Nucleic Acid Detection in Anti-Viral Responses and Autoimmunity

Rebecca Lynn Brunette

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2012

Reading Committee:

Daniel B. Stetson, Chair

Michael J. Gale

Jessica A. Hamerman

Program Authorized to Offer Degree:

Immunology

University of Washington

Abstract

Nucleic Acid Detection in Anti-Viral Responses and Autoimmunity

Rebecca Lynn Brunette

Chair of Supervisory Committee:

Assistant Professor Daniel B. Stetson

Department of Immunology

Viruses infect all living organisms. Therefore, hosts evolved extensive means to detect and destroy viral pathogens. Pattern recognition receptors that detect viral RNA or DNA genome are important components of the innate immune response against viruses. In vertebrates, viral nucleic acid is detected by two mechanisms that both result in the production of antiviral type I IFN. The first mechanism involves toll-like receptor detection of non-cell autonomous viral nucleic acid and allows uninfected, specialized innate immune cells to initiate an adaptive immune response to eliminate infection. Second, cell intrinsic receptors allow all infected cells to detect viral infection and induce a type I IFN response. The interferon stimulatory DNA (ISD) pathway specifically detects of viral DNA genomes. Chronic activation of the ISD pathway is the cause of lethal autoimmune disease in Trex1 deficient mice and humans thereby coupling innate immune detection to initiation of autoimmunity. The ISD pathway induces type I IFN by STING-dependent activation of TBK1 and IRF3. However, little is known about the DNA sensors proteins of the ISD pathway. Here we use the novel model of Trex1 autoimmunity to identify Trex1 interacting proteins that may be involved in ISD sensing. We discovered that

Trex1 is localized to the inner nuclear membrane in addition to the endoplasmic reticulum and interacts with nuclear partners Brd7, Btf3l4, Chaf1a, Pias1, Rfc2 and Sae2. This data implies that negative regulation of the ISD pathway may occur in the nucleus and not the cytoplasm as previously thought. In addition, we identified new members and characterized the 13 members of the AIM2-like receptor (ALR) family of ISD sensors. Here we show the ALR family of nuclear proteins is highly divergent across mammals reflecting each species' unique struggle with specific viral pathogens. We also discovered the ALR proteins fit into two functional categories based on their ability to interact with STING and induce type I IFN or interact with ASC to activate the IL-1β inflammasome. Together our data imply nuclear detection of DNA occurs utilizing highly species-specific sensors and provides a functional framework for further study of the ISD sensors.

Table of Contents

List of Figures

Section 2

Section 3

List of Tables

Section 3

Acknowledgements

The author wishes to acknowledge all of the scientific and personal support provided during her graduate studies. Special thanks to Dan Stetson for being a fantastic mentor and all his support during my graduate training. Also, thank you to all past and present members of the Stetson Lab including Tina Gall, Hannah Volkman, Debbie Whitley, Sterling Eckard and Laura Lau for ongoing scientific support, technical advice and making the lab a fun place to work. Thanks to C.J. Cambier, Shivani Srivastava and Jeff Duggan for scientific contributions to Sections 2 and 3 during their rotations. More thanks to Harmit Malik and Janet Young for their scientific advice and contributions over the course of our collaboration and manuscript publication. All of the members of my thesis committee not previously mentioned, Jessica Hamerman, Michael Gale and Keith Elkon, for contributions to my scientific training. Lastly, a warm thank you to all my family and friends especially my parents Steve and Karen Brunette and fiancé Dmitry Gorn. Completion of this work would not have been possible without your support.

Section 1: Introduction

Viruses are obligate intracellular parasites that replicate and transmit their genomes from infected to non-infected host cells. Because virions are simple particles incapable of metabolism or self-motility, viruses exploit the host cell machinery for all stages of infection including entry, replication, particle assembly and export. Viruses infect every organism from single celled bacteria, Archaea and Eukaryotes to complex, multicellular Eukaryotes such as plants and animals. Viruses and phages (viruses that infect bacteria and Archaea) are the most abundant biological entities on earth [1]. Studies have shown that one milliliter of seawater contains around 10 million viral particles [2]. Structurally virus and phage particles are generally comprised of a nucleic acid genome consisting of either ribonucleic acid (RNA) or deoxyribonucleic acid (DNA) in either single (ss) or double stranded (ds) form enclosed by a protein capsid and in the case of enveloped viruses, surrounded by a host-derived lipid membrane. Additionally, some viral particles also contain accessory proteins such as reverse transcriptases, RNA polymerases, kinases and/or proteases that aid in establishing and propagating infection.

