesis, potentially placing Reep4 in the pathway of viruses that require lipid-raft formation to trigger endocytosis[200]. Cellular entry of the non-enveloped DNA Polyomaviruses requires caveolar/lipid-raft formation for endocytosis. Perhaps Reep4 plays a similar sentinel role to HMGB family proteins by surveying lipid-raft mediated viral entry. As a recent study predicted that Abcf1 interacts with polyomavirus, it is possible that Reep4 is similarly involved[179]. Polyomavirus infection of GM-CSF-differentiated dendritic cells induces the dramatic expansion of CD8+ T-lymphocytes[201]. To elucidate a potential role of Reep4 in polyomavirus-induced antiviral response, cDCs infected with lentiviral shRNAs targeting Reep4 could subsequently be infected with polyomavirus. Analysis of antiviral signatures in Reep4 knockout cDCs may provide a link to Reep4 regulation of the ISD-pathway.

Taken together, these data suggest that *Reep4* is a putative regulator of the ISD pathway. Although we were unable to rescue the phenotype with siRNA-resistant cDNA clones, and could therefore not rule out off-target effects, we demonstrated nucleic acid

specificity and described a potential regulatory role in the response to retroviral infection. Future investigations of Reep4 may focus on demonstrating a potential role in the regulation of lipid-raft mediated viral entry.

5.6 – *Putative ISD-sensing pathway signaling molecules*

In our recent publication, we demonstrated a role for novel regulators that interact with primary signaling molecules in the ISD-sensing pathway including Cdc37, a molecular chaperone that interacts with Hsp90, a putative interacting partner of Tbk1[140, 202]. We demonstrated that siRNA-mediated knockdown of *Cdc37* decreased Tbk1 protein expression and subsequently abrogated phosphorylation at key sites on Irf3. Furthermore, we demonstrated that small molecules targeting Cdc37, Hsp90 or Tbk1 decreased Ifn-β and Cxcl10 production in ISD-stimulated mouse lung fibroblast and human monocyte derived dendritic cells (MoDCs).

Similarly, signaling events downstream of the interferon receptor are also critical in the ISD-sending pathway. In addition to the identification of known mediators *Irf9* and *Stat1*, we investigated the role of putative ISD-sensing signaling molecules including the serine-threonine phosphatase Ppp6c and protein tyrosine phosphatase Ptpn1[140]. Although IkB-e is a proposed substrate for Ppp6c, and while tyrosine-phosphorylated proteins JAK2 and TYK2 are established substrates for Ptpn1, no targets in the ISD pathway have been established[202, 203]. In agreement with our siRNA screen, small molecule inhibition of Ppp6c by okadaic acid increased ISD-stimulation CXCL10 production in MoDCs. We also demonstrated that inhibition of PTPN1 in MoDCs increased CXCL10 expression following ISD-stimulation. Furthermore, consistent with our screen and small molecule-directed inhibition, *Ptpn1* deficient MEFs produced up to

2.4-fold more Cxcl10 than rescued MEFs in response to stimulation with increasing doses of ISD (Figure 5.6).

5.7 – *Putative ISD-sensing pathway candidates with no known ISD-interaction partners*

We also investigated the role of a number of candidate genes for which there are no known molecular interaction partners in the ISD-pathway including the interferon-regulated nuclear body protein Sp110 and the helicase Hells.

First detected in the *Mycobacterium tuberculosis*-susceptible C3HeB/FeJ mice, positional cloning of the tuberculosis susceptibility locus, *sst1* (for super-susceptibility to tuberculosis 1) identified the candidate gene Sp110[204]. Upregulated following infection with tuberculosis or *Listeria monocytogenes*, Sp110 limits bacterial multiplication in macrophages and mediates a switch in the cell death pathway from necrosis. Additionally, mutations in *SP110* have been associated with hepatitis C virus susceptibility and may interact physically with hepatitis C virus core protein as well as the Epstein-Barr virus SM protein[205, 206].

We demonstrated a putative role for Sp110 in the ISD-sensing pathway by stimulating conventional dendritic cells (cDCs) from C3H-*sst1*^{s} mice (*sst1* susceptible strain carrying a natural deleterious mutation in Sp110) with dsDNA. C3H-*sst1*^{s} cDCs produced 40% less Ifnb1 compared to wild-type cDCs confirming our siRNA screen finding (Figure 5.7). While interacting partners to Sp110 remain unclear, a putative pro-apoptotic binding partner, Mybbp1a, has been identified[207]. In addition to activation of the ISD-sensing pathway and the Aim2-dependent inflammasome, cytosolic DNA induces DNA-damage signaling proteins that trigger mitochondrial apoptosis[208]. Perhaps, Sp110, via a pro-apoptotic intermediary Mybbp1a, mediates cross-talk generated by intracellular DNA or pathogens that regulates activation, gene expression

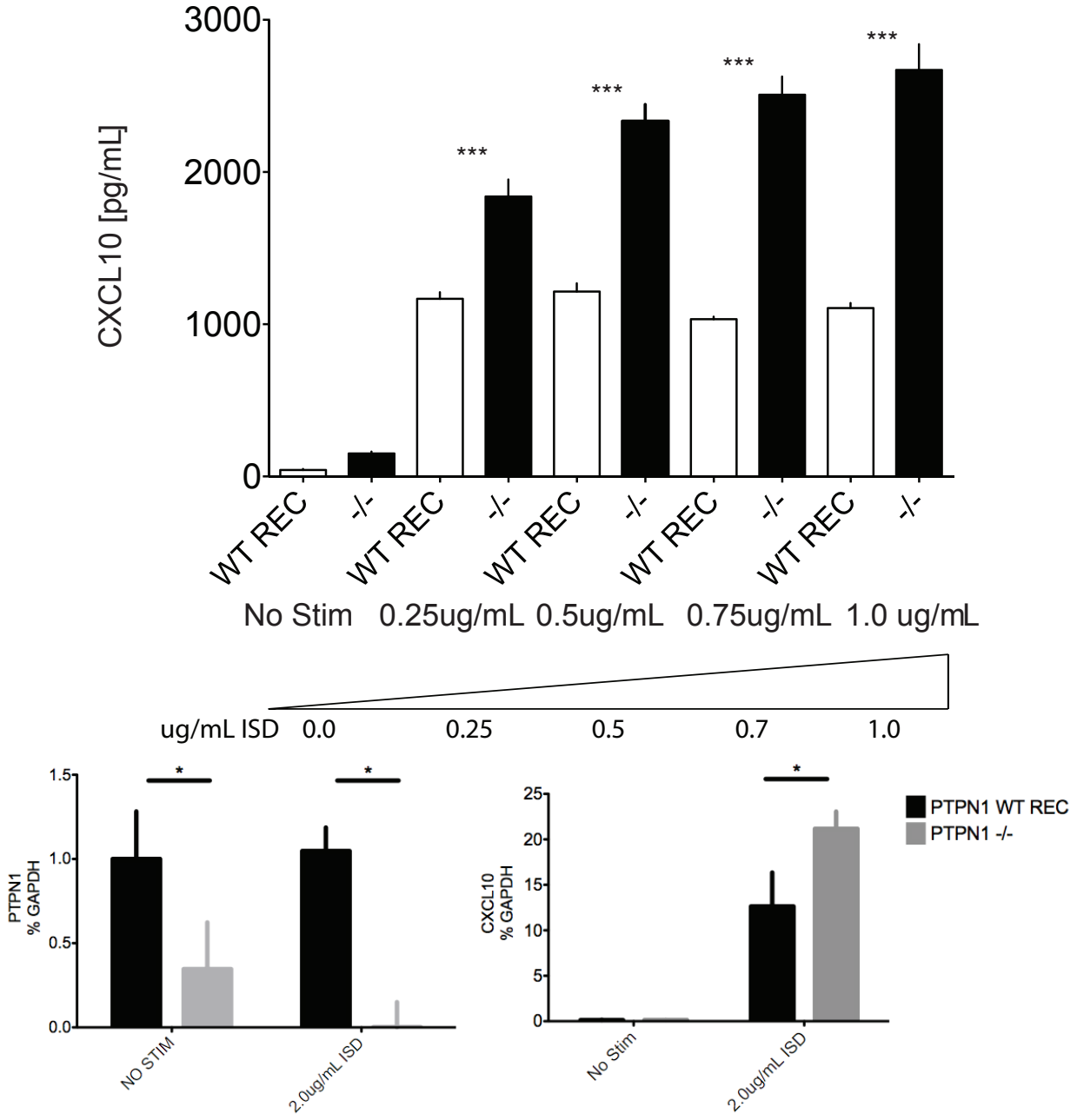


Figure 5.6

Ptpn1 is a putative negative regulator of the ISD pathway. Top, Cxcl10 expression was detected by ELISA following stimulation with indicated ug/mL of ISD for 26 hours in MEFs from *Ptpn1*$^{-/-}$ (-/-) and *Ptpn1*$^{-/-}$ reconstituted with wild-type (WT REC) mice. Bottom, mRNA expression of *Ptpn1* and *Cxcl10* following simulation with ISD in *Ptpn1* WT REC and -/- MEFs, measured by quantitative RT-PCR. P-value <0.01, 0.0001 (*, ***, respectively), Student's t-test.

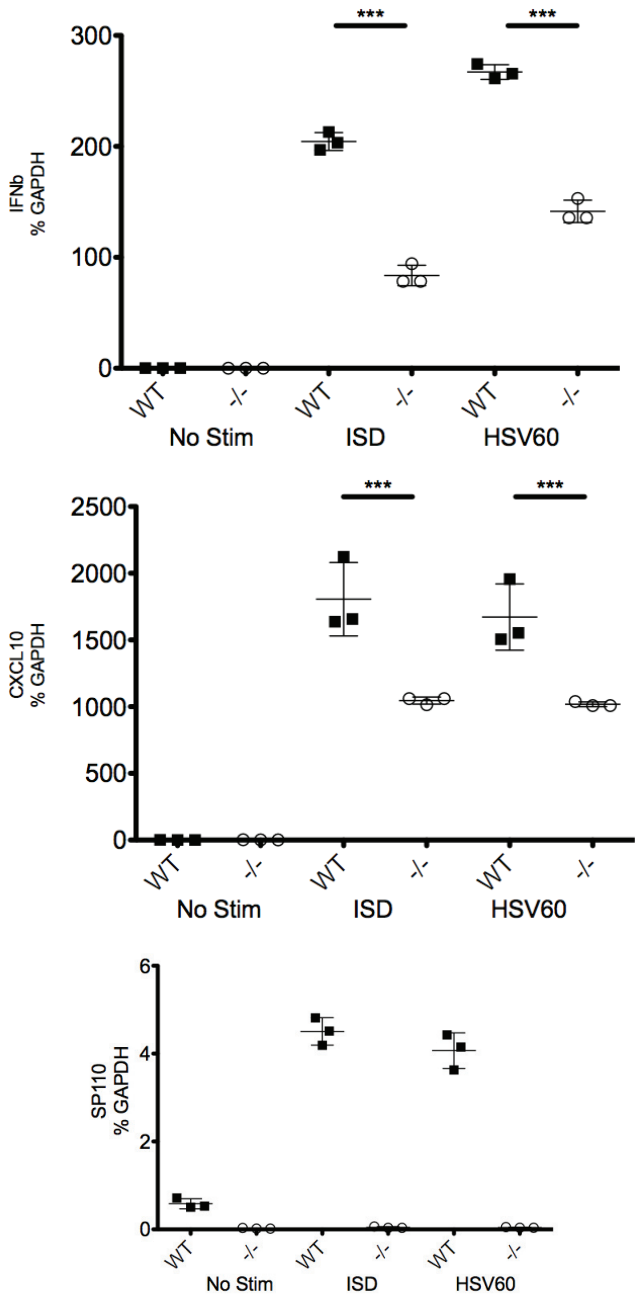


Figure 5.7

Sp110 is a putative regulator of the ISD pathway. Mouse conventional dendritic cells (cDCs) were prepared from wild-type (WT) or B6.C3H-sst1 (Sp110 LoF)(-/-) and stimulated with the indicated dsDNA ligands for 6 hours. From top to bottom, *Ifnb*, *Cxcl10* and *Sp110* mRNA expression is measured by quantitative RT-PCR. P-value <0.0001 (***), Student's t-test.

and apoptosis of host cells through unknown ISD-sensing pathway components[204]. Sp110 interacting partner Mybbp1a co-localizes with the nuclease NME1, a critical component apoptosis, regulated by SET complex proteins[209, 210]. Sp110 may therefore play a role in SET-complex-mediated ISD-sensing responses through yet-to-be-discovered interactions with pro-apoptotic molecules regulated by SET-complex proteins.

We also identified the lymphoid specific helicase Hells as a putative regulator of the ISD-sensing pathway. Hells is a chromatin remodeling ATPase, similar to the SWI/SNF family of chromatin remodelers, that modulates genome-wide cytosine methylation patterns at non-repeat sequences[211]. In addition to DNA hypomethylation, Hells-deficient mice show delayed growth, multiorgan and skeletal defects, premature graying, kyphosis, cachexia and early death[212]. An independent animal model deficient in Hells shows neonatal death, low birth weight, lymphocyte defects (T-cells are reduced by 60% and B-cells are reduced by 40%) and renal lesions[213].

Though there is no known interaction with ISD-sensing sensing components, we investigated the role of Hells as a putative regulator in the ISD-sensing pathway. The IFN response to dsDNA stimulated *Hells* deficient MEFs was significantly reduced by more than 90% compared to matching wild-type MEFs (Figure 5.8A). To demonstrate specificity of the ISD-sensing effect, we stimulated *Hells*$^{-/-}$ MEFs with Adenovirus, Sendai Virus, dsDNA and dsRNA ligands. The reduced type I IFN response appeared to be specific to DNA as stimulation with Poly I:C and Sendai Virus did not significantly reduce Ifn-β production (Figure 5.8B). Because Adenovirus failed to induce an Ifn-β response, we infected *Hells*$^{-/-}$ MEFs with HSVd109, inducing a robust *Cxcl10* response in wild-type MEFs but demonstrated a 95% reduction in *Hells*-deficient MEFs (Figure 5.8C). Because HSVd109 is replication deficient, we infected *Hells*$^{-/-}$ MEFs with replication competent

Figure 5.8

Hells is a putative regulator of the ISD pathway. A) Wild-type (HELLS WT, gray bars) and *Hells*$^{-/-}$ MEFs (HELLS -/-, black bars) were stimulated with the indicated dsDNA ligands for 6 hours. Left to right, *Ifnb*, *Cxcl10* and bottom, *Hells* mRNA expression was measured by quantitative RT-PCR. B) Wild-type and *Hells*$^{-/-}$ MEFs were infected with adenovirus (AdV), Sendai virus (SeV) or stimulated with dsDNA or dsRNA ligands for 26 hours. Ifnβ (pg/mL) was detected by ELISA. C) MEFs were infected with HSV variant d109 or transfected with increasing doses of dsDNA for 6 hours. *Cxcl10* expression was measured by quantitative RT-PCR. D) MEFs were infected with replication-competent adenovirus stably expressing GFP for 24 hours. Mean fluorescence intensity and percent GFP positive were measured by FACS. E) Global changes in gene expression in Hells/ MEFs compared with wild-type controls as detected by expression microarrays. Black circles, Log2 values of normalized intensities for Hells/ MEFs (y-axis) versus wild-type MEFs (x-axis) are shown. Blue circles, 893 Putative immunome genes (ImmTree/ (http://bioinf.uta.fi/ImmTree/)). Red circles, 1003 ISD-pathway candidates, gene names included for 150 top linear Z-ranked screen hits. P-value <0.01, 0.001, 0.0001 (*, **, ***, respectively), Student's t-test.