Host Defense Against Viral Infection

Because of their ubiquity and abundance, all organisms have developed defense mechanisms to prevent and control virus infection. Broadly, defense against pathogens across all kingdoms is based on the ability of organisms to discriminate self versus non-self molecules. This is achieved at a cellular level by specialized host proteins that detect foreign molecules such as lipids, proteins and nucleic acids that are either molecularly distinct or improperly localized from host molecules. In the case of viral infections, discrimination of viral nucleic acid (either RNA or DNA) from self nucleic acid is an important for recognition and elimination of the viral

pathogen. Upon recognition of viral pathogens, organisms or individual cells in multicellular organisms initiate a wide variety of antiviral responses that range from direct degradation or silencing of the viral genome to expression of signaling proteins (cytokines) that alert specialized cells within the organism to destroy virally infected cells. Because the mechanisms of antiviral defense are broad, specific examples of viral detection and defense in Prokaryotes and Eukaryotes will be discussed below.

Studies in aquatic environments have identified that the relative ratio of phages to unicellular bacteria and archaea is on the order of 5-10 phage particles per bacterial cell [3]. Additionally, phages are responsible for the death of 5-50% of bacteria produced each day [4, 5]. Despite the high mortality rate of Prokaryotes by phages, bacteria and archaea are capable of detecting and inhibiting phages. Many bacteria prevent phage entry and DNA injection by production of extracellular matrix or specific proteins that block phage receptors. For example, both *Azotobacter* and *Pseudomonas* species of bacteria produce extracellular alginate matrices that confer resistance to phage infection by blocking cell wall receptors [6-8]. Many bacteria also encode and express proteins that block phage entry receptors. *Staphylococcus aureus* produces a phage receptor masking protein called protein A that prevents phage absorption upon expression [9]. These systems are the first line of defense against phage infection and as such many phages have evolved ways to overcome inhibition. One example of this is phage F116, which specifically targets Pseudomonas species of bacteria. This phage produces an alginate lyase that degrades *Pseudomonas*' protective alginate in order to facilitate its entry [7].

Another mechanism of bacterial and archaeal resistance to phage infection is detection and targeted degradation of phage genome. Most bacteria and archaea achieve this by a specialized restriction-modification (R-M) system wherein phage DNA is detected due to its lack

of methylation and cleaved at specific sites [10]. There are four types R-M systems (type I-IV) that vary based on target substrate specificity, protein subunits involved and the requirement of specific co-factors. In general, R-M systems are composed of restriction endonuclease (REases) enzymes that recognize a specific DNA sequence for cleavage and another enzyme that that has methyltransferase activity to confer protection to the host by methylating adenosines and cytosines in the same sequence that the REase recognizes (reviewed in [11]). As with viral entry, the evolutionary pressure placed on phages to overcome R-M systems has led to phages that are completely resistant such as the T4 phage, which contains hydroxymethylcytosine instead of cytosine, rendering it completely resistant to R-M recognition and cleavage [12].

Recent studies have identified a novel phage defense system present in bacteria and archaea. Nearly 40% of bacteria and most (90%) archaea species have at least one clustered, regularly interspaced short palindromic repeats (CRISPR) locus containing palindromic repeats separated by spacers derived from phage genomic DNA and associated CRISPR associated genes (CAS) [13-17]. The spacers derived from phage DNA are added upon infection with predatory phages as demonstrated with *Streptococcus mutans* and *Streptococcus thermophilus* [18-21]. Upon re-infection with a particular phage, CRISPR RNAs are transcribed, processed by the CAS family of proteins and direct sequence specific silencing of phage (reviewed in [22-24]). In essence the CRISPR system confers bacteria and archaea the ability to acquire "memory" of past phage infections in order to restrict phage growth upon subsequent infection, an ability that was previously thought to be restricted to vertebrates.

RNA interference (RNAi) in non-vertebrate eukaryotes is similar to the CRISPR system in prokaryotes. RNA interference is a major component of antiviral defense in fungi, plants, nematode worms and insects [25-36]. Double stranded RNA from viruses or parasitic

retrotransposon RNA is recognized and processed by Dicer into 21-24 nucleotide small interfering RNAs (siRNAs) and loaded into the RNA-induced silencing complex (RISC) containing an Argonaute protein. The active Argonaute protein then mediates cleavage and silencing of viral RNA (reviewed in [37-39]). In contrast to the prokaryotic CRISPR system, viral sequence is not incorporated into the eukaryotic genome during RNAi. The RNAi system diversified in animals into a microRNA system that mediates endogenous gene silencing during development or differentiation. MicroRNA gene silencing works in a manner analogous to RNA interference except the RNAs that are processed and loaded into the Argonaute-containing RISC are derived from endogenous, non-coding mRNAs [40-43]. Despite the prevalence of RNAi and microRNA systems across all eukaryotes, there is no strong evidence for an antiviral RNAi pathway in vertebrates [44, 45].