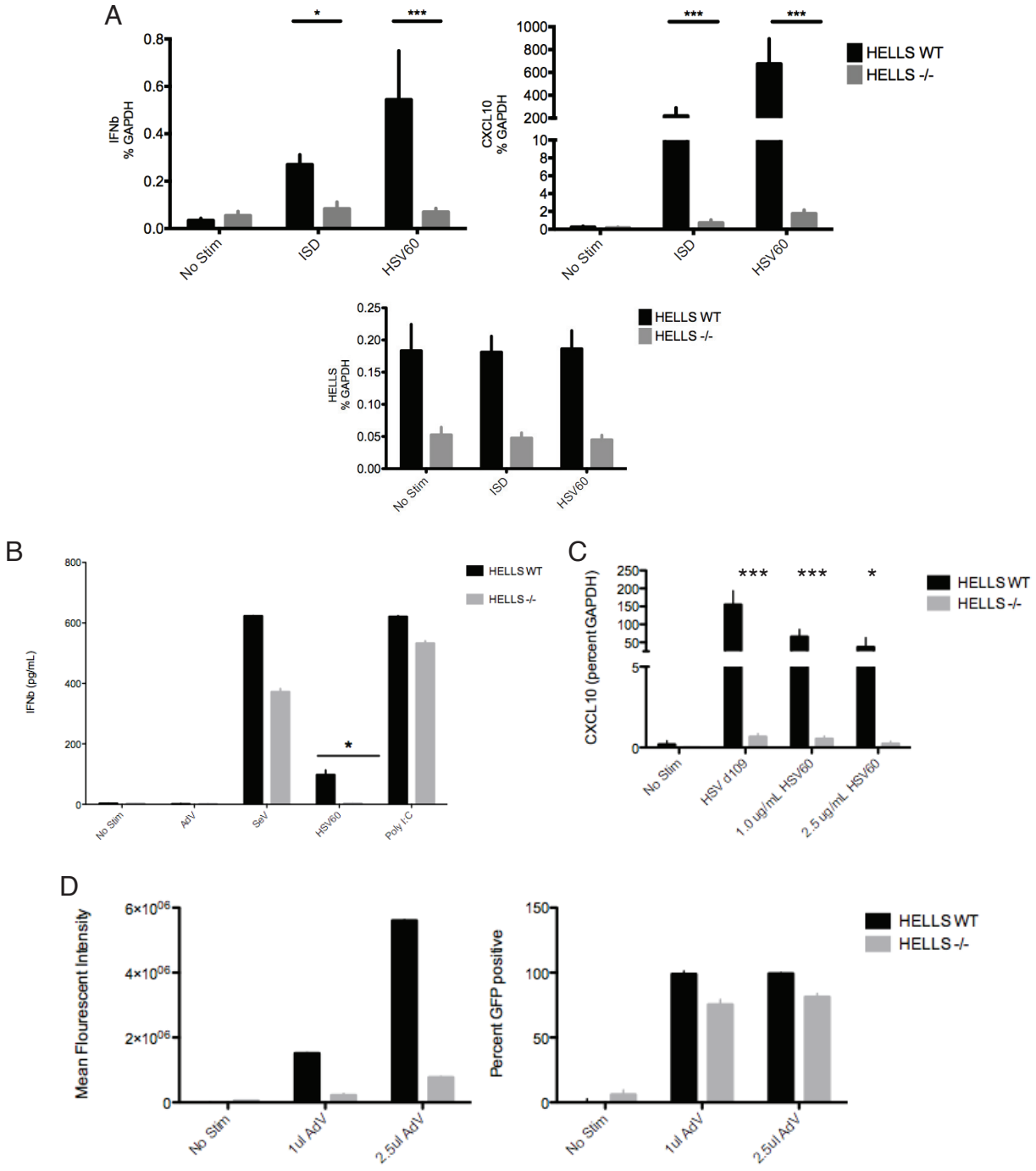


Figure 5.8 (continued)

E

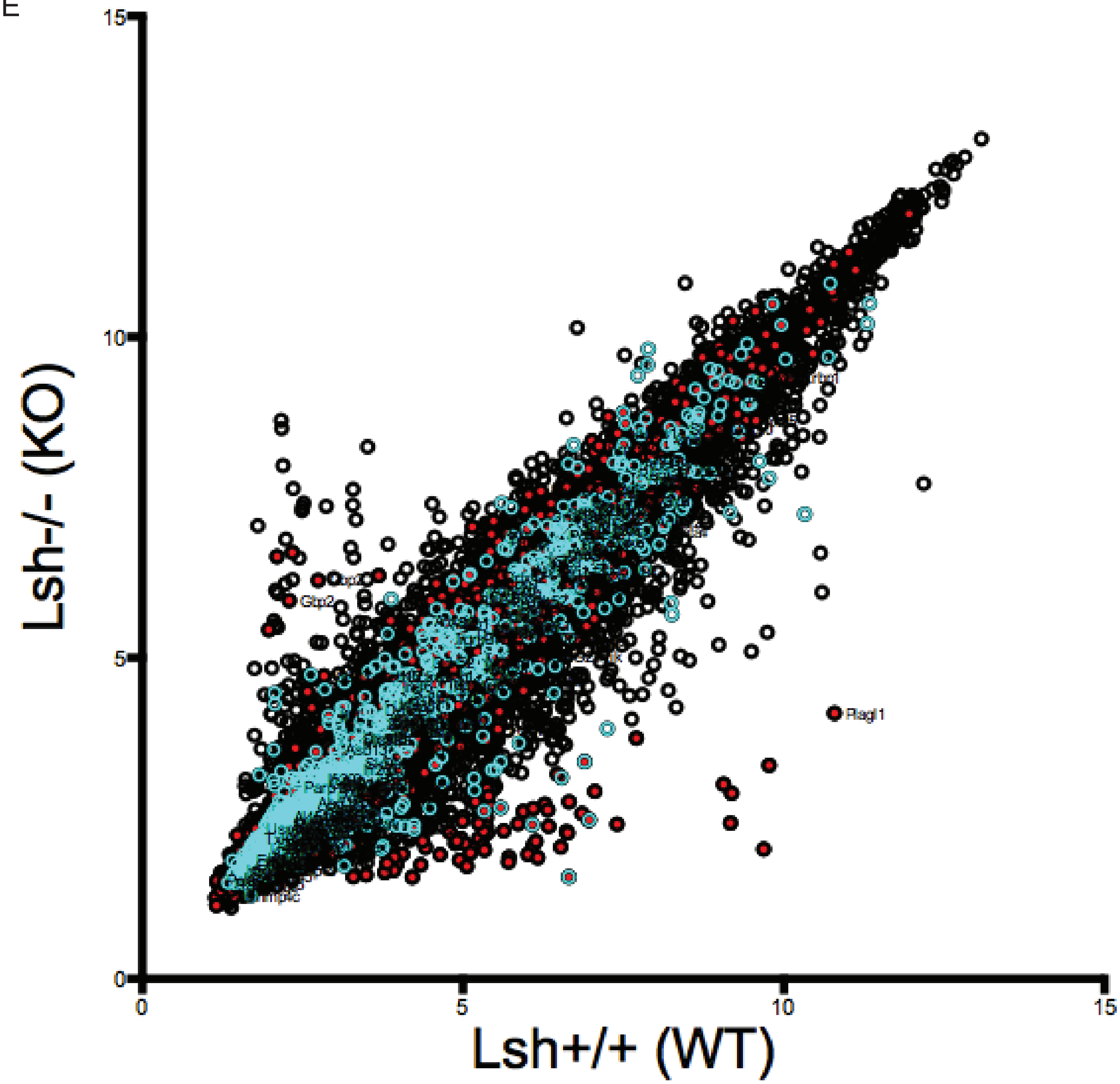


Figure 5.8 (continued)

Adenovirus that constitutively expresses GFP to further dissect the role of Hells in the ISD-pathway. Wild-type and *Hells*$^{-/-}$ MEFs are both infectible, however viral load was greatly reduced in *Hells*$^{-/-}$ MEFs as detected by GFP mean fluorescence intensity.

The putative role of Hells in chromatin remodeling suggested that reduced ISD response could be the result of an epigenetic modification of an ISD-sensing pathway component. A comprehensive genomic map of cytosine methylation for wild-type and *Hells*$^{-/-}$ MEFs revealed global changes in gene expression in *Hells*$^{-/-}$ MEFs compared with wild-type controls as detected by microarrays[211]. Changes in expression coincided with hypo- and hypermethylated promoters suggesting the Hells is a critical epigenetic modulator required for the normal distribution of cytosine methylation throughout the murine genome. Because *Hells* deficiency resulted in significant modification of promoters in more than five percent of the murine genome, we investigated whether changes in gene expression in the absence of *Hells* resulted in the modification of known immune genes (sourced from ImmTree Immunome[214]) or candidates from our siRNA screen (Figure 5.8E). Twenty-eight siRNA candidates were expressed at least 2.5 fold less in *Hells* deficient MEFs than in wild-type MEFs including the ISD-regulated zinc finger protein Plagl1[44]. Additionally, *Hells* deficiency led to the increases in three ISD-sensing candidate proteins including the guanylate binding protein Gbp2. Highly induced following stimulation with Poly (dA:dT), Gbp2 has been implicated in the inhibition of VSV and EMCV replication and may mediate early resistance to *Toxoplasma gondii* infection in mice[43, 215, 216]. Furthermore, our screen for putative cytosolic DNA-binding proteins identified SET complex components including the chromatin-modifying proteins SET and ANP32A, along with the nucleic acid co-receptor HMGB2[140]. Along with recent evidence implicating a role for the SET complex in viral DNA recognition[217], our findings suggest that the SET complex may form a DNA-sensing sensing complex coordinating

the detection, response, modification and degradation of viral and retroelement DNA[140]. The chromatin modifications directed by SET and ANP32A further implicate Hells as a regulator of the ISD-sensing pathway.

Global changes in Hells-mediated gene expression, including putative ISD-sensing components, suggests that epigenetic modifications may influence unknown ISD-sensing pathway components during any aspect of the response to cytosolic DNA.

Finally, we tested the role of putative ISD-sensing factors Endog, Gpr34, Polq and Rasgpr1 by investigating ISD-specific responses in targeted knockouts. In each of these models, we failed to demonstrate the phenotype first recognized in our siRNA screen.

5.8 – *Conclusion*

By combining genomic, proteomic and domain-based data sets with a loss-of-function screen we identified several novel components of the ISD-sensing pathway. The DNA-associated protein Abcf1 appears to be a critical factor in the ISD-sensing pathway. Abcf1 interacts with the SET complex, Hmgb2 and Ifi204, and, as we demonstrated a role in the innate immune response to retroviral infection, Abcf1 may play a part in the early detection of retroviruses. Furthermore, Abcf1 knockdown significantly decreases Ifn-β and ISG responses following infection with HSV-1. We demonstrated that perturbation of Ifit1 and Reep4 leads to reduced type I IFN responses to ISD stimulation. Identified as putative cytosolic DNA-binding components, Ifit1 and Reep4 may play a role in sequestration and delivery of cytosolic DNA. However, the evidence to date (with some caveats) suggests that the siRNAs targeting Ifit1 and Reep4 were off-target, and may not be worth pursuing.

We also identified novel primary and secondary signaling components in the ISD-pathway. We demonstrated that chemical inhibition of several molecules, including the serine-threonine phosphatase Ppp6c and protein tyrosine phosphatase Ptpn1 (also shown with knockout cells), modulates the ISD response. The administration of okadaic acid mimics genetic perturbation of these putative negative regulators, suggesting that these small-molecules may provide a means to boost immune response to DNA viruses or retroviruses[140]. In our recent publication we also demonstrated that administration of small molecules targeting DNA-sensing components, including Tbk1, Hsp90 or Cdc37, to diseases with overactive DNA-sensing pathways (AGS and SLE), may be useful in providing therapeutic benefits [104]. Furthermore, we identified novel components with no known ISD-sensing interactors. The interferon-regulated nuclear body protein Sp110 and the helicase Hells provide new lines of evidence for regulation of the ISD-pathway including the cross-talk with apoptotic pathway components and epigenetic regulation, respectively.

Chapter 6:

Concluding Remarks and Network Analysis

6.1 – Overview: Screening development, analysis and outcome

With the goal of identifying novel components of the ISD-sensing pathway, we integrated genomic, proteomic and domain based datasets to reveal novel components implicated in the detection, signaling and response to cytosolic DNA by way of a loss-of-function genetic perturbation screen and subsequent validation. The resulting candidates provide novel insights into the ISD-sensing pathway and will be a valuable resource to future researchers investigating the cytosolic DNA response.

The development of the screen required the dissection of many basic ISD-pathway attributes. We recognized critical differences in the ISD-pathway between mice and human (such as the RIG-I-dependent response to Poly (dA:dT) in primary human bronchial epithelial cells), ligand-specific responses (ISD vs. Poly (dA:dT), and nucleic acid receptor expression across multiple cell types. We considered many different approaches to detect cell-autonomous ISD-pathway responses. In doing so we developed a number of useful approaches for detecting IFN responses that may be useful for future research. For example, we developed Ifn-β- and Cxcl10-GFP reporters that, while ineffective for detection of ISD-induced signals, may be useful for detecting RNA-based ligands and viral responses. Similarly, we developed robust tools reporting ISRE by luciferase for human and murine systems. While we did not successfully produce a CD-tagging-vector-based YFP-reporter, this unique system is worth revisiting for use in large-scale genomic loss-of-function or gain-of-function screens. Furthermore, we optimized conditions for use of a highly accurate tool for detecting the simultaneous expression of multiple genes by RT-qPCR directly from cell lysates.

The development of our siRNA-based screening system required critical assessment of many variables. To maximize siRNA-directed knockdown, we considered many factors including transfection reagent volume, time of knockdown, media changes

and sell seeding density. The resulting product will continue to be a useful screening tool for future genetic perturbation screens, including the cDNA over-expression library currently being developed at the Broad Institute.

Our candidate gene selection integrated data culled from multiple resources including proteomic, genomic and domain-based datasets. The subsequent siRNA screen revealed many novel factors contributing to the ISD response. Secondary screening of our top candidates yielded putative regulatory components specific to the ISD-pathway. We selected the most compelling of these candidates for subsequent validations. The remaining candidates may prove to be an invaluable resource for future studies.

Most notably, the SILAC screens described in our recent publication will continue to be a valuable resource for cytosolic DNA-binding proteins, as well as Abcf1- and Sting-interacting proteins[140]. Three of our strongest candidates, Abcf1, Ifit1 and Reep4, were identified in our SILAC-based screens. In addition to these candidates, components of the RNA Polymerase III complex, the SET complex, as well as Hmgb2 and Ifit16 were purified. Identification of the SET complex-association with the ISD-sensing pathway yielded one of the most interesting aspects of our screen, potentially linking Trex1 and Abcf1 to early control of exogenous retroelements from HIV infection[140].

The combined microarray dataset yielded compelling candidates that were revealed in our ISD-sensing pathway screen, including Ifit1 and Sp110. The inclusion of helicases in our screening set proved useful as well, resulting in the identification of the chromatin remodeling factor Hells. We also included genes with putative DNA-binding domains with predicted cytoplasmic localization. While we did not pursue candidates from this set further, we identified a number of genes with no previously ascribed ISD-

pathway role, including the nuclear matrix protein Matr3 and DNA fragmentation factor B (DFFB). Our pilot screen of putative negative regulators, consisting of phosphatases and deubiquitinases, yielded many valuable candidates; most notably Ptpn1 and Ppp6c. Candidates from our negative regulator screen may continue to be a valuable source of candidates for small molecule screens targeting regulators of the IFN response. Like Ptpn1 and Ppp6c, we may identify chemical inhibitors with therapeutic potential in treating patients with over-active ISD-pathway etiologies, including SLE, AGS and chilblain lupus.

In addition to Hmgb2, validated in our screen as a DNA-specific component of the nucleic-acid sentinel HMGB-complex, we investigated seven candidates for involvement in the ISD-sensing pathway (Table 6.1). Discussed further below, we proposed putative roles for each of these candidates in the ISD-sensing pathway as sensors (Abcf1, Ifit1, Reep4 and Sp110), negative regulators (Ppp6c and Ptpn1) and transcriptional modifiers of the ISD sensing pathway (Hells), described previously.

6.2 – Protein-protein interaction network analysis

We have proposed roles for each of these candidates based on our own and existing experimental data. To better define their roles in ISD-sensing, it will be important to find their interacting partners. Our candidate genes and their interacting partners make attractive targets for therapeutics. In an effort to find associated proteins, we performed network analysis of each candidate group (sensor, negative regulators, et cetera) using a network association tool (GeneMANIA) that derives associations between proteins from large function data sets that include protein and genetic interactions, pathways, co-expression, co-localization and protein domain similarity[218]. To this end, we generated four network association maps corresponding to the following

Table 6.1
Summary of putative ISD-sensing pathway candidates.