The Vertebrate Immune System

Vertebrates evolved complex means to recognize and eliminate viruses that extend beyond antiviral RNA interference. Multicellularity introduced new challenges and necessitated the development of many different specialized cells that work in concert to eliminate viruses as well as other vertebrate pathogens such as bacteria and eukaryotic parasites. Broadly, the vertebrate immune system is derived from hematopoietic precursors and comprised of the innate and the adaptive arms that primarily differ based on the receptors used to recognize pathogens and their effector functions. The invariant pattern recognition receptors (PRRs) of the innate immune system recognize conserved pathogen-associated molecular patterns (PAMPs) while adaptive immune receptors diversify by random recombination of gene segments to recognize specific antigen sequences. In general, pattern recognition receptor activation of innate immune cells is required to activate and inform the cells of the adaptive immune response. An overview

of the vertebrate immune system with additional focus on the innate immune response to viral infection follows.

The Innate Immune System

The innate immune system is composed of a variety of hematopoietic cells with specialized effector functions such as macrophages, dendritic cells (DCs), neutrophils, Natural Killer (NK) cells, eosinophils, basophils and mast cells. The hallmark of the innate immune system is its ability to detect pathogen-derived molecules from self-molecules through PRRs. Pathogen recognition by these cells directly or indirectly activates the production of cytokines to initiate an inflammatory response. The inflammatory response then serves to activate specialized effector functions ranging from direct killing of infected cells, phagocytosis of extracellular microbes, activation of the adaptive immune system, and/or amplification of the inflammatory response. There are several different families of PRRs expressed by innate immune cells that recognize conserved PAMPs.

The 12 members of the Toll-like receptor (TLR) family are primarily expressed on the surface and in endosomal compartments of macrophages, DCs and some adaptive immune cells. The TLR family recognizes a wide range of PAMPs. TLR1, TLR2 and TLR6 recognize lipid molecules such as lipopeptides found in the cell wall of bacteria and zymosan, a lipid found on the surface of yeast. TLR3, TLR7, TLR8 and TLR9 all recognize foreign nucleic acids produced by viruses or intracellular bacteria such as double-stranded RNA (TLR3), single-stranded RNA (TLR7 and 8) and hypomethylated DNA respectively (TLR9) (discussed in greater detail below). Lastly, TLR5 recognizes the flagellin protein produced by flagellated bacteria, and TLR4 has a well-characterized role in recognizing bacterial lipopolysaccharide, but has been shown to recognize many bacterial and viral PAMPs (TLRs reviewed in [46]). Recognition of PAMPs by

TLRs on macrophages and DCs generally results in the production of inflammatory proteins including cytokines and chemokines such as tumor necrosis factor (TNF), interleukin-1β (IL-1β), and interferon-β (IFN-β). These cytokines then activate the expression of soluble complement components to aid in the phagocytosis and destruction of extracellular pathogens by macrophages and neutrophils and to recruit other immune cells to the site of infection. C-type lectin receptors (CLRs) are another class of PRRs expressed primarily on the surface of macrophages and DCs that are important for antifungal defense. For example, the CLR Dectin-1 binds the fungal PAMP β-glucan to induce antifungal effector functions such as IL-2 and IL-10 production and phagocytosis of the pathogen [47, 48]. However, the function of the other 16 CLRs is still largely unknown.

Another group of pattern recognition receptors utilized by the innate immune system to detect microbial PAMPs are the intracellular NOD-like receptors (NLRs). The NOD-1 and NOD-2 receptors are primarily responsible for recognizing intracellular bacterial peptidoglycan and activating inflammatory cytokines production and neutrophil recruitment and degranulation in an NF-κB-dependent manner [49]. In addition, NLRP3, NLRC4, and NLRP1 activate caspase-1-dependent inflammasome processing and secretion of pro-IL-1β and IL-18 into their mature biologically active forms. Except for NLRC4 which has a well described role in recognizing bacterial flagellin, the ligands of NLRP1 and NLRP3 are not well understood and range from host molecules and processes such as ATP and ion influx/efflux to bacterial toxins and muramyl dipeptides. In addition to inflammatory cytokine production, activation of the caspase-1 inflammasome can also lead to an inflammatory altruistic form of cell death termed pyropoptosis (NLR inflammasome reviewed in [50]). Lastly, there are a group of intracellular PRRs that recognize viral RNA (RIG-I and MDA5) and DNA (DAI, IFI16, DDX41) and activate an

antiviral type I IFN response. These PRRs are not exclusively found in specialized innate immune cells, but instead are expressed in most cell types. The ability of any cell to detect viral infection and participate in the antiviral response is an important feature of the immune system in multicellular organisms (discussed in detail below).

In addition to recognition of pathogens by pattern recognition receptors, the innate immune system activates and participates in antimicrobial effector functions. The cytokines and effector functions activated by the innate immune system during an infection directly depend on the pattern recognition receptors triggered. Direct killing of extracellular microbes or virally-infected cells is a common innate immune effector response to bacterial and viral infection. For example, neutrophils are recruited to the site of infection based on cytokines and chemokines produced by activation of PRR such as TLR4 that indicate extracellular infection. These specialized cells mediate killing of microbes and amplification of inflammation by releasing their secretory granules containing proteases, antimicrobial peptides and reactive oxygen species [51]. Additionally, professional phagocytes such as macrophages and dendritic cells are responsible for pathogen killing and clearance. Once activated by pro-inflammatory cytokines, these cells phagocytosis microbes into endosomal compartments where acidification and microbe death occurs [52]. Although a subset of lymphocyte, natural killer (NK) cells play an important role in the innate immune response to viruses. NK cells recognize virally-infected cells by the absence of MHC class I antigen receptors and expression of activating ligands. When NK cells detect both lack of MHC class I and activating ligand on the same cell, they release preformed vesicles containing granzymes and perforin that kill the virally infected cell [53]. In the case of multicellular parasitic infections, innate effector cells such as mast cells, basophils and eosinophils contribute to pathogen clearance by secreting cytokines and other protein factors that

act on the tissue microenvironment to induce expulsion of the parasite. These factors act to induce tissue remodeling, increase vascular permeability, and secretions in the case of the mucosa and amplify the adaptive immune response to result in the expulsion rather than death of the parasite. However, unlike the well-characterized pattern recognition receptors for microbial products on macrophages and dendritic cells, it is not understood what activates mast cells, basophils and eosinophils to secrete soluble factors [54]. Lastly, antigen presentation to CD4 and CD8 T cells is a critical innate effector function that links innate immune activation by PRR to induction of adaptive immune responses. PRR-activated phagocytes proteolytically process phagocytosed pathogen proteins into peptides that they present on major histocompatibility molecules (MHC) I and II, and upregulate the expression of co-stimulatory molecules such as CD80, CD86 and ICOSL. Dendritic cells in particular play a specific role in migrating from the initial site of infection to the nearest lymph node to active CD4 and CD8 subsets of T cells [55].

The Adaptive Immune System

The cells of the adaptive immune system are derived from lymphoid precursors and include conventional T cells and B cells as well as innate-like lymphocytes such as B1 cells, marginal zone B cells, natural killer T (NKT) cells and γδ T cells. One of the primary differences between cells of the innate and adaptive immune system is the receptors they used to identify pathogens. Conventional T and B cell receptors are highly variable and randomly generated from a series of V, D and J gene segments by recombination-activating gene (RAG)-mediated recombination. This confers the ability for conventional T and B cells to specifically recognize nearly any antigen. The exception to this is the innate-like lymphocyte antigen receptors that preferentially recognize specific types of antigen (discussed in [56-58]). Both T cells and B cells

express single antigen receptors that are selected during development termed the T cell receptor (TCR) and the B cell receptor (BCR), respectively.

T cells and B cells reside in the lymph nodes and spleen where antigen drains from the surrounding tissue increasing the likelihood of antigen encounter. T cells and B cells clonally expand upon recognition of antigen and exert effector functions such as cytokine secretion and antibody production to eliminate invading pathogens. T cells consist of two subsets based on co-receptor expression and effector function. The TCRs of CD4 T cells recognize antigen presented by MHC class II on professional phagocytes and coordinate the adaptive immune response through production of a variety of cytokines such as IL-2, IL-4 and/or IFN-γ. CD4 T cells also provide signals necessary for B cell activation and antibody class switching. CD8 T cell TCRs detect antigen presented by professional phagocytes on MHC class I and migrate to the site of infection to produce the antiviral cytokine IFN-γ and kill infected cells expressing their cognate MHC class I: antigen complex by release of vesicles containing perforin/granzyme. B cells however, recognize soluble antigen through their BCRs, and, with the exception of innate-like B1 and marginal zone B cells, generally require CD4 T cell-mediated help in the form of cytokines and receptor: ligand interactions to become fully activated antibody secreting cells.

The adaptive immune response is also tailored to the specific type of pathogen encountered (bacteria, virus or multicellular parasite). However, adaptive immune responses are dependent on the cytokines and chemokines produced by innate immune cells after pattern recognition receptor triggering. A feature unique to the adaptive immune system is the ability of lymphocytes to form immunological memory after pathogen exposure. After antigen recognition, clonal expansion and exerting effector functions during an infection, a small fraction of antigen-specific effector CD4/CD8 T cells and B cells remain in lymphoid tissues poised to exert quick

and efficient effector functions upon re-exposure to their cognate antigen. The adaptive immune memory response is the basis of vaccine-induced immune protection (adaptive immune system reviewed by [59]). In summary, the vertebrate immune system is complex and contains many different cell types that work in concert to eliminate pathogens. Ultimately, the pattern recognition receptors of the innate immune system coordinate the initiation and activation of the entire immune response including innate and adaptive effector functions through detection of pathogen associated molecular patterns and inflammatory cytokine production. We will next focus on the role of antiviral cytokines induced during viral infections.