GeneID	Gene	Candidate List	Fold change _ linear	Validation	Putative ISD pathway role	Knockdown/ knockout ISD-pathway response to pathogen	Putative ISD-sensing pathway
224742	Abcf1	SILAC / ARRAY	30.18	multiple siRNAs cDNA rescue	SENSOR	↓ HSV-d109 ↓ VSV A Polyomavirus	SET-complex associated. SILAC-based evidence.
15957	Ifit1	SILAC / ARRAY	12.57	siRNA* - potential off target animal model did not replicate	SENSOR	↓ VSV ↓ WNV ↓ Influenza	Promiscuous nucleic acid sensor, predicted STING interaction
72549	Reep4	SILAC	21.22	multiple siRNAS cDNA rescue unsuccessful	SENSOR	↓ HIV-based-retrovrius A Polyomavirus	Lipid-raft localized sensor?
67857	Ppp6c	PHOSPHATASE	0.64	small molecule inhibtion	NEGATIVE REGULATOR	↑ HIV-based-retrovrius	Secondary signaling, Predicted indirect TBK1 interaction. Ikbke interaction
19246	Ptpn1	PHOSPHATASE	0.68	small molecule inhibtion animal model	NEGATIVE REGULATOR	↑ HIV-based-retrovrius	Secondary signaling, Predicted indirect TBK1 interaction. Ikbke interaction.
15201	Hells	HELICASE	3.55	multiple siRNAs animal model	TRANSCRIPTIONAL MODIFICATION	↓ HSV-d109 ↓ Adenovirus	Transcriptional modification of ISD pathway. Putative IFNAR1 genetic interaction
109032	Sp110	ARRAY	2.57	animal model	SENSOR	Tuberculosis susceptibility gene	SET-complex associated, pro-apoptotis. Indirect association with Mybbp1a

groups: SILAC candidates (Abcf1, Ifit1 and Reep4), Sp110, Hells and putative negative regulators (Ppp6c and Ptpn1)(Figures 6.1-6.4).

To identify novel components of the SILAC-ISD-sensing candidates, we constructed a network map that includes SET complex proteins (Trex1, Apex1 and Hmgb2), Ifi16, Samhd1 and signaling components of the ISD-sensing pathway (Sting, Tbk1, Irf3 and Ifnar1)(Figure 6.1). Putative connections are delineated by functional connection (i.e., physical interactions, co-expression, etc.). Network analysis of the putative sensors reveals interesting connections, including physical interactions between Sting and Ifit family members Ifit1 and Ifi2.

Similarly, we constructed network maps for Hells and Sp110 (Figures 6.2 and 6.3). As there are no known ISD-pathway components associated with Hells, we generated a network map based on proteins previously revealed in our SILAC screen and ISD-signaling pathway components as described previously. Mitogen-activated protein kinase 1 (Mapk1) is predicted to associate with Hells based on yeast-2-hybrid protein-protein interaction data, providing a compelling candidate connecting Hells to Nod-like-receptor and TLR signaling. The pro-apoptotic protein Mybbp1a is predicted to interact with Sp110 and was therefore included in the predictive network analysis of Sp110 and ISD-sensing components. Network analysis revealed a number of co-expressed proteins including Ifit1, Samhd1, and Trex1. Furthermore, Sp110 shares domain similarities to the Autoimmune Regulator gene Aire including the SAND domain, which has been linked to various human diseases including autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED).

To construct a network for putative negative regulators Ppp6c and Ptpn1, in addition to known ISD-pathway signaling components, we added known interactors Jak2, Tyk2, and Ikbke to network analyses (Figure 6.4). Of note is the physical

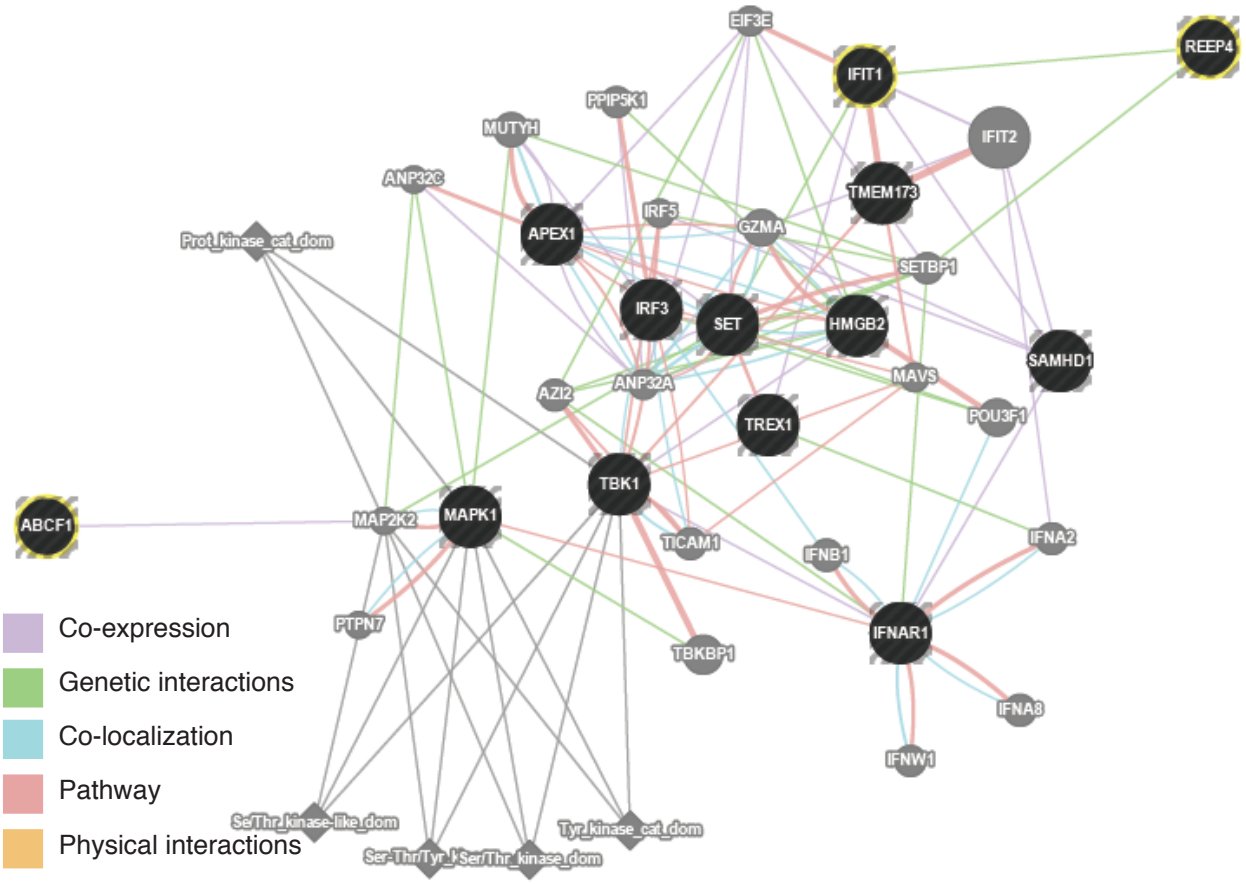


Figure 6.1

Functional association predictions of SILAC candidates Abcf1, Ifit1 and Reep4. Functional association prediction of SILAC candidates (yellow circles) were evaluated in combination with known SET complex members (Set, Apex1, Hmgb2), Trex1, Samhd1, and ISD-sensing pathway signaling components including Sting, Tbk1, Irf3, and Ifnar1 (black circles). Colored lines correspond to network prediction method (see legend).

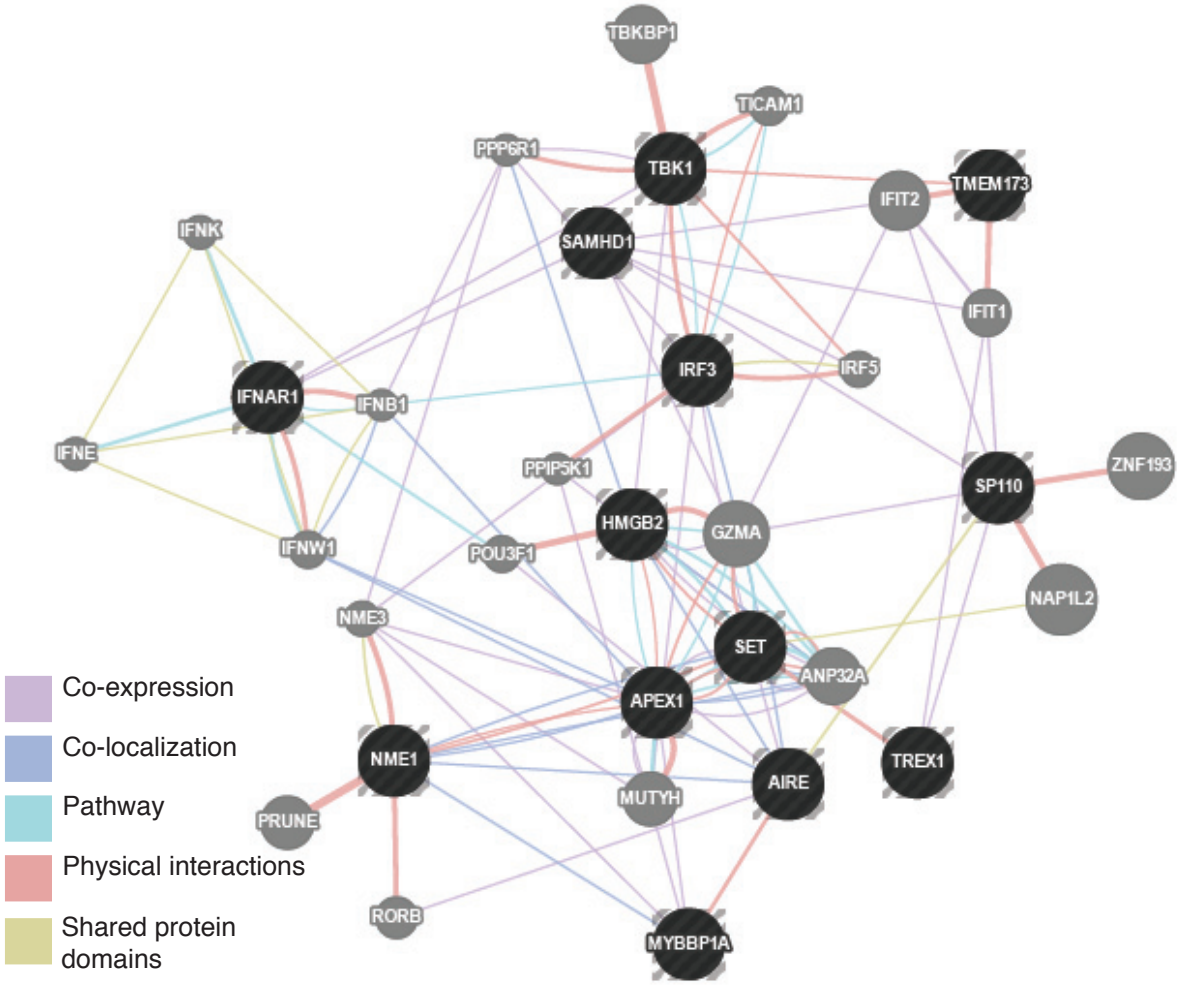


Figure 6.2

Functional association predictions of Sp110 and putative pro-apoptotic binding partners. Functional association prediction of Sp110 and putative binding partner Mybbp1a (yellow circles) were evaluated in combination with known SET complex members (Set, Apex1, Hmgb2), Trex1, Samhd1, and ISD-sensing pathway signaling components including Sting, Tbk1, Irf3, and Ifnar1 (black circles). Colored lines correspond to network prediction method (see legend).

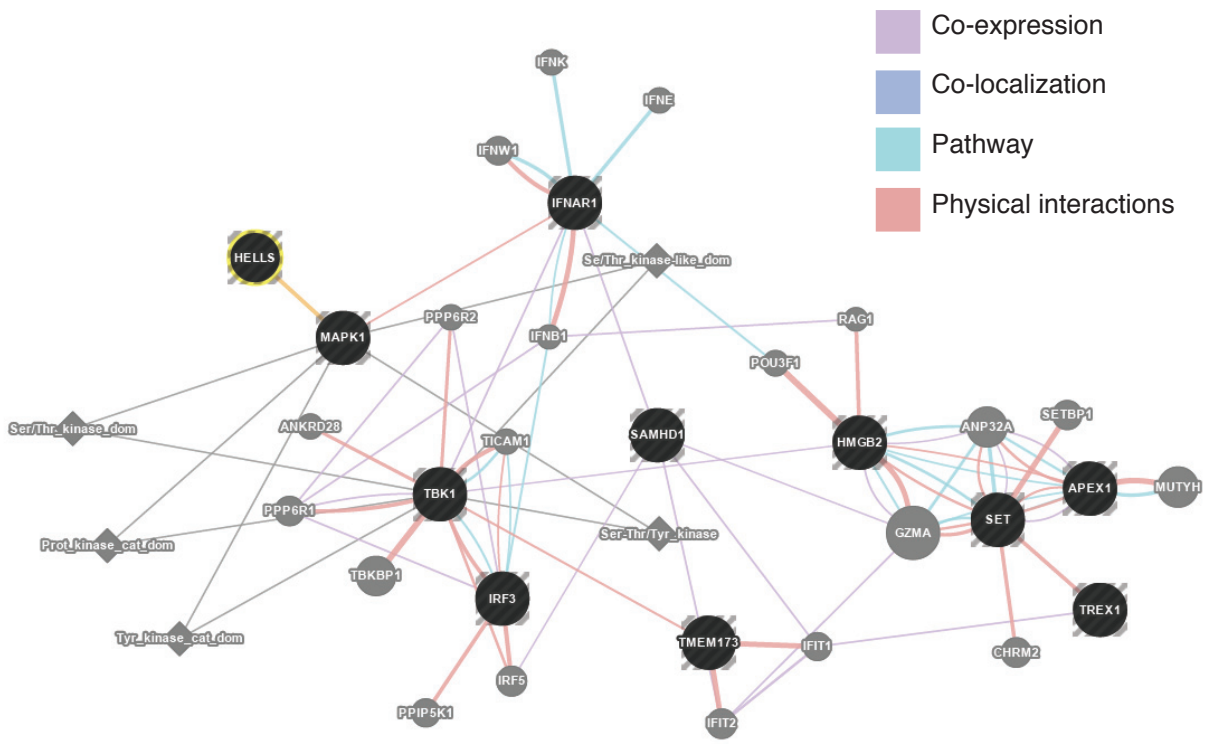


Figure 6.3

Functional association predictions of Hells. Functional association prediction of Hells (yellow circle) was evaluated in combination with known SET complex members (Set, Apex1, Hmgb2), Trex1, Samhd1, and ISD-sensing pathway signaling components including Sting, Tbk1, Irf3, and Ifnar1 (black circles). Colored lines correspond to network prediction method (see legend).

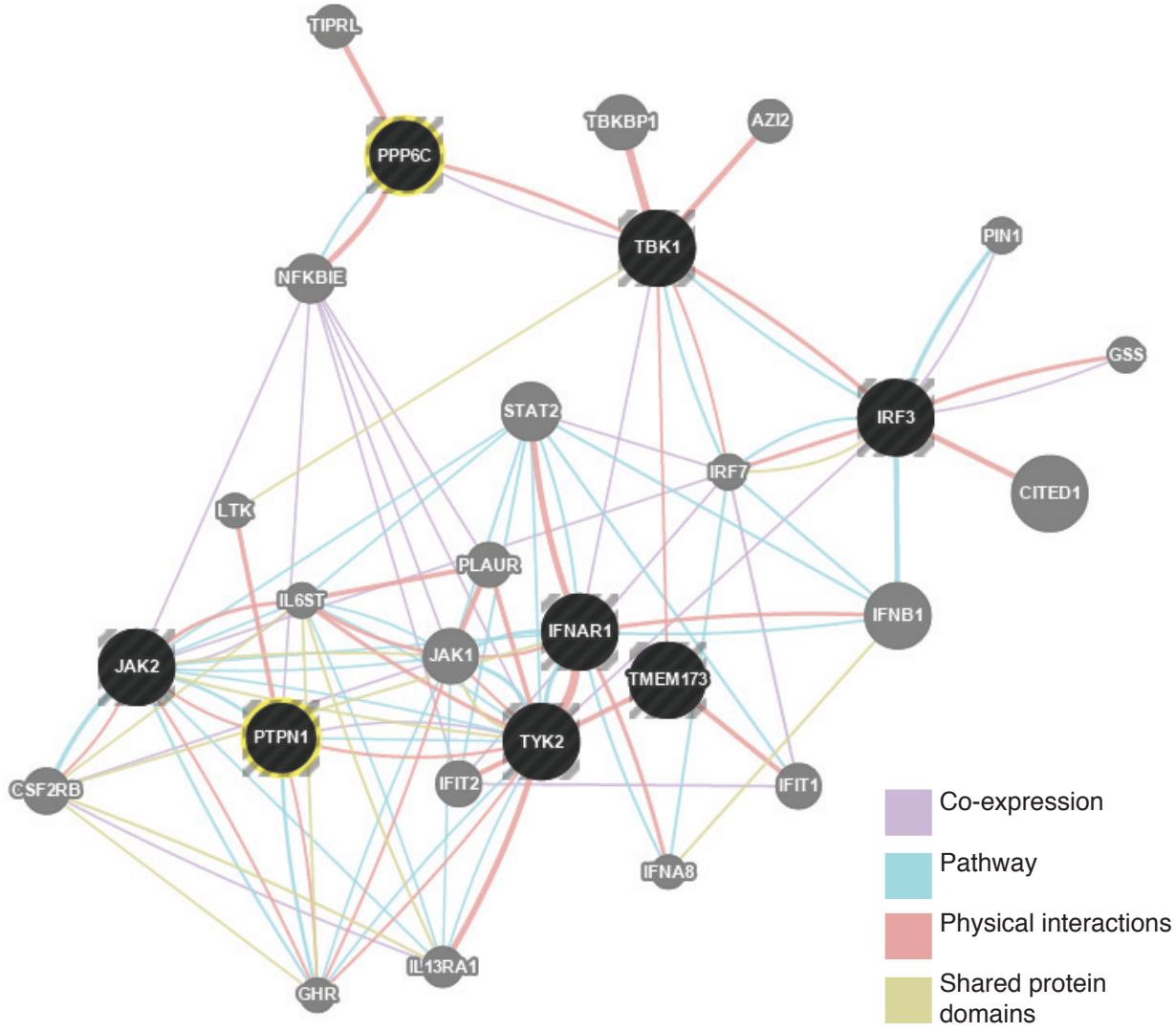


Figure 6.4

Functional association predictions of putative negative regulators Ppp6c and Ptpn1. Functional association prediction of Ppp6c and Ptpn1 (yellow circles) were evaluated in combination with known ISD-sensing pathway signaling components including Sting, Tbk1, Irf3, and Ifnar1 (black circles). Colored lines correspond to network prediction method (see legend).

interaction between Ptpn1 and leukocyte receptor tyrosine kinase Tyk1. Tyk1 shares protein domains with Tbk1 and is interferon inducible. Additionally, Ppp6c interacts directly with Tbk1.

The integration of network analysis provides a compelling layer of data to the role of our candidates in regulating the ISD-sensing pathway. Future studies will interrogate network predictions in the context of our candidate genes to connect and reveal novel roles in the ISD-sensing pathway. Furthermore, proteins revealed in network analysis may be drug-able targets with therapeutic potential for patients with an overacting ISD-sensing pathway.

6.3 – Predictions and concluding remarks

The putative roles of the candidate genes identified in our screen of the ISD-sensing pathway as sensors, negative regulators and chromatin remodelers are summarized in Figure 6.5. The putative cytosolic DNA sensor Abcf1 binds DNA and interacts with the SET complex, whose role in nucleic acid sensing is only beginning to be understood. Recent evidence suggests a role for the SET complex in the early detection of viral infection, including HIV infection, during which SET complex member *Trex1* is sequestered by HIV to eliminate retroviral DNA thus abrogating early viral detection[106]. Although pull down of Abcf1 did not include Sting, knockdown of *Abcf1* reduces the expression of *Sting* regardless of stimulation (Figure 5.3). Perhaps, Abcf1, as part of the SET complex, modulates the expression of Sting through an unknown mechanism.

The putative DNA sensor Sp110 may also play a role in SET complex-based response to DNA. Sp110-interacting protein Mybbp1a interacts with Aire, which in turn co-localizes with SET complex members including Set, Apex1 and Nme1. Though we did

Figure 6.5

Putative roles of candidate genes as sensors, negative regulators and chromatin remodelers of the ISD pathway. Putative DNA sensors Abcf1 and Sp110 may interact with SET complex components and direct ISD-sensing responses via Sting and Tbk1 or through an unknown Sting-independent pathway. Reep4, by way of lipid-raft mediated viral endocytosis may directly detect DNA or recruit a DNA sensor, akin to the antimicrobial peptide LL37 leading to what is presumed to be Sting-dependent-IFN expression. Ifit1 and homologous family members Ifit1b and Ifit1c may separately or in conjunction be involved in the promiscuous detection of viral RNA and DNA. Putative negative regulators Ppp6c and Ptpn1 may inhibit IFN expression by blocking NFκB or Jak1/Tyk2 phosphorylation, respectively. Cytosine methylation patterns directed by Hells may influence the expression of ISD-sensing pathway components during any phase of detection.

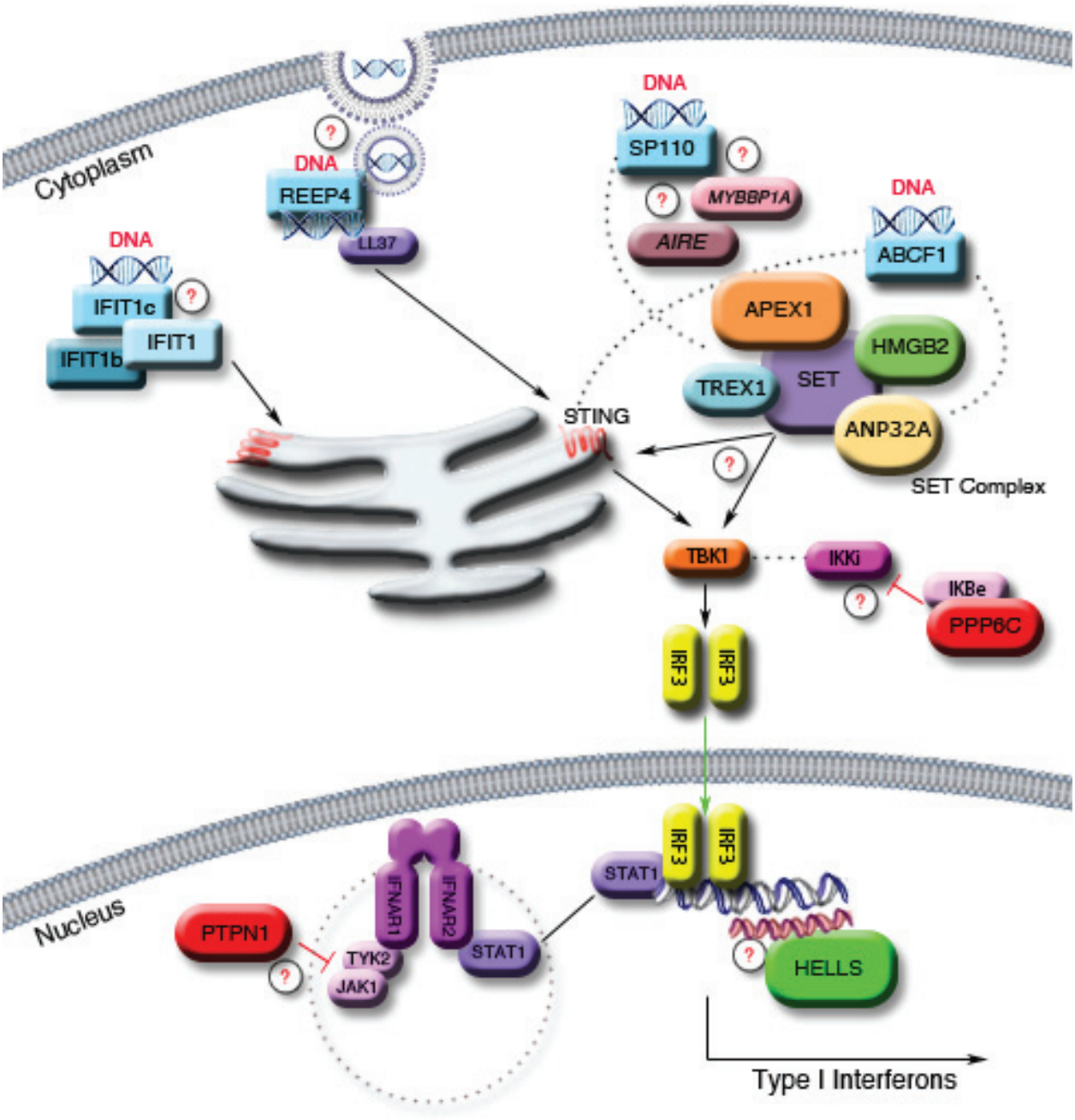


Figure 6.5 (continued)

not identify Sp110 in our initial SILAC-based DNA-binding study, Sp110 may interact with SET complex members through an unknown sensing pathway that further implicates a novel role for Aire in early pathogen detection. While there are many unanswered questions, these new insights into Sp110 suggest a broader role in pathogen response beyond the reported antimicrobial roles following tuberculosis or Listeria infection.

Potential siRNA off-target effects and the high homology of Ifit1 family members Ifit1b and Ifit1c cloud the role of Ifit1 as a putative sensor of the ISD-sensing pathway. However, recent evidence implicates Ifit1 in the sensing of 5' triphosphorylated RNA, supporting the hypothesis that the tetratcopeptide repeats that comprise the Ifit1 protein behave promiscuously and may thus also bind DNA and promote antiviral sensing. We propose that Ifit1, in conjunction with homologous family members Ifit1b and/or Ifit1c, may be involved in the direct detection of DNA leading to the activation of Type I IFNs. An Ifit1-locus targeted knockout approach may reveal the nature of the ISD response obscured by the knockout of only a single Ifit1-superfamily gene.

We also implicated the trafficking protein Reep4 as a novel component of the ISD-sensing pathway. Although we were unable to fully validate the role of Reep4 in the response to DNA, it was isolated as a DNA-interacting protein in our DNA-SILAC screen and additionally demonstrated abrogated DNA-induced responses following knockdown with multiple siRNAs. Reep4 may be involved in lipid-raft mediated endocytosis, similar to REEP family member REEP2. Perhaps Reep4 recruits LL-37, an antimicrobial peptide implicated in the transport of self-DNA into monocytes via lipid rafts, to aid in the detection of DNA viruses that enter the cell via lipid-raft mediated endocytosis, such as Polyomaviruses. Thus, Reep4 may act as a DNA sensor itself or as a trafficking mechanism associated with lipid-rafts that directly recruits DNA sensing components.

In addition to putative DNA sensors, we identified two potential novel negative regulators of the ISD response, Ppp6c and Ptpn1. Though the mechanism of these components needs further clarification, protein-protein interaction data provides compelling insight as to their roles in the ISD-sensing pathway. Ppp6c has been implicated as a substrate for Ikbke, an NFκB inhibitor and may thus regulate ISD-induced IFN expression by preventing phosphorylation of Ikki, thereby inhibiting Tbk1/Ikki co-activation of Irf3 and subsequent IFN expression. Consistent with this hypothesis, knockdown or chemical inhibition of Ppp6c dramatically increases ISD-induced IFN responses. Additionally, we demonstrated a role for protein tyrosine phosphatase Ptpn1 as a negative regulator of the ISD-sensing pathway. Chemical inhibition, siRNA knockdown and MEFs from Ptpn1-deficient mice demonstrate increased IFN responses to DNA. While we have not identified the substrate for Ptpn1 in the DNA-sensing pathway, we predict that Ptpn1 acts as an inhibitory substrate to interferon-receptor signaling molecule Jak1 and/or Tyk2. Understanding the exact mechanism by which DNA stimulation induces the inhibitory capacity of Ptpn1 may have broad therapeutic implications in the regulation of inflammatory responses across autoimmune and infectious diseases.

Finally, we discussed the potential role of the helicase Hells in the ISD-sensing pathway. The genome-wide effect of Hells deficiency on chromatin modification and subsequent gene expression implicates Hells in a potential epigenetic role in the ISD-sensing pathway. We demonstrated that the expression of many immune function genes was significantly changed in the absence of Hells. How global changes in Hells-mediated gene expression influences ISD-sensing pathway components remains unclear and opens a novel avenue to pursue additional ISD-pathway components. It is possible that, in the absence of Hells, DNA sensors are not expressed and thus, in-depth analysis and

loss-of-function screens of differentially expression genes may reveal novel components of the ISD-sensing pathway.

In conclusion, we identified a number of novel ISD pathway components including Abcf1, Ptpn1 and Hells. We validated hits through siRNA-resistant cDNA rescue, chemical inhibition, or targeted knockout. Additionally, we evaluated protein-protein interactions of our strongest validated hits to develop a network model of the ISD pathway. In addition to the identification of novel ISD pathway components, our enriched screening data set may provide a useful resource of candidate genes involved in the response to cytosolic DNA. The resulting data set may prove a useful resource to immunologists seeking to identify factors involved in any aspect of the response to cytosolic DNA and may further reveal therapeutic targets for patients with ISD-sensing pathway-driven diseases.

Chapter 7:

Materials and Methods

Cells, viruses and reagents

HEK293T, A549, HELA, and RAW 264.7 cells were obtained from the American Type Culture Collection. Cells were maintained in DMEM (Mediatech) with 10% FBS (Sigma). RS4-11 cells were a gift from T. Means (Massachusetts General Hospital). Primary HBECs (Lonza, Basel, Switzerland) derived from normal human bronchial epithelium were cultured as previously described[134]. KBM7 cells were provided by J.E. Carette and T.R. Brummelkamp (Whitehead Institute) and maintained as previously described[122]. $p53^{-/-}$ MEFs were derived from $p53^{-/-}$ mice (gift from D.J. Kwiatkowski (Harvard Medical School) and D.M. Sabatini (Massachusetts Institute of Technology)). ISRE-luciferase 293T and $p53^{-/-}$ MEFs were generated as previously described[134]. GFP-expressing IFNb and CXCL10-promoter reporters, IFNb-PLJM6-GFP and CXCL10-PLJM6-GFP, were a gift from A. Luster (Massachusetts General Hospital). $Abcf1^{-/-}$ and wild-type and MEFs were a gift from S. Wilcox (University of British Columbia)[178]. $Ifit1^{-/-}$ and wild-type and MEFs were a gift from M. Diamond (Lerner Research Institute, Cleveland Clinic, Cleveland, Ohio)[189]. $Ptpn1^{-/-}$ MEFs and $Ptpn1^{-/-}$MEFs rescued with Ptpn1 cDNA were a gift from B.G. Neel (Ontario Cancer Institute)[219]. Primary murine lung fibroblasts were derived from lung tissue of 4-8 wk old female C57BL/6 mice. Mouse conventional dendritic cells (cDCs) were prepared from wild-type or B6.C3H-sst1 (Sp110 LoF) mice as previously described[126, 204]. $Hells^{-/-}$ and wild-type and MEFs were a gift from K. Muegge(NCI)[211]. Endog MEFs were a gift from J. Chung (National Heart, Lung, and Blood Institute)[220]. Gpr34 MEFs were provided by T.Schöneberg (University of Leipzig)[221]. Polq MEFs were a gift from J. Schimenti (Cornell)[222]. Rasgpr1 MEFs were a gift from J. Stone (University of Alberta)[223]. Adenovirus, Sendai virus and *Listeria monocytogenes* genomic DNA, were obtained from ATCC. Viruses were used at a multiplicity of infection (MOI) of 1 unless otherwise

indicated. HSV-1 d109 was obtained as a gift from N.A. DeLuca (University of Pittsburgh) and used at an MOI of 1[180].