Type I IFN Response

The interferon (IFN) proteins are an important antiviral adaptation in vertebrates. IFN was discovered over 50 years ago as a viral restriction factor in influenza infection, and since then IFNs have been well characterized as the primary cytokine involved in coordinating the antiviral innate and adaptive immune responses [60]. There are three types of interferons; Type I, Type II and Type III. Type I IFNs consist of a single IFN-β gene and 13 different IFN-α genes that are activated during the initial sensing of viruses. IFN-γ is the only Type II IFN and serves to mediate a virus-specific immune response through activation of CD4 T cells, CD8 T cells, B cells and NK cells during viral infections. Lastly, the Type III IFNs consist of IFN-λ 1-3 and are thought to be ancestral IFNs (IFNs reviewed in [61]).

IFN-β is the first IFN induced after recognition of viral PAMPs and can be produced by any infected cell. IFN-β signaling occurs in an autocrine and paracrine manner through the heterodimeric interferon-alpha receptor 1 (IFNAR1) and interferon-alpha receptor 2 (IFNAR2) to activate Janus kinase 1 (Jak1) and tyrosine kinase 2 (Tyk2) that phosphorylate the receptor chains. Signal transducers and activators of transcription 1 and 2 (STAT1/STAT2) are recruited

to the phosphorylated receptor chains and are phosphorylated themselves to form a heterodimer. STAT1/STAT2 complexes recruit and activate interferon regulatory factor 9 (IRF-9) to form the interferon-stimulated gene factor 3 (ISGF3) complex that translocates to the nucleus to drive IFN-inducible genes by binding the interferon-stimulated response element (ISRE) (IFN signaling reviewed in [62] and [63]). Type I IFN signaling induces the expression of over 300 interferon-stimulated genes (ISGs) that induce an antiviral state, including upregulation of genes involved in inhibiting proliferation, inducing apoptosis, and positively regulating the interferon response [64]. In addition, type I IFN signaling induces MHC class I expression and antigen processing machinery to allow CD8 T cell-mediated recognition and killing of infected cells. Some ISGs are involved in direct inhibition of viral replication within infected cells. For example, the ISG PKR is activated upon binding of viral RNA and mediates inhibition of translation by phosphorylating the alpha subunit of initiation factor eIF2 [65]. Additionally, the 2-5A oligoadenylates (OAS) ISGs recognize viral RNA and produce 2-5A oligoadenylates from ATP. The 2-5A oligoadenylates then activate the endoribonuclease RNAseL (also an ISG) that degrades viral RNA [66]. Lastly, the Mx GTPase ISGs catalyze the hydrolysis of GTP to GDP to restrict viral replication by reducing cellular deoxynucleotide triphosphates (dNTPs) [67]. In summary, type I IFNs are necessary for immune responses against viruses. IFN-α/-β knockout mice are extremely susceptible to viral infections illustrating their importance to antiviral immunity [68]. Because of their potent antiviral activity, type I IFN induction is tightly regulated and controlled by PRR.

Recognition of Viral Nucleic Acid

Because viruses can infect all mammalian cells, it is extremely important that all cells are capable of detecting viral infection through PRRs and initiating a potent type I IFN response to

alert the immune system of infection. Viruses utilize all of the components of the host cell replication machinery and are essentially composed of host components making self versus non-self discrimination a challenge. All viruses possess either an RNA or DNA nucleic acid genome. These are made of host materials, but are often found in structures and locations that are distinct from host nucleic acid. Therefore, detection of virally-derived nucleic acids by PRR and activation of type I IFN production are key components of the innate antiviral immune responses. Viral nucleic acid is detected by two distinct mechanisms. First, specialized innate immune cells detect viral nucleic acid through activation of TLRs. Additionally, all cells contain intracellular nucleic acid sensors that sense viral DNA and RNA. The outcome of detection of viral nucleic acids by both mechanisms is the production of type I IFN and initiation of an antiviral immune response.