ISD, shISD, 12bpISD, Random-ISD and HSV60 dsDNA were annealed from oligonucleotides (IDT) as described[44, 106]; sequences are listed in Table 7.1. PR-8 and 3p-RNA *In vitro*–transcribed RNA ligands were synthesized as described[31]. ODN 1668 CpG type B TLR and Poly (dA:dT) were from Invivogen and Poly (I:C) from Enzo Life Sciences. Nucleic acids were mixed with Lipofectamine LTX (Life Technologies) at ratio of 1:3 (wt/vol) in Opti-MEM (Life Technologies) and added to cells at 1 μg/ml (DNA) or 0.1 μg/ml (RNA) unless otherwise indicated. Recombinant IFN-β was obtained from PBL InterferonSource, murine CXCL10 ELISA kit from R&D, NE-PER from Pierce, Luminescent cell viability assay was from Promega (CellTiter Glo). Antibodies used were anti-Human CD40 (14-0409, eBioscience), anti-CD80 (ab64116, Abcam), anti-IFIT1 (SC-134949, Santa Cruz Biotechnology), anti-β-actin (ab6276, Abcam), and anti-HA (High Affinity 3F10; Roche).

CD-Tagging

The pBabe-YFP1 CD tagging plasmid was a gift from A. Sigal (Weizmann Institute of Science) and were utilized following established protocols[130].

Quantitative RT-PCR

Total RNA was extracted using RNeasy Mini kit (Qiagen). cDNA was synthesized using High Capacity cDNA Reverse Transcription Kit (Applied Biosystems). Real-time qPCR was performed using SYBR Green and LightCycler 480 system (Roche). Dual reporter real-time PCR was performed using AptaTaq Master Mix (Roche), Universal Probe Library Probe # 3 (Cat. No. 04685008001, Roche), UPL Mouse GAPDH Gene Assay (Cat. No. 05046211001, Roche). The primers used for qPCR are listed in Table 7.2.

Table 7.1
DNA and RNA ligands

Name	Sequences
ISD	5'-TACAGATCTACTAGTGATCTATGACTGATCTGTACATGATCTACA-3'
	5'-TGTAGATCATGTACAGATCAGTCATAGATCACTAGTAGATCTGTA-3'
shISD	5'-TACAGATCTACTAGTGATCTA-3'
	5'-TAGATCACTAGTAGATCTGTA-3'
12bpISD	5'-AGTAGATCTGTA-3'
	5'-TACAGATCTACT-3'
HSV60	5'-TAAGACACGATGCGATAAAATCTGTTTGTAAAATTTATTAAGGGTACAAATTGCCCTAGC-3'
	5'-GCTAGGGCAATTTGTACCCTTAATAAATTTTACAAACAGATTTTATCGCATCGTGTCTTA-3'
5'-biotin-ISD	5'-biotin-TACAGATCTACTAGTGATCTATGACTGATCTGTACATGATCTACA-3'
	5'-TGTAGATCATGTACAGATCAGTCATAGATCACTAGTAGATCTGTA-3'
Random-ISD	5'-AGTAGAAACAAGGGTGTTTTTTATTATTAAATAAGCTGAAATGAGAAAGT-3'
	5'-ACTTTCTCATTTCAGCTTATTTAATAATAAAAAACACCCTTGTTTCTACT-3'
pBluescript primers of IVT-RNA	
IVT-RNA -PR8	5'-ACTTTCTCATTTCAGCTTATTTAATAATAAAAAACACCCTTGTTTCTACTCCTGTCTC-3'
IVT-RNA -3p-RNA	5'-AAATGTGTGTGTGTGTGGTGCCTGTCTCCCCCTGTCTC-3'

Table 7.2

Quantitative RT-PCR Primers

Species	Gene	GeneID	Forward	Reverse
Mo	Abcf1	224742	5'-AGAAAGCCCGAGTTGTGTTTG-3'	5'-GCCCCCTTGTAGTCGTTGATG-3'
Mo	Aim2	383619	5'-GGCCGCATAGTCATCCTTTA-3'	5'-CAACAGCATTTCCCGGTACT-3'
Mo	Ccl5	20304	5'-GCTGCTTTGCCTACCTCTCC-3'	5'-TCGAGTGACAAACACGACTGC-3'
Mo	Cdc37	12539	5'-GACTACAGCGTTTGGGATCAC-3'	5'-CCCCGGTCCAGTTCCTCTT-3'
Mo	Cxcl10	15945	5'-CCAAGTGCTGCCGTCATTTTC-3'	5'-GGCTCGCAGGGATGATTTCAA-3'
Mo	Dai	58203	5'-TAAGCACCTTCTGAGCTATGACG-3'	5'-AGGGCTACATGGCAAGACTAT-3'
Mo	Dhx58	80861	5'-GGAAGTGATCTTACCTGCTCTGG-3'	5'-TTGCCTCTGTCTACCGTCTCT-3'
Mo	Gapdh	14433	5'-GGCAAATTCAACGGCACAGT-3'	5'-AGATGGTGATGGGCTTCCC-3'
Mo	Hells	15201	5'-TGAGGATGAAAGCTCTTCCACT-3'	5'-ACATTTCCGAACTGGGTCAAAA-3'
Mo	Hmgb2	97165	5'-CGGGGCAAAATGTCCTCGTA-3'	5'-ATGGTCTTCCATCTCTCGGAG-3'
Mo	Ifi44	99899	5'-AACTGACTGCTCGCAATAATGT-3'	5'-GTAACACAGCAATGCCTCTTGT-3'
Mo	Ifih1	71586	5'-AGATCAACACCTGTGGTAACACC-3'	5'-CTCTAGGGCCTCCACGAACA-3'
Mo	Ifit1	15957	5'-CTGAGATGTCACTTCACATGGAA-3'	5'-GTGCATCCCCAATGGGTTCT-3'
Mo	Ifit1 super-family	-	5'-GACCTGGGGCAACTGTGC-3'	5'-CAGCCTTCCTCACAGTCCAT-3'
Mo	Ifit1b	112419	5'-CCAGAGCACAGCAAAATCAA-3'	5'-CGGCTTAAGCTGACTCAACC-3'
Mo	Ifit1c	667373	5'-GCTTTACTGCAGCCAGAACC-3'	5'-GTTGTTGGCACATTCCACAG-3'
Mo	Ifit2	15958	5'-GGAGAGCAATCTGCGACAG-3'	5'-GCTGCCTCATTTAGACCTCTG-3'
Mo	Ifit3	15959	5'-CCTACATAAAGCACCTAGATGGC-3'	5'-ATGTGATAGTAGATCCAGGCGT-3'
Mo	Ifit3 super-family	-	5'-CGAGCAAAAATGTGCTTTGA-3'	5'-GCTCCCCTTCAGCTTCTTCT-3'
Mo	Ifit3B	667370	5'-GTTTGGGAGGCAACACACTT-3'	5'-ATTGTCCCCATAAGCAGCAC-3'
Mo	Ifnb1	15977	5'-CTGGCTTCCATCATGAACAA-3'	5'-AGAGGGCTGTGGTGGAGAA-3'
Mo	IL6	16193	5'-TAGTCCTTCCTACCCCAATTTCC-3'	5'-TTGGTCCTTAGCCACTCCTTC-3'
Mo	Irf3	54131	5'-GAGAGCCGAACGAGGTTCAG-3'	5'-CTTCCAGGTTGACACGTCCG-3'
Mo	Irf7	54123	5'-GAGACTGGCTATTGGGGGAG-3'	5'-GACCGAAATGCTTCCAGGG-3'
Mo	Isg15	100038882	5'-GGTGTCCGTGACTAACTCCAT-3'	5'-TGGAAAGGGTAAGACCGTCCT-3'
Mo	Mtmr3	74302	5'-ATGACTCGTTGGCTACCTGAC-3'	5'-GAACCGGAACCTTCTGGTTAC-3'
Mo	Mx1	17857	5'-GACCATAGGGGTCTTGACCAA-3'	5'-AGACTTGCTCTTTCTGAAAAGCC-3'
Mo	Ptpn1	19246	5'-GGAACTGGGCGGCTATTTACC-3'	5'-CAAAAGGGCTGACATCTCGGT-3'
Mo	Reep4	72549	5'-GCCTGGTAGTGCTCATATTTGG-3'	5'-GCCATGAAGATCGCAAAGACAA-3'
Mo	RIG-I	230073	5'-ACTTGGGTACAACATTGCGAG-3'	5'-GTTCACAAGAATCTGGGGTGTC-3'
Mo	Sp110	109032	5'-ATGAAGGTGAACATCGCCTATG-3'	5'-GGACAGAGGGACCAGATTTTG-3'
Mo	Stat1	20846	5'-TCACAGTGGTTCGAGCTTCAG-3'	5'-GCAAACGAGACATCATAGGCA-3'
Mo	Sting	72512	5'-GGTCACCGCTCCAAATATGTAG-3'	5'-CAGTAGTCCAAGTTCGTGCGA-3'
Mo	Tbk1	56480	5'-ACTGGTGATCTCTATGCTGTCA-3'	5'-TTCTGGAAGTCCATACGCATTG-3'
Hu	CXCL10	3627	5'-GTGGCATTCAAGGAGTACCTC-3'	5'-TGATGGCCTTCGATTCTGGATT-3'
Hu	IFIT1	3434	5'-TCAGGTCAAGGATAGTCTGGAG-3'	5'-AGGTTGTGTATTCCCACACTGTA-3'
Hu	IL6	3569	5'-ACTCACCTCTTCAGAACGAATTG-3'	5'-CCATCTTTGGAAGGTTCAGGTTG-3'
Hu	MX1	4599	5'-TGTTTCCGAAGTGGACATCGC-3'	5'-CCATTCAGTAATAGAGGGTGGGA-3'
Hu	AIM2	9447	5'-CACCAAAAGTCTCTCCTCATGTT-3'	5'-AAACCCTTCTCTGATAGATTCCTG-3'
Hu	DAI	81030	5'-GAGGAGTCGCAGGGTCTG-3'	5'-GACTTCTGGATTGTGTGTCTGC-3'
Hu	IFNB1	3456	5'-ATGACCAACAAGTGTCTCCTCC-3'	5'-GGAATCCAAGCAAGTTGTAGCTC-3'
Hu	RIGi	23586	5'-TGGACCCTACCTACATCCTGA-3'	5'-GGCCCTTGTTGTTTTTCTCA-3'
Hu	TNFA	7124	5'-CCTCTCTCTAATCAGCCCTCTG-3'	5'-GAGGACCTGGGAGTAGATGAG-3'

RNA interference screen

We seeded 750 $p53^{-/-}$ MEFs per well in 96-well plates in 60% DMEM and 40% Opti-MEM. siRNA (25 nM) was complexed with 0.5 µl Lipofectamine RNAiMax (Life Technologies) in Opti-MEM, incubated for 12 min and added to wells. Cells were transfected with DNA 72 h later. Supernatants were collected 26 h later, and CXCL10 was quantified by ELISA. Cell viability was estimated by CellTiter-Glo Luminescent Cell Viability Assay (Promega); CellTiter-Glo values below 3.75×10^5were considered toxic. Dharmacon siGENOME SMARTpools from Harvard ICCB were used for screening. ON-TARGETplus Nontargeting Pool (Dharmacon) was used as negative control (siCtrl). Individual siRNAs are listed in Table 7.3.

shRNA Knockdowns

High-titer lentiviruses expressing shRNAs were obtained from The Broad RNAi Platform and used to infect BMDCs as previously described[126].

Statistics

Statistical significance was determined by paired Student's t-test, unless otherwise noted. $P < 0.05$ was considered statistically significant.

Table 7.3

siRNA sequences

Species	Gene	GeneID	Sense sequence	Notes
Mo	Abcf1	224742	Dharmacon siGENOME SMARTpool	siABCF1_1
Mo	Abcf1	224742	5'-AGGAAGUCCUGACUCGAAA-3'	siABCF1_2
Mo	Abcf1	224742	5'-CGATGATAGTGATGAGAGA-3'	siABCF1_3
Mo	Irf3	54131	5'-CGGACAAGCUUGUGAAGGA-3'	siIrf3
Mo	Cdc37	12539	5'-GCGCCAAGCUGCGAAUAGA-3'	siCdc37
Mo	Ptpn1	19246	5'-GACCACAGUCGGAUUAAAU-3'	siPtpn1
Mo	Tiparp	99929	5'-GAAACAUCACACCGUAUUG-3'	siTiparp
Mo	Mdp1	67881	5'-AACGUUAACUCAAGGAUUA-3'	siMdp1
Mo	Ppp6c	67857	5'-GCACGAAGGCUAUAAGUUU-3'	siPpp6c
Mo	Asb13	142688	5'-UAGAGAAAGUCGCCAAGUU-3'	siAsb13
Mo	Trim56	384309	5'-GAAGCCAACUUGCGCUCUG-3'	siTrim56
Mo	Usp49	224836	5'-GACCUGAAGUUGCUGAGAA-3'	siUsp49
Mo	Reep4	72549	5'-GAUGAUCUGUCGCCUGGUA-3'	siReep4
Mo	Cyb5r3	109754	5'-CCAAUGGGCUACUGGUCUA-3'	siCyb5r3
Mo	Numa1	101706	5'-GGACGGCCAUUCUCUAGUA-3'	siNuma1
Mo	Wdr77	70465	5'-GGGUGUCACUAGACUGGUA-3'	siWdr77
Mo	Asf1a	66403	5'-AAUCUACAGUCCCUUCUUU-3'	siAsf1a
Mo	Tbk1	56480	5'-CAGACUAGCUUAUAAUGAA-3'	siTbk1
Mo	Mtmr3	74302	5'-GGGAAGAGGUGCCUGCUAU-3'	siMtmr3
Mo	Tmem173	72512	5'-GGAUCCGAAUGUUCAAUCA-3'	siSting
Mo	RIG-I	230073	Dharmacon ON-TARGETplus SMARTpool	siRIG-I
Mo	Ifnar1	15975	Dharmacon siGENOME SMARTpool	siIfnar1
Mo	Jak1	16451	Dharmacon siGENOME SMARTpool	siJak1
Mo	Abcf1	224742	Sigma Cat. No. SASI_Mm01_00125927	siABCF1_4
Mo	Abcf1	224742	Sigma Cat. No. SASI_Mm02_00322866	siABCF1_5
Mo	Abcf1	224742	Sigma Cat. No. SASI_Mm02_00322867	siABCF1_6
Mo	Abcf1	224742	Sigma Cat. No. SASI_Mm01_00125919	siABCF1_7
Mo	Abcf1	224742	Sigma Cat. No. SASI_Mm02_00322868	siABCF1_8
Mo	Abcf1	224742	Sigma Cat. No. SASI_Mm02_00322869	siABCF1_9
Mo	Abcf1	224742	Sigma Cat. No. SASI_Mm02_00322870	siABCF1_10
Mo	Abcf1	224742	Sigma Cat. No. SASI_Mm01_00125920	siABCF1_11
Mo	Abcf1	224742	Sigma Cat. No. SASI_Mm02_00322871	siABCF1_12
Mo	Abcf1	224742	Sigma Cat. No. SASI_Mm02_00322872	siABCF1_13
Mo	Abcf1	224742	Sigma Cat. No. SASI_Mm01_00125918	siABCF1_14
Mo	Abcf1	224742	Sigma Cat. No. SASI_Mm01_00125921	siABCF1_15
Mo	Abcf1	224742	Sigma Cat. No. SASI_Mm01_00125922	siABCF1_16
Mo	Abcf1	224742	Sigma Cat. No. SASI_Mm01_00125923	siABCF1_17
Mo	Abcf1	224742	Sigma Cat. No. SASI_Mm01_00125924	siABCF1_18
Mo	Abcf1	224742	Sigma Cat. No. SASI_Mm01_00125925	siABCF1_19
Mo	Abcf1	224742	Sigma Cat. No. SASI_Mm01_00125926	siABCF1_20
Mo	Abcf1	224742	Sigma esiRNA Cat. No. EMU008591	esiABCF1
Mo	Ifit1	15957	5'-CCACATTTGAGATGGCCTA-3'	siIFIT1_1
Mo	Ifit1	15957	5'-ACATAGAGCAGGATATTCA-3'	siIFIT1_2
Mo	Ifit1	15957	5'-GAGGAAGGCTGTCCGGTTA-3'	siIFIT1_3
Mo	Ifit1	15957	5'-GAAACATCGCGTAGACAAA-3'	siIFIT1_4
Mo	Ifit1	15957	5'-CAACAAAUCUCCCAACUGATT-3'	siIFIT1_4
Mo	Ifit1	15957	5'-GGCUUACUAGAGAAUUAAATT-3'	siIFIT1_5
Mo	Ifit1	15957	5'-CCUGUGAGCUGUAGAAUCATT-3'	siIFIT1_6
Mo	Ifit1	15957	5'-CAGGATATTCACCTCCGCTAT-3'	siIFIT1_7
Mo	Ifit1	15957	5'-AAGCGACCCACAAATAACAAA-3'	siIFIT1_8
Mo	Ifit1	15957	5'-CCAAGGATACATACTCACTTA-3'	siIFIT1_9
Mo	Ifit1	15957	5'-CAGGTAGTCTTGGCCTATATA-3'	siIFIT1_10

Species	Gene	GeneID	Sense sequence	Notes
Mo	Ifit1	15957	5'-CTGATGGATTATGATCTCTAA-3'	siIFIT1_11
Mo	Ifit1	15957	5'-TTGGGCTAGTCTACAAACTGA-3'	siIFIT1_12
Mo	Ifit1	15957	5'-CAGGCCATTCATGAATTTCAA-3'	siIFIT1_13
Mo	Ifit1	15957	5'-CACCTCTGTGACGTTAATATA-3'	siIFIT1_14
Mo	Ifit1	15957	5'-AAGATTTAGATGACAACTTTA-3'	siIFIT1_15
Mo	Ifit1	15957	5'-CTGGCTACTTACATTATCAAA-3'	siIFIT1_16
Mo	Ifit1b	112419	5'-GUAGGAAUGUUACUAAUCATT-3'	siIFIT1B_1
Mo	Ifit1b	112419	5'-ATCCCAGTTGTTATCAATAAA-3'	siIFIT1B_2
Mo	Ifit1b	112419	5'-GGCUUAAAAUUUAGUAGGATT-3'	siIFIT1B_3
Mo	Ifit1b	112419	5'-GCUCUGAAACGGAUACAGATT-3'	siIFIT1B_4
Mo	Ifit1b	112419	5'-TAACCCTATTATATTCGATAA-3'	siIFIT1B_5
Mo	Ifit1b	112419	5'-CAGATTGGCCTTTGCCATAAA-3'	siIFIT1B_6
Mo	Ifit1b	112419	5'-AAGGAGGATCATGTCCTGGAA-3'	siIFIT1B_7
Mo	Ifit1b	112419	5'-TTGGTCTTTATTCACCACTAA-3'	siIFIT1B_8
Mo	Ifti1c	667373	5'-AAACGACUGAUUCAAAUAATT-3'	siIFIT1C_4
Mo	Ifti1c	667373	5'-CUUCAAACAUGCAACCAAGTT-3'	siIFIT1C_5
Mo	Ifti1c	667373	5'-GCACUUCGCUCCAACAUUUTT-3'	siIFIT1C_6
Mo	Ifti1c	667373	5'-ATGCCTGATTTGGAAGTAAGA-3'	siIFIT1C_7
Mo	Ifti1c	667373	5'-TTGGAAGATCATATTCAGCAA-3'	siIFIT1_C_7
Mo	Ifti1c	667373	5'-AAGAAACGACTGATTCAAATA-3'	siIFIT1_C_8
Mo	Ifti1c	667373	5'-AAGCTTCAAACATGCAACCAA-3'	siIFIT1_C_9
Mo	Ifti1c	667373	5'-CAGGATATCCTCGCAGCCCTA-3'	siIFIT1_C_10
Mo	Ifit2	15958	Dharmacon ON-TARGETplus SMARTpool	siIFT2
Mo	Ifit1 super-family	-	5'-AGAAGAUCCAUACCUUAAATT-3'	siIFIT1_F_1
Mo	Ifit1 super-family	-	5'-CCAGAAGAUCCAUACCUUATT-3'	siIFIT1_F_2
Mo	Ifit1 super-family	-	5'-GGUGUGCAAGGAAUUUUCATT-3'	siIFIT1_F_3
Mo	Ifit1 super-family	-	5'-GCCUAGAAGAAAAACCUGATT-3'	siIFIT1_F_4
Mo	Ifit1 super-family	-	5'-CGCCUAAAGUGUGGAAGAATT-3'	siIFIT1_F_5
Mo	Ifit1 super-family	-	5'-GCCCAGAACUUGAAUGUGATT-3'	siIFIT1_F_6
Mo	Ifit1 super-family	-	5'-CAGAGCTTGAAAGAAGCTGAA-3'	siIFIT1_F_7
Mo	Ifit1 super-family	-	5'-CAAGGTGGAGAAGGTGTGCAA-3'	siIFIT1_F_8
Mo	Ifit1 super-family	-	5'-TAAAGTTCTCCTTGCCCTAAA-3'	siIFIT1_F_9
Mo	Ifit1 super-family	-	5'-CAACTGTGCCTGGCTGCATTA-3'	siIFIT1_F_10
Mo	Ifit1 super-family	-	5'-CAGGCTGGAGTGTGCTGAGAT-3'	siIFIT1_F_11
Mo	Ifit1 super-family	-	5'-AAGGAAGTATGTCCAGTCATA-3'	siIFIT1_F_12
Mo	Ifit3	15959	Dharmacon ON-TARGETplus SMARTpool	siIFIT3
Mo	Ifit3 super-family	-	5'-ATGGCAGAACTGAGACGATTA-3'	siIFIT3_F_1
Mo	Ifit3 super-family	-	5'-ACGCCTGGATCTACTATCACA-3'	siIFIT3_F_2
Mo	Ifit3 super-family	-	5'-TACGTTGACAAGGTGAGACAA-3'	siIFIT3_F_3
Mo	Ifit3 super-family	-	5'-CACCTGGAATTTATTCAGGGA-3'	siIFIT3_F_4
Hu	AIM2	9447	Dharmacon ON-TARGETplus SMARTpool	siAIM2
Hu	DAI	81030	Dharmacon ON-TARGETplus SMARTpool	siDAI
Hu	IFNB1	3456	Dharmacon ON-TARGETplus SMARTpool	siIFNb
Hu	IRF3	3661	Dharmacon ON-TARGETplus SMARTpool	siIRF3
Hu	RIGi	23586	Dharmacon ON-TARGETplus SMARTpool	siRIGi

References

1. Medzhitov, R. and C.A. Janeway, Jr., *Innate immunity: the virtues of a nonclonal system of recognition.* Cell, 1997. **91**(3): p. 295-8.

2. Iwasaki, A. and R. Medzhitov, *Regulation of adaptive immunity by the innate immune system.* Science, 2010. **327**(5963): p. 291-5.

3. Ausubel, F.M., *Are innate immune signaling pathways in plants and animals conserved?* Nat Immunol, 2005. **6**(10): p. 973-9.

4. Didierlaurent, A., M. Simonet, and J.C. Sirard, *Innate and acquired plasticity of the intestinal immune system.* Cell Mol Life Sci, 2005. **62**(12): p. 1285-7.

5. Barbalat, R., et al., *Nucleic acid recognition by the innate immune system.* Annual review of immunology, 2011. **29**: p. 185-214.

6. Atianand, M.K. and K.A. Fitzgerald, *Molecular basis of DNA recognition in the immune system.* J Immunol, 2013. **190**(5): p. 1911-8.

7. Alexopoulou, L., et al., *Recognition of double-stranded RNA and activation of NF-kappaB by Toll-like receptor 3.* Nature, 2001. **413**(6857): p. 732-8.

8. Lund, J.M., et al., *Recognition of single-stranded RNA viruses by Toll-like receptor 7.* Proc Natl Acad Sci U S A, 2004. **101**(15): p. 5598-603.

9. Heil, F., et al., *Species-specific recognition of single-stranded RNA via toll-like receptor 7 and 8.* Science, 2004. **303**(5663): p. 1526-9.

10. Diebold, S.S., et al., *Innate antiviral responses by means of TLR7-mediated recognition of single-stranded RNA.* Science, 2004. **303**(5663): p. 1529-31.

11. Bauer, S., et al., *Human TLR9 confers responsiveness to bacterial DNA via species-specific CpG motif recognition.* Proc Natl Acad Sci U S A, 2001. **98**(16): p. 9237-42.

12. Hemmi, H., et al., *A Toll-like receptor recognizes bacterial DNA.* Nature, 2000. **408**(6813): p. 740-5.

13. Kawai, T. and S. Akira, *The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors.* Nat Immunol, 2010. **11**(5): p. 373-84.

14. Honda, K. and T. Taniguchi, *IRFs: master regulators of signalling by Toll-like receptors and cytosolic pattern-recognition receptors.* Nat Rev Immunol, 2006. **6**(9): p. 644-58.

15. Thompson, M.R., et al., *Pattern recognition receptors and the innate immune response to viral infection.* Viruses, 2011. **3**(6): p. 920-40.

16. Tabeta, K., et al., *Toll-like receptors 9 and 3 as essential components of innate immune defense against mouse cytomegalovirus infection.* Proc Natl Acad Sci U S A, 2004. **101**(10): p. 3516-21.

17. Krug, A., et al., *TLR9-dependent recognition of MCMV by IPC and DC generates coordinated cytokine responses that activate antiviral NK cell function.* Immunity, 2004. **21**(1): p. 107-19.

18. Delale, T., et al., *MyD88-dependent and -independent murine cytomegalovirus sensing for IFN-alpha release and initiation of immune responses in vivo.* J Immunol, 2005. **175**(10): p. 6723-32.

19. Lund, J., et al., *Toll-like receptor 9-mediated recognition of Herpes simplex virus-2 by plasmacytoid dendritic cells.* J Exp Med, 2003. **198**(3): p. 513-20.

20. Hochrein, H., et al., *Herpes simplex virus type-1 induces IFN-alpha production via Toll-like receptor 9-dependent and -independent pathways.* Proc Natl Acad Sci U S A, 2004. **101**(31): p. 11416-21.

21. Zhu, J., X. Huang, and Y. Yang, *Innate immune response to adenoviral vectors is mediated by both Toll-like receptor-dependent and -independent pathways.* J Virol, 2007. **81**(7): p. 3170-80.

22. Kawai, T., et al., *IPS-1, an adaptor triggering RIG-I- and Mda5-mediated type I interferon induction.* Nature immunology, 2005. **6**(10): p. 981-8.

23. Meylan, E., et al., *Cardif is an adaptor protein in the RIG-I antiviral pathway and is targeted by hepatitis C virus.* Nature, 2005. **437**(7062): p. 1167-72.

24. Seth, R.B., et al., *Identification and characterization of MAVS, a mitochondrial antiviral signaling protein that activates NF-kappaB and IRF 3.* Cell, 2005. **122**(5): p. 669-82.

25. Xu, L.G., et al., *VISA is an adapter protein required for virus-triggered IFN-beta signaling.* Molecular cell, 2005. **19**(6): p. 727-40.

26. Ishikawa, H. and G.N. Barber, *STING is an endoplasmic reticulum adaptor that facilitates innate immune signalling.* Nature, 2008. **455**(7213): p. 674-8.

27. Barber, G.N., *Cytoplasmic DNA innate immune pathways.* Immunol Rev, 2011. **243**(1): p. 99-108.

28. Diebold, S.S., et al., *Viral infection switches non-plasmacytoid dendritic cells into high interferon producers.* Nature, 2003. **424**(6946): p. 324-8.

29. Yoneyama, M., et al., *The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses.* Nature immunology, 2004. **5**(7): p. 730-7.

30. Kato, H., et al., *Length-dependent recognition of double-stranded ribonucleic acids by retinoic acid-inducible gene-I and melanoma differentiation-associated gene 5.* J Exp Med, 2008. **205**(7): p. 1601-10.

31. Hornung, V., et al., *5'-Triphosphate RNA is the ligand for RIG-I.* Science, 2006. **314**(5801): p. 994-7.

32. Yoneyama, M. and T. Fujita, *RNA recognition and signal transduction by RIG-I-like receptors.* Immunol Rev, 2009. **227**(1): p. 54-65.

33. Satoh, T., et al., *LGP2 is a positive regulator of RIG-I- and MDA5-mediated antiviral responses.* Proc Natl Acad Sci U S A, 2010. **107**(4): p. 1512-7.

34. Venkataraman, T., et al., *Loss of DExD/H box RNA helicase LGP2 manifests disparate antiviral responses.* J Immunol, 2007. **178**(10): p. 6444-55.

35. Girardin, S.E., et al., *Nod2 is a general sensor of peptidoglycan through muramyl dipeptide (MDP) detection.* J Biol Chem, 2003. **278**(11): p. 8869-72.

36. Sabbah, A., et al., *Activation of innate immune antiviral responses by Nod2.* Nat Immunol, 2009. **10**(10): p. 1073-80.

37. Rotem, Z., R.A. Cox, and A. Isaacs, *Inhibition of virus multiplication by foreign nucleic acid.* Nature, 1963. **197**: p. 564-6.

38. Jensen, K.E., et al., *Interferon Responses of Chick Embryo Fibroblasts to Nucleic Acids and Related Compounds.* Nature, 1963. **200**: p. 433-4.

39. Suzuki, K., et al., *Activation of target-tissue immune-recognition molecules by double-stranded polynucleotides.* Proc Natl Acad Sci U S A, 1999. **96**(5): p. 2285-90.

40. Ishii, K.J., et al., *Genomic DNA released by dying cells induces the maturation of APCs.* J Immunol, 2001. **167**(5): p. 2602-7.

41. Yasuda, K., et al., *Macrophage activation by a DNA/cationic liposome complex requires endosomal acidification and TLR9-dependent and -independent pathways.* J Leukoc Biol, 2005. **77**(1): p. 71-9.

42. Martin, D.A. and K.B. Elkon, *Intracellular mammalian DNA stimulates myeloid dendritic cells to produce type I interferons predominantly through a toll-like receptor 9-independent pathway.* Arthritis Rheum, 2006. **54**(3): p. 951-62.

43. Ishii, K.J., et al., *A Toll-like receptor-independent antiviral response induced by double-stranded B-form DNA.* Nature immunology, 2006. **7**(1): p. 40-8.

44. Stetson, D.B. and R. Medzhitov, *Recognition of cytosolic DNA activates an IRF3-dependent innate immune response.* Immunity, 2006. **24**(1): p. 93-103.

45. Davis, B.K., H. Wen, and J.P. Ting, *The inflammasome NLRs in immunity, inflammation, and associated diseases.* Annu Rev Immunol, 2011. **29**: p. 707-35.

46. Mariathasan, S. and D.M. Monack, *Inflammasome adaptors and sensors: intracellular regulators of infection and inflammation.* Nat Rev Immunol, 2007. **7**(1): p. 31-40.

47. Muruve, D.A., et al., *The inflammasome recognizes cytosolic microbial and host DNA and triggers an innate immune response.* Nature, 2008. **452**(7183): p. 103-7.

48. Burckstummer, T., et al., *An orthogonal proteomic-genomic screen identifies AIM2 as a cytoplasmic DNA sensor for the inflammasome.* Nature immunology, 2009. **10**(3): p. 266-72.

49. Roberts, T.L., et al., *HIN-200 proteins regulate caspase activation in response to foreign cytoplasmic DNA.* Science, 2009. **323**(5917): p. 1057-60.

50. Fernandes-Alnemri, T., et al., *AIM2 activates the inflammasome and cell death in response to cytoplasmic DNA.* Nature, 2009. **458**(7237): p. 509-13.

51. Hornung, V., et al., *AIM2 recognizes cytosolic dsDNA and forms a caspase-1-activating inflammasome with ASC.* Nature, 2009. **458**(7237): p. 514-8.