Detection of Exogenous Nucleic Acids

Specialized hematopoietic cells, such as macrophages and dendritic cells, detect viral nucleic acid and induce potent type I IFN and inflammatory cytokine responses. Importantly, this system depends on phagocytosis or endocytosis of extracellular nucleic acid produced in neighboring cells by the detecting cell and serves to alert these specialized innate immune cells to the presence of a viral infection. The pattern recognition receptors responsible for detection of exogenous viral nucleic acid are the TLRs. As previously discussed, toll-like receptors are a subset of pattern recognition receptors that recognize a variety of microbial products. All TLRs are structurally similar in that they contain extracellular leucine-rich repeats (LRRs) and a cytoplasmic Toll/Interluekin-1 receptor (TIR) domain [69]. All TLRs signal by ligand-induced dimerization and recruitment of either myeloid differentiation primary response gene (MyD88) and/or TIR-domain-containing adaptor inducing IFN-β (TRIF) and Trif-related adaptor molecule

(TRAM). In the case of MyD88-mediated signaling, IL-1R-associated kinase 1 (IRAK1) and/or IRAK2 are recruited to MyD88 and IRAK4 mediates IRAK1 phosphorylation and activation. Activated IRAK-1 is released and binds TNF-α receptor-associated factor 6 (TRAF6), which in turn promotes the ubiquitination of TGF-β-activated kinase-1 (TAK-1). TAK-1 then mediates NF-κB release from IκB and mitogen-activated protein kinase (MAPK) activation all of which serve to induce transcription pro-inflammatory cytokines such as TNF-α, interleukin-6 (IL-6) and interleukin-12 (IL-12). TLR-3 is the only exception to this general signaling cascade and instead of recruiting MyD88, TRIF is recruited to the activated receptor, which directly recruits and activates TRAF6 (TLR signaling reviewed in [70]). The TLRs responsible for recognizing viral nucleic acids are TLR3, 7, 8 and 9. These TLRs are located on the endosomal compartments of phagocytes where they are poised to recognize exogenous viral nucleic acids and induce type I IFN.

Endosomal TLR3 was initially characterized as a dsRNA binding protein by its ability to activate signaling after binding the dsRNA mimetic poly(I:C) [71]. TLR3 activates IFN-β transcription via TRIF-mediated activation of TBK-1 and subsequent phosphorylation and dimerization of IRF-3 [70]. Since dsRNA is a well-characterized viral pathogen associated molecular pattern, TLR3 was initially thought to play a role in RNA virus infections. However, subsequent in vivo studies with TLR3 deficient mice showed no role for protective role for TLR-3 in lymphocytic choriomeningitis virus (LCMV), vesicular stomatitis virus (VSV) and reovirus infections [72]. In further studies, TLR3 exerted a protective effect in murine cytomegalovirus (MCMV) and encephalomyocarditis virus (ECMV) infection during which cytokine responses were reduced in both infections and mortality increased (ECMV infection only) in TLR3 deficient mice [73-75]. Oddly TLR3 seemed to amplify inflammation, pathogenicity and/or

increase mortality in a number of infections including West Nile virus (WNV), Punta Toro virus (PTV) and influenza virus infections [76-78]. From these differing reports, we can surmise that whether TLR3 is protective or pathogenic during RNA virus infections is largely virus-specific.

TLR7 and TLR8 are highly related endosomal toll-like receptors that recognize long, viral ssRNAs as well as synthetic RNA mimetics [79]. Despite their shared ability to recognize long ssRNAs, TLR7 is also capable of detecting shorter dsRNAs rich in uridine [80]. In addition, cell type-specific physiological roles have been ascribed to TLR7 and TLR8. TLR7 is highly expressed on plasmacytoid dendritic cells (pDCs) that specialize in secreting type I IFNs (primarily IFN-α) after TLR7 or TLR9-mediated pattern recognition[81]. TLR8 however is expressed more highly on conventional DCs and monocytes and activates other pro-inflammatory cytokines such as TNF-α and IL-12 more efficiently than type I IFN [82]. Unlike TLR3, TLR7 utilizes the adaptor proteins IRAK1 and TRAF6 to phosphorylate and activate IRF7 dimerization for induction of type I IFN [70]. TLR7 plays an important role in pDC-mediated type I IFN and inflammatory cytokine responses to viral pathogens. In murine models, IFN-α production by pDCs during VSV, influenza, sendai virus (SV) and Coxsackievirus B (CVB) infections was dependent on TLR7 [83-86]. The only known physiological role for TLR8 is in human peripheral blood mononuclear cell (PBMC) recognition of human immunodeficiency virus (HIV) ssRNA [87].

The final TLR responsible for inducing type I IFN downstream of nucleic acid detection is TLR9. TLR9 recognizes unmethylated deoxycytidylate-phosphate-deoxyguanylate (CpG) DNA patterns in bacterial and DNA virus genomes [88]. Like TLR7, TLR9 is highly expressed on pDCs and induces IFN-α expression via the same signaling pathway as TLR7 upon ligand activation [70, 81]. TLR9 has a well-documented role in the pDC IFN-α response to many DNA

viruses. TLR9-deficient mice or cells infected with a variety of DNA viruses such as MCMV, herpes simplex virus 1 (HSV-1) and HSV-2 and Epstein-Barr virus (EBV) all have decreased IFN-α production by pDCs which often correlates with increased mortality [73, 89-93]. In summary, TLR3, TLR7/TLR8 and TLR9 all have important roles in the IFN response to exogenous nucleic acid in specialized innate immune cells. However, the TLRs alone do not account for the fact that *all* mammalian cells are capable of detecting viral infection to induce type I IFN.

Detection of Endogenous Nucleic Acids

All cells can detect infection by viral pathogens via mechanisms distinct from TLRs, through expression of cytosolic pattern recognition receptors that detect viral RNA or DNA genomes. Importantly, this sensing pathway enables the recognition of foreign, non-self nucleic acids introduced during viral infection and/or replication within the infected cell. Similar to toll-like receptor detection of viral nucleic acids, cytosolic detection of viral nucleic acid results in the production of type I IFN to alert innate and adaptive immune cells to the presence of viral pathogens. The pattern recognition receptors responsible for detecting cytosolic viral RNA are the retinoic acid-inducible gene I (RIG-I)-like receptors (RLRs). This family of receptors contains three related RNA helicases; retinoic acid-inducible gene-I (RIG-I), melanoma differentiation-associated gene 5 (MDA5) and laboratory of genetics and physiology 2 (LGP2) [94]. All of these three PRRs contain C-terminal DExD/H box-containing RNA helicase domains that physically bind viral RNA. However, only RIG-I and MDA5 contain N-terminal caspase recruitment domains (CARDs) that mediate protein interactions necessary for signaling [94]. Upon ligand binding, RIG-I and MDA5 are activated and bind to the adaptor mitochondrial antiviral signaling (MAVS) protein via CARD domain interactions [95-98]. This sensor/adaptor

complex then mediates the recruitment of several different TRAF family members to activate type I IFN and NF-κB signaling. TRAF3 recruitment allows TBK1-dependent phosphorylation of IRF-3 and IRF-7 that (as homo or heterodimers) induce IFN-β gene transcription. TRAF2/6 recruitment to the mitochondrial signaling complex activates the inhibitor of NF-κB kinase (IKK) complex that results in degradation of the inhibitor of NF-κB (IκB) and NF-κB translocation to the nucleus to induce target gene expression (Signaling reviewed in [99]).

The ligand specificities of RIG-I and MDA5 are not yet completely understood. RIG-I recognizes several distinct RNA ligands including single or double-stranded 5' triphosphate viral RNAs and short double-stranded RNAs [100-104]. The ligand specificity of MDA5 is even less clear. In addition to binding the synthetic RNA polymer poly I:C, MDA5 also detects long, double-stranded RNAs greater than one kilobase in length [104, 105]. Both long, double-stranded and 5' triphosphate RNAs are distinct pathogen-associated molecular patterns since host RNA is not found in long double-stranded forms and the 5' triphosphate is either masked by a 5' cap in the case of messenger RNA or removed as with transfer and ribosomal RNAs [99]. For the most part, RIG-I and MDA5 recognize distinct viral pathogens. RIG-I deficient cells and/or mice had reduce type I IFN responses to SV, VSV, HCV, Newcastle disease virus (NDV), influenza A virus and Japanese encephalitis virus (JEV) infections [105, 106]. Whereas MDA5 has a role in specifically recognizing and inducing type I IFN in response to encephalomyocarditis virus (EMCV), Theiler's virus (TV), and Mengo virus (MV) and murine hepatitis virus (MHV) infections [105, 107]. However, RIG-I and MDA5 appear to be redundant in reovirus, west nile virus and dengue virus infections [104, 106, 108]. The broad specificities of RIG-I and MDA5 allow this family to recognize the entire range of RNA viruses that infect vertebrates. The role of LGP2 in recognition of RNA viruses is not understood. Since LGP2 does

not contain a CARD domain, it was initially thought to be a negative regulator of RIG-I and MDA5 immune responses, but new studies implicate LGP2 in the IFN response to ECMV infection [109-112]. Therefore, more work is needed to determine the precise role of LGP2 in RNA virus infection.

In contrast to the well-characterized cytosolic RNA sensing pathway mediated by the RIG-I-like-receptors that leads to type I IFN activation, little is known about the IFN response to endogenous viral DNA. At least three distinct intracellular DNA detection pathways exist in vertebrates. The interferon stimulatory DNA (ISD) pathway activates a TBK1/IRF3 dependent type I IFN response after stimulation with double-stranded interferon-stimulatory DNA (ISD) [113]. Subsequent work led to the discovery of an ISD-specific endoplasmic reticulum adaptor protein stimulator of IFN genes (STING) that is responsible for directly activating TBK1 to facilitate IRF3 phosphorylation, dimerization and induction of IFN-β [114-119]. The STING-dependent ISD pathway is important for activating type I IFN upon detection of HSV-1, HCMV, vaccinia virus and *Plasmodium falciparum* DNA [118, 120]. Cytosolic DNA rich in adenosine and thymidine such as the synthetic DNA mimetic poly dA:dT can activate type I IFN through both the ISD pathway and the RNA polymerase III pathway [121-123]. In the RNA Polymerase III pathway, RNA Pol III transcribes A:T rich DNA into RNA that activates the cytosolic RNA sensor RIG-I and induces type I IFN in a MAVS-dependent manner [121, 122]. The physiological relevance of this alternative pathway is unclear, although it has been implicated in type I IFN responses to HSV-1 and EBV infections [121, 122, 124, 125]. In addition to activating type I IFN, cytosolic DNA activates an IL-1β inflammasome response. In the DNA-activated inflammasome, intracellular DNA is directly sensed by absent in melanoma 2 (AIM2). AIM2 oliogomerizes and promotes the recruitment of the adaptor protein associated speck-like

protein (ASC) and subsequent activation of caspase-1 through CARD-domain interactions. Caspase-1 then mediates processing of pro-IL-1β to the biologically active form of IL-1β for secretion (AIM2 inflammasome discussed in [126-129]). The AIM2 inflammasome is responsible for IL-1β production and amplification of inflammation in response to DNA virus infections such as murine cytomegalovirus and vaccinia virus infections [130]. The mechanisms of RNA Polymerase III interferon induction and AIM2 inflammasome activation are well characterized. However, despite intensive research, the exact proteins involved in ISD sensing remain largely unknown.

Several ISD receptors have been described, but the use of different cell types (immortalized cell lines, primary cells) and ligands (Poly dA:dT, ISD, viral infections) has complicated our understanding of DNA recognition upstream of STING. Interestingly, STING itself was recently identified as a direct sensor for bacterial second messenger cyclic-dinucleotides, but not interferon stimulatory DNA upstream of type I IFN induction [131]. One of the first proposed cytosolic DNA sensors identified was DNA-dependent activator of IFN-regulatory factors (DAI). This sensor was shown to induce type I IFN on overexpression, show reduced IFN responses during siRNA mediated knockdowns, and directly interact and activate TBK1 and IRF3 [132]. However, DAI deficient mice were generated and found to have normal type I IFN responses to DNA ligands and DNA virus infections, indicating that DAI may not actually be a cytosolic DNA sensor or that it shares functional redundancy with other DNA sensor proteins [133]. Additionally, a member of the DExD/H helicase family, DDX41, was implicated in sensing interferon stimulatory DNA specifically in dendritic cells. DDX41 bound to cytoplasmic DNA, interacted directly with the ISD adaptor protein STING and DDX41 siRNA knockdown reduced IRF3-dependent type I IFN activation in response to DNA ligands

and HSV-1 infection [134]. Human IFI16 and its murine homolog, IFI204, were also implicated in ISD sensing [135]. IFI16 and IFI204 (also called p204) are members of the PYHIN (also known as AIM2-like receptor, ALR) family of proteins that contain Pyrin signaling adaptor domains and HIN DNA binding domains. The cytosolic DNA inflammasome sensor AIM2 is also part of this family making IFI16 and IFI204 attractive candidate ISD sensors (PYHIN family reviewed in [136]). IFI16 and IFI204 were identified as DNA binding proteins that interacted with the ISD adaptor protein STING. Additionally, IFI16 or IFI204 knockdown resulted in reduced IRF3 phosphorylation and IFN-β induction in response to transfected DNA ligands as well as HSV-1 infection [135]. In addition to its role in type I IFN induction, IFI16 has been implicated in IL-1β inflammasome activation after sensing Kaposi sarcoma-associated Herpes virus (KSHV) genome [137].

Additionally, two cytosolic DNA sensors that enhance type I IFN activation independent of STING, TBK1 and IRF3 were identified. Leucine-rich repeat in flightless I-interacting protein 1 (LRRFIP1) was identified as a DNA-binding protein that activated the transcription factor β-catenin. β-catenin enhanced IRF3-mediated type I IFN production by partnering with IRF3 on the IFN-β promoter to recruit the transcriptional activator p300 [138-140]. The DNA damage repair protein Ku70 was identified as a DNA-binding protein that mediated activation of type III IFN, IFN-λ1, instead of type I IFN [141]. Upon knockdown of Ku70, *Zhang et al.* found reduced IFR1/IRF7-dependent IFN-λ1 induction after DNA ligand stimulation. Despite the identification of many candidate cytosolic DNA sensors, further studies using primary cells, ISD ligand instead of poly dA:dT and functional assays in knockout mice are necessary to determine their relevance and potential redundancy in DNA