52. Rathinam, V.A., et al., *The AIM2 inflammasome is essential for host defense against cytosolic bacteria and DNA viruses.* Nat Immunol, 2010. **11**(5): p. 395-402.

53. Belhocine, K. and D.M. Monack, *Francisella infection triggers activation of the AIM2 inflammasome in murine dendritic cells.* Cell Microbiol, 2012. **14**(1): p. 71-80.

54. Kim, S., et al., *Listeria monocytogenes is sensed by the NLRP3 and AIM2 inflammasome.* Eur J Immunol, 2010. **40**(6): p. 1545-51.

55. Warren, S.E., et al., *Cutting edge: Cytosolic bacterial DNA activates the inflammasome via Aim2.* J Immunol, 2010. **185**(2): p. 818-21.

56. Saiga, H., et al., *Critical role of AIM2 in Mycobacterium tuberculosis infection.* Int Immunol, 2012. **24**(10): p. 637-44.

57. Ishikawa, H., Z. Ma, and G.N. Barber, *STING regulates intracellular DNA-mediated, type I interferon-dependent innate immunity.* Nature, 2009. **461**(7265): p. 788-92.

58. Takaoka, A., et al., *DAI (DLM-1/ZBP1) is a cytosolic DNA sensor and an activator of innate immune response.* Nature, 2007. **448**(7152): p. 501-5.

59. Wang, Z., et al., *Regulation of innate immune responses by DAI (DLM-1/ZBP1) and other DNA-sensing molecules.* Proc Natl Acad Sci U S A, 2008. **105**(14): p. 5477-82.

60. Ishii, K.J., et al., *TANK-binding kinase-1 delineates innate and adaptive immune responses to DNA vaccines.* Nature, 2008. **451**(7179): p. 725-9.

61. Ablasser, A., et al., *RIG-I-dependent sensing of poly(dA:dT) through the induction of an RNA polymerase III-transcribed RNA intermediate.* Nature immunology, 2009. **10**(10): p. 1065-72.

62. Chiu, Y.H., J.B. Macmillan, and Z.J. Chen, *RNA polymerase III detects cytosolic DNA and induces type I interferons through the RIG-I pathway.* Cell, 2009. **138**(3): p. 576-91.

63. Monroe, K.M., S.M. McWhirter, and R.E. Vance, *Identification of host cytosolic sensors and bacterial factors regulating the type I interferon response to Legionella pneumophila.* PLoS Pathog, 2009. **5**(11): p. e1000665.

64. Melchjorsen, J., et al., *Early innate recognition of herpes simplex virus in human primary macrophages is mediated via the MDA5/MAVS-dependent and MDA5/MAVS/RNA polymerase III-independent pathways.* J Virol, 2010. **84**(21): p. 11350-8.

65. Unterholzner, L., et al., *IFI16 is an innate immune sensor for intracellular DNA.* Nature immunology, 2010. **11**(11): p. 997-1004.

66. Orzalli, M.H., N.A. DeLuca, and D.M. Knipe, *Nuclear IFI16 induction of IRF-3 signaling during herpesviral infection and degradation of IFI16 by the viral ICP0 protein.* Proc Natl Acad Sci U S A, 2012. **109**(44): p. E3008-17.

67. Jin, T., et al., *Structures of the HIN domain:DNA complexes reveal ligand binding and activation mechanisms of the AIM2 inflammasome and IFI16 receptor.* Immunity, 2012. **36**(4): p. 561-71.

68. Kim, T., et al., *Aspartate-glutamate-alanine-histidine box motif (DEAH)/RNA helicase A helicases sense microbial DNA in human plasmacytoid dendritic cells.* Proc Natl Acad Sci U S A, 2010. **107**(34): p. 15181-6.

69. Krieg, A.M., *CpG motifs in bacterial DNA and their immune effects.* Annu Rev Immunol, 2002. **20**: p. 709-60.

70. Hokeness-Antonelli, K.L., et al., *IFN-alphabeta-mediated inflammatory responses and antiviral defense in liver is TLR9-independent but MyD88-dependent during murine cytomegalovirus infection.* J Immunol, 2007. **179**(9): p. 6176-83.

71. Zhang, Z., et al., *DDX1, DDX21, and DHX36 helicases form a complex with the adaptor molecule TRIF to sense dsRNA in dendritic cells.* Immunity, 2011. **34**(6): p. 866-78.

72. Zhang, Z., et al., *DHX9 pairs with IPS-1 to sense double-stranded RNA in myeloid dendritic cells.* J Immunol, 2011. **187**(9): p. 4501-8.

73. Zhang, Z., et al., *The helicase DDX41 senses intracellular DNA mediated by the adaptor STING in dendritic cells.* Nature immunology, 2011. **12**(10): p. 959-65.

74. Wilson, S.A., et al., *TRIP: a novel double stranded RNA binding protein which interacts with the leucine rich repeat of flightless I.* Nucleic Acids Res, 1998. **26**(15): p. 3460-7.

75. Yang, P., et al., *The cytosolic nucleic acid sensor LRRFIP1 mediates the production of type I interferon via a beta-catenin-dependent pathway.* Nat Immunol, 2010. **11**(6): p. 487-94.

76. Mosimann, C., G. Hausmann, and K. Basler, *Beta-catenin hits chromatin: regulation of Wnt target gene activation.* Nat Rev Mol Cell Biol, 2009. **10**(4): p. 276-86.

77. Zhang, X., et al., *Cutting edge: Ku70 is a novel cytosolic DNA sensor that induces type III rather than type I IFN.* J Immunol, 2011. **186**(8): p. 4541-5.

78. Cavlar, T., A. Ablasser, and V. Hornung, *Induction of type I IFNs by intracellular DNA-sensing pathways.* Immunol Cell Biol, 2012. **90**(5): p. 474-82.

79. Burdette, D.L., et al., *STING is a direct innate immune sensor of cyclic di-GMP.* Nature, 2011. **478**(7370): p. 515-8.

80. Woodward, J.J., A.T. Iavarone, and D.A. Portnoy, *c-di-AMP secreted by intracellular Listeria monocytogenes activates a host type I interferon response.* Science, 2010. **328**(5986): p. 1703-5.

81. Sun, L., et al., *Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway.* Science, 2013. **339**(6121): p. 786-91.

82. Wu, J., et al., *Cyclic GMP-AMP is an endogenous second messenger in innate immune signaling by cytosolic DNA.* Science, 2013. **339**(6121): p. 826-30.

83. Yin, Q., et al., *Cyclic di-GMP sensing via the innate immune signaling protein STING.* Mol Cell, 2012. **46**(6): p. 735-45.

84. Huang, Y.H., et al., *The structural basis for the sensing and binding of cyclic di-GMP by STING.* Nat Struct Mol Biol, 2012. **19**(7): p. 728-30.

85. Ouyang, S., et al., *Structural analysis of the STING adaptor protein reveals a hydrophobic dimer interface and mode of cyclic di-GMP binding.* Immunity, 2012. **36**(6): p. 1073-86.

86. Parvatiyar, K., et al., *The helicase DDX41 recognizes the bacterial secondary messengers cyclic di-GMP and cyclic di-AMP to activate a type I interferon immune response.* Nat Immunol, 2012. **13**(12): p. 1155-61.

87. Schoggins, J.W., et al., *A diverse range of gene products are effectors of the type I interferon antiviral response.* Nature, 2011. **472**(7344): p. 481-5.

88. Ablasser, A. and V. Hornung, *DNA sensing unchained.* Cell Res, 2013.

89. Napirei, M., et al., *Features of systemic lupus erythematosus in Dnase1-deficient mice.* Nat Genet, 2000. **25**(2): p. 177-81.

90. Martinez Valle, F., et al., *DNase 1 and systemic lupus erythematosus.* Autoimmun Rev, 2008. **7**(5): p. 359-63.

91. Chitrabamrung, S., R.L. Rubin, and E.M. Tan, *Serum deoxyribonuclease I and clinical activity in systemic lupus erythematosus.* Rheumatol Int, 1981. **1**(2): p. 55-60.

92. Yasutomo, K., et al., *Mutation of DNASE1 in people with systemic lupus erythematosus.* Nat Genet, 2001. **28**(4): p. 313-4.

93. Mukae, N., et al., *Activation of the innate immunity in Drosophila by endogenous chromosomal DNA that escaped apoptotic degradation.* Genes Dev, 2002. **16**(20): p. 2662-71.

94. Kawane, K., et al., *Requirement of DNase II for definitive erythropoiesis in the mouse fetal liver.* Science, 2001. **292**(5521): p. 1546-9.

95. Yoshida, H., et al., *Lethal anemia caused by interferon-beta produced in mouse embryos carrying undigested DNA.* Nature immunology, 2005. **6**(1): p. 49-56.

96. Okabe, Y., et al., *Toll-like receptor-independent gene induction program activated by mammalian DNA escaped from apoptotic DNA degradation.* J Exp Med, 2005. **202**(10): p. 1333-9.

97. Kawane, K., et al., *Chronic polyarthritis caused by mammalian DNA that escapes from degradation in macrophages.* Nature, 2006. **443**(7114): p. 998-1002.

98. Ahn, J., et al., *STING manifests self DNA-dependent inflammatory disease.* Proc Natl Acad Sci U S A, 2012. **109**(47): p. 19386-91.

99. Lee-Kirsch, M.A., et al., *Mutations in the gene encoding the 3'-5' DNA exonuclease TREX1 are associated with systemic lupus erythematosus.* Nature genetics, 2007. **39**(9): p. 1065-7.

100. Crow, Y.J., et al., *Mutations in the gene encoding the 3'-5' DNA exonuclease TREX1 cause Aicardi-Goutieres syndrome at the AGS1 locus.* Nature genetics, 2006. **38**(8): p. 917-20.

101. Yang, Y.G., T. Lindahl, and D.E. Barnes, *Trex1 exonuclease degrades ssDNA to prevent chronic checkpoint activation and autoimmune disease.* Cell, 2007. **131**(5): p. 873-86.

102. Lehtinen, D.A., et al., *The TREX1 double-stranded DNA degradation activity is defective in dominant mutations associated with autoimmune disease.* J Biol Chem, 2008. **283**(46): p. 31649-56.

103. O'Driscoll, M., *TREX1 DNA exonuclease deficiency, accumulation of single stranded DNA and complex human genetic disorders.* DNA Repair (Amst), 2008. **7**(6): p. 997-1003.

104. Stetson, D.B., et al., *Trex1 prevents cell-intrinsic initiation of autoimmunity.* Cell, 2008. **134**(4): p. 587-98.

105. Goutieres, F., *Aicardi-Goutieres syndrome.* Brain Dev, 2005. **27**(3): p. 201-6.

106. Yan, N., et al., *The cytosolic exonuclease TREX1 inhibits the innate immune response to human immunodeficiency virus type 1.* Nature immunology, 2010. **11**(11): p. 1005-13.

107. Gallucci, S., M. Lolkema, and P. Matzinger, *Natural adjuvants: endogenous activators of dendritic cells.* Nat Med, 1999. **5**(11): p. 1249-55.

108. Andrade, F., L. Casciola-Rosen, and A. Rosen, *Apoptosis in systemic lupus erythematosus. Clinical implications.* Rheum Dis Clin North Am, 2000. **26**(2): p. 215-27, v.

109. Botto, M., et al., *Homozygous C1q deficiency causes glomerulonephritis associated with multiple apoptotic bodies.* Nat Genet, 1998. **19**(1): p. 56-9.

110. Gaipl, U.S., et al., *Clearance deficiency and systemic lupus erythematosus (SLE).* J Autoimmun, 2007. **28**(2-3): p. 114-21.

111. Munoz, L.E., et al., *SLE--a disease of clearance deficiency?* Rheumatology (Oxford), 2005. **44**(9): p. 1101-7.

112. Janssen, E., et al., *Efficient T cell activation via a Toll-Interleukin 1 Receptor-independent pathway.* Immunity, 2006. **24**(6): p. 787-99.

113. Barrat, F.J., et al., *Nucleic acids of mammalian origin can act as endogenous ligands for Toll-like receptors and may promote systemic lupus erythematosus.* J Exp Med, 2005. **202**(8): p. 1131-9.

114. Decker, P., et al., *Nucleosome, the main autoantigen in systemic lupus erythematosus, induces direct dendritic cell activation via a MyD88-independent pathway: consequences on inflammation.* J Immunol, 2005. **174**(6): p. 3326-34.

115. Lu, Q. and G. Lemke, *Homeostatic regulation of the immune system by receptor tyrosine kinases of the Tyro 3 family.* Science, 2001. **293**(5528): p. 306-11.

116. Scott, R.S., et al., *Phagocytosis and clearance of apoptotic cells is mediated by MER.* Nature, 2001. **411**(6834): p. 207-11.

117. Suzuki, K., et al., *Activation of target-tissue immune-recognition molecules by double-stranded polynucleotides.* Proceedings of the National Academy of Sciences of the United States of America, 1999. **96**(5): p. 2285-90.

118. Pichlmair, A., et al., *RIG-I-mediated antiviral responses to single-stranded RNA bearing 5'-phosphates.* Science, 2006. **314**(5801): p. 997-1001.

119. Hornung, V., et al., *Sequence-specific potent induction of IFN-alpha by short interfering RNA in plasmacytoid dendritic cells through TLR7.* Nat Med, 2005. **11**(3): p. 263-70.

120. Diebold, S.S., et al., *Nucleic acid agonists for Toll-like receptor 7 are defined by the presence of uridine ribonucleotides.* Eur J Immunol, 2006. **36**(12): p. 3256-67.

121. Okabe, Y., T. Sano, and S. Nagata, *Regulation of the innate immune response by threonine-phosphatase of Eyes absent.* Nature, 2009. **460**(7254): p. 520-4.

122. Carette, J.E., et al., *Haploid genetic screens in human cells identify host factors used by pathogens.* Science, 2009. **326**(5957): p. 1231-5.

123. Nociari, M., et al., *Sensing infection by adenovirus: Toll-like receptor-independent viral DNA recognition signals activation of the interferon regulatory factor 3 master regulator.* J Virol, 2007. **81**(8): p. 4145-57.

124. Cheng, G., et al., *Double-stranded DNA and double-stranded RNA induce a common antiviral signaling pathway in human cells.* Proc Natl Acad Sci U S A, 2007. **104**(21): p. 9035-40.

125. Ablasser, A., et al., *RIG-I-dependent sensing of poly(dA:dT) through the induction of an RNA polymerase III-transcribed RNA intermediate.* Nat Immunol, 2009. **10**(10): p. 1065-72.

126. Amit, I., et al., *Unbiased reconstruction of a mammalian transcriptional network mediating pathogen responses.* Science, 2009. **326**(5950): p. 257-63.

127. Lahita, R.G., *Systemic lupus erythematosus.* 5th ed2011, Amsterdam ; New York: Elsevier Academic Press. xix, 1134 p.

128. Baechler, E.C., et al., *Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus.* Proc Natl Acad Sci U S A, 2003. **100**(5): p. 2610-5.

129. Fitzgerald, K.A., et al., *IKKepsilon and TBK1 are essential components of the IRF3 signaling pathway.* Nat Immunol, 2003. **4**(5): p. 491-6.

130. Sigal, A., et al., *Generation of a fluorescently labeled endogenous protein library in living human cells.* Nat Protoc, 2007. **2**(6): p. 1515-27.

131. Van Peer, G., P. Mestdagh, and J. Vandesompele, *Accurate RT-qPCR gene expression analysis on cell culture lysates.* Sci Rep, 2012. **2**: p. 222.

132. Shames, D.S., et al., *A genome-wide screen for promoter methylation in lung cancer identifies novel methylation markers for multiple malignancies.* PLoS Med, 2006. **3**(12): p. e486.

133. Choe, J., H.H. Guo, and G. van den Engh, *A dual-fluorescence reporter system for high-throughput clone characterization and selection by cell sorting.* Nucleic Acids Res, 2005. **33**(5): p. e49.

134. Shapira, S.D., et al., *A physical and regulatory map of host-influenza interactions reveals pathways in H1N1 infection.* Cell, 2009. **139**(7): p. 1255-67.

135. Moffat, J., et al., *A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen.* Cell, 2006. **124**(6): p. 1283-98.

136. Blakely, K., T. Ketela, and J. Moffat, *Pooled lentiviral shRNA screening for functional genomics in mammalian cells.* Methods Mol Biol, 2011. **781**: p. 161-82.

137. Luo, H., et al., *Inhibition of cell growth and VEGF expression in ovarian cancer cells by flavonoids.* Nutr Cancer, 2008. **60**(6): p. 800-9.

138. Zhang, J.H., T.D. Chung, and K.R. Oldenburg, *A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays.* J Biomol Screen, 1999. **4**(2): p. 67-73.

139. Tsuchida, T., et al., *The ubiquitin ligase TRIM56 regulates innate immune responses to intracellular double-stranded DNA.* Immunity, 2010. **33**(5): p. 765-76.

140. Lee, M.N., et al., *Identification of regulators of the innate immune response to cytosolic DNA and retroviral infection by an integrative approach.* Nat Immunol, 2013. **14**(2): p. 179-85.

141. Ong, S.E., et al., *Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics.* Molecular & cellular proteomics : MCP, 2002. **1**(5): p. 376-86.

142. Ong, S.E., G. Mittler, and M. Mann, *Identifying and quantifying in vivo methylation sites by heavy methyl SILAC.* Nature methods, 2004. **1**(2): p. 119-26.

143. Yanai, H., et al., *HMGB proteins function as universal sentinels for nucleic-acid-mediated innate immune responses.* Nature, 2009. **462**(7269): p. 99-103.

144. Yan, N., et al., *The SET complex acts as a barrier to autointegration of HIV-1.* PLoS pathogens, 2009. **5**(3): p. e1000327.

145. Rice, G.I., et al., *Mutations involved in Aicardi-Goutieres syndrome implicate SAMHD1 as regulator of the innate immune response.* Nature genetics, 2009. **41**(7): p. 829-32.

146. Laguette, N. and M. Benkirane, *How SAMHD1 changes our view of viral restriction.* Trends Immunol, 2012. **33**(1): p. 26-33.

147. Oshiumi, H., et al., *DEAD/H BOX 3 (DDX3) helicase binds the RIG-I adaptor IPS-1 to up-regulate IFN-beta-inducing potential.* Eur J Immunol, 2010. **40**(4): p. 940-8.

148. Soulat, D., et al., *The DEAD-box helicase DDX3X is a critical component of the TANK-binding kinase 1-dependent innate immune response.* EMBO J, 2008. **27**(15): p. 2135-46.

149. Ronald, P.C. and B. Beutler, *Plant and animal sensors of conserved microbial signatures.* Science, 2010. **330**(6007): p. 1061-4.

150. Kawai, T. and S. Akira, *The roles of TLRs, RLRs and NLRs in pathogen recognition.* Int Immunol, 2009. **21**(4): p. 317-37.

151. Kato, H., et al., *Differential roles of MDA5 and RIG-I helicases in the recognition of RNA viruses.* Nature, 2006. **441**(7089): p. 101-5.

152. Yoneyama, M., et al., *Shared and unique functions of the DExD/H-box helicases RIG-I, MDA5, and LGP2 in antiviral innate immunity.* J Immunol, 2005. **175**(5): p. 2851-8.

153. Schroder, M., M. Baran, and A.G. Bowie, *Viral targeting of DEAD box protein 3 reveals its role in TBK1/IKKepsilon-mediated IRF activation.* EMBO J, 2008. **27**(15): p. 2147-57.

154. DeFilippis, V.R., et al., *Human cytomegalovirus induces the interferon response via the DNA sensor ZBP1.* Journal of virology, 2010. **84**(1): p. 585-98.

155. Mi, H., et al., *The PANTHER database of protein families, subfamilies, functions and pathways.* Nucleic Acids Res, 2005. **33**(Database issue): p. D284-8.

156. Thomas, P.D., et al., *PANTHER: a library of protein families and subfamilies indexed by function.* Genome Res, 2003. **13**(9): p. 2129-41.

157. Thomas, P.D., et al., *PANTHER: a browsable database of gene products organized by biological function, using curated protein family and subfamily classification.* Nucleic Acids Res, 2003. **31**(1): p. 334-41.

158. Fink, J.L., et al., *LOCATE: a mouse protein subcellular localization database.* Nucleic Acids Res, 2006. **34**(Database issue): p. D213-7.

159. Carninci, P., et al., *The transcriptional landscape of the mammalian genome.* Science, 2005. **309**(5740): p. 1559-63.

160. Fulton, D.L., et al., *TFCat: the curated catalog of mouse and human transcription factors.* Genome Biol, 2009. **10**(3): p. R29.

161. Yu, C.S., C.J. Lin, and J.K. Hwang, *Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions.* Protein Sci, 2004. **13**(5): p. 1402-6.

162. Yu, C.S., et al., *Prediction of protein subcellular localization.* Proteins, 2006. **64**(3): p. 643-51.

163. Horton, P., et al., *WoLF PSORT: protein localization predictor.* Nucleic Acids Res, 2007. **35**(Web Server issue): p. W585-7.

164. McWhirter, S.M., et al., *IFN-regulatory factor 3-dependent gene expression is defective in Tbk1-deficient mouse embryonic fibroblasts.* Proc Natl Acad Sci U S A, 2004. **101**(1): p. 233-8.

165. Trompouki, E., et al., *CYLD is a deubiquitinating enzyme that negatively regulates NF-kappaB activation by TNFR family members.* Nature, 2003. **424**(6950): p. 793-6.

166. Friedman, C.S., et al., *The tumour suppressor CYLD is a negative regulator of RIG-I-mediated antiviral response.* EMBO Rep, 2008. **9**(9): p. 930-6.

167. Wertz, I.E., et al., *De-ubiquitination and ubiquitin ligase domains of A20 downregulate NF-kappaB signalling.* Nature, 2004. **430**(7000): p. 694-9.

168. Kayagaki, N., et al., *DUBA: a deubiquitinase that regulates type I interferon production.* Science, 2007. **318**(5856): p. 1628-32.

169. Lee, E.G., et al., *Failure to regulate TNF-induced NF-kappaB and cell death responses in A20-deficient mice.* Science, 2000. **289**(5488): p. 2350-4.

170. Rice, G., et al., *Heterozygous mutations in TREX1 cause familial chilblain lupus and dominant Aicardi-Goutieres syndrome.* American journal of human genetics, 2007. **80**(4): p. 811-5.

171. Gall, A., et al., *Autoimmunity initiates in nonhematopoietic cells and progresses via lymphocytes in an interferon-dependent autoimmune disease.* Immunity, 2012. **36**(1): p. 120-31.

172. Nijman, S.M., et al., *A genomic and functional inventory of deubiquitinating enzymes.* Cell, 2005. **123**(5): p. 773-86.

173. Hindorff, L.A., et al., *Potential etiologic and functional implications of genome-wide association loci for human diseases and traits.* Proc Natl Acad Sci U S A, 2009. **106**(23): p. 9362-7.

174. Wu, C., et al., *BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources.* Genome Biol, 2009. **10**(11): p. R130.

175. Letunic, I., T. Doerks, and P. Bork, *SMART 7: recent updates to the protein domain annotation resource.* Nucleic Acids Res, 2012. **40**(Database issue): p. D302-5.

176. Jackson, A.L., et al., *Position-specific chemical modification of siRNAs reduces "off-target" transcript silencing.* RNA, 2006. **12**(7): p. 1197-205.

177. Paytubi, S., et al., *ABC50 promotes translation initiation in mammalian cells.* The Journal of biological chemistry, 2009. **284**(36): p. 24061-73.

178. Wilcox, S.M., *The Function Of ABCF1 in immunity and mouse developent.*, in *The Faculty of Graduate Studies*2011, The University of British Columbia: Vancouver, BC. p. 246.

179. Rozenblatt-Rosen, O., et al., *Interpreting cancer genomes using systematic host network perturbations by tumour virus proteins.* Nature, 2012.

180. Samaniego, L.A., L. Neiderhiser, and N.A. DeLuca, *Persistence and expression of the herpes simplex virus genome in the absence of immediate-early proteins.* Journal of virology, 1998. **72**(4): p. 3307-20.

181. Kim, K., et al., *Recombination signatures distinguish embryonic stem cells derived by parthenogenesis and somatic cell nuclear transfer.* Cell Stem Cell, 2007. **1**(3): p. 346-52.

182. Guo, L., E. Allen, and W.A. Miller, *Structure and function of a cap-independent translation element that functions in either the 3' or the 5' untranslated region.* RNA, 2000. **6**(12): p. 1808-20.

183. Fensterl, V. and G.C. Sen, *The ISG56/IFIT1 gene family.* J Interferon Cytokine Res, 2011. **31**(1): p. 71-8.

184. Pichlmair, A., et al., *IFIT1 is an antiviral protein that recognizes 5'-triphosphate RNA.* Nat Immunol, 2011. **12**(7): p. 624-30.

185. Bordon, Y., *Innate immunity: an inducible RNA sensor? IFITs the bill.* Nat Rev Immunol, 2011. **11**(7): p. 440.

186. Terenzi, F., et al., *Distinct induction patterns and functions of two closely related interferon-inducible human genes, ISG54 and ISG56.* J Biol Chem, 2006. **281**(45): p. 34064-71.

187. Hui, D.J., et al., *Mouse p56 blocks a distinct function of eukaryotic initiation factor 3 in translation initiation.* J Biol Chem, 2005. **280**(5): p. 3433-40.

188. Hui, D.J., et al., *Viral stress-inducible protein p56 inhibits translation by blocking the interaction of eIF3 with the ternary complex eIF2.GTP.Met-tRNAi.* J Biol Chem, 2003. **278**(41): p. 39477-82.

189. Fensterl, V., et al., *Interferon-induced Ifit2/ISG54 protects mice from lethal VSV neuropathogenesis.* PLoS Pathog, 2012. **8**(5): p. e1002712.

190. McDermott, J.E., et al., *Identification and validation of Ifit1 as an important innate immune bottleneck.* PLoS One, 2012. **7**(6): p. e36465.

191. Cermak, T., et al., *Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting.* Nucleic Acids Res, 2011. **39**(12): p. e82.

192. Zhang, F., et al., *Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription.* Nat Biotechnol, 2011. **29**(2): p. 149-53.

193. Saito, H., et al., *RTP family members induce functional expression of mammalian odorant receptors.* Cell, 2004. **119**(5): p. 679-91.

194. Argasinska, J., et al., *Loss of REEP4 causes paralysis of the Xenopus embryo.* Int J Dev Biol, 2009. **53**(1): p. 37-43.

195. Beetz, C., et al., *Exome sequencing identifies a REEP1 mutation involved in distal hereditary motor neuropathy type V.* Am J Hum Genet, 2012. **91**(1): p. 139-45.

196. Zuchner, S., et al., *A new locus for dominant hereditary spastic paraplegia maps to chromosome 2p12.* Neurogenetics, 2006. **7**(2): p. 127-9.

197. Chamilos, G., et al., *Cytosolic sensing of extracellular self-DNA transported into monocytes by the antimicrobial peptide LL37.* Blood, 2012. **120**(18): p. 3699-707.

198. Sandgren, S., et al., *The human antimicrobial peptide LL-37 transfers extracellular DNA plasmid to the nuclear compartment of mammalian cells via lipid rafts and proteoglycan-dependent endocytosis.* J Biol Chem, 2004. **279**(17): p. 17951-6.

199. Ilegems, E., et al., *REEP2 enhances sweet receptor function by recruitment to lipid rafts.* J Neurosci, 2010. **30**(41): p. 13774-83.

200. Mercer, J., M. Schelhaas, and A. Helenius, *Virus entry by endocytosis.* Annu Rev Biochem, 2010. **79**: p. 803-33.

201. Drake, D.R., 3rd, et al., *Polyomavirus-infected dendritic cells induce antiviral CD8(+) T lymphocytes.* J Virol, 2000. **74**(9): p. 4093-101.

202. Bouwmeester, T., et al., *A physical and functional map of the human TNF-alpha/NF-kappa B signal transduction pathway.* Nature cell biology, 2004. **6**(2): p. 97-105.

203. Myers, M.P., et al., *TYK2 and JAK2 are substrates of protein-tyrosine phosphatase 1B.* The Journal of biological chemistry, 2001. **276**(51): p. 47771-4.

204. Pan, H., et al., *Ipr1 gene mediates innate immunity to tuberculosis.* Nature, 2005. **434**(7034): p. 767-72.

205. Watashi, K., et al., *Modulation of retinoid signaling by a cytoplasmic viral protein via sequestration of Sp110b, a potent transcriptional corepressor of retinoic acid receptor, from the nucleus.* Mol Cell Biol, 2003. **23**(21): p. 7498-509.

206. Nicewonger, J., et al., *Epstein-Barr virus (EBV) SM protein induces and recruits cellular Sp110b to stabilize mRNAs and enhance EBV lytic gene expression.* J Virol, 2004. **78**(17): p. 9412-22.

207. Cai, L., et al., *MYBBP1A: a new Ipr1's binding protein in mice.* Mol Biol Rep, 2010. **37**(8): p. 3863-8.

208. Wenzel, M., et al., *Cytosolic DNA triggers mitochondrial apoptosis via DNA damage signaling proteins independently of AIM2 and RNA polymerase III.* J Immunol, 2012. **188**(1): p. 394-403.

209. Johnson, J.M., et al., *Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays.* Science, 2003. **302**(5653): p. 2141-4.

210. Elmore, S., *Apoptosis: a review of programmed cell death.* Toxicol Pathol, 2007. **35**(4): p. 495-516.

211. Tao, Y., et al., *Lsh, chromatin remodeling family member, modulates genome-wide cytosine methylation patterns at nonrepeat sequences.* Proc Natl Acad Sci U S A, 2011. **108**(14): p. 5626-31.

212. Sun, L.Q., et al., *Growth retardation and premature aging phenotypes in mice with disruption of the SNF2-like gene, PASG.* Genes Dev, 2004. **18**(9): p. 1035-46.

213. Geiman, T.M. and K. Muegge, *Lsh, an SNF2/helicase family member, is required for proliferation of mature T lymphocytes.* Proc Natl Acad Sci U S A, 2000. **97**(9): p. 4772-7.

214. Ortutay, C., M. Siermala, and M. Vihinen, *ImmTree: database of evolutionary relationships of genes and proteins in the human immune system.* Immunome Res, 2007. **3**: p. 4.

215. Carter, C.C., V.Y. Gorbacheva, and D.J. Vestal, *Inhibition of VSV and EMCV replication by the interferon-induced GTPase, mGBP-2: differential requirement for wild-type GTP binding domain.* Arch Virol, 2005. **150**(6): p. 1213-20.

216. Zhao, Y.O., et al., *Toxoplasma gondii and the Immunity-Related GTPase (IRG) resistance system in mice: a review.* Mem Inst Oswaldo Cruz, 2009. **104**(2): p. 234-40.

217. Yan, N., et al., *HIV DNA is heavily uracilated, which protects it from autointegration.* Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(22): p. 9244-9.

218. Mostafavi, S., et al., *GeneMANIA: a real-time multiple association network integration algorithm for predicting gene function.* Genome Biol, 2008. **9 Suppl 1**: p. S4.

219. Klaman, L.D., et al., *Increased energy expenditure, decreased adiposity, and tissue-specific insulin sensitivity in protein-tyrosine phosphatase 1B-deficient mice.* Molecular and cellular biology, 2000. **20**(15): p. 5479-89.

220. Zhang, J., et al., *Endonuclease G is required for early embryogenesis and normal apoptosis in mice.* Proc Natl Acad Sci U S A, 2003. **100**(26): p. 15782-7.

221. Liebscher, I., et al., *Altered immune response in mice deficient for the G protein-coupled receptor GPR34.* J Biol Chem, 2011. **286**(3): p. 2101-10.

222. Shima, N., R.J. Munroe, and J.C. Schimenti, *The mouse genomic instability mutation chaos1 is an allele of Polq that exhibits genetic interaction with Atm.* Mol Cell Biol, 2004. **24**(23): p. 10381-9.

223. Dower, N.A., et al., *RasGRP is essential for mouse thymocyte differentiation and TCR signaling.* Nat Immunol, 2000. **1**(4): p. 317-21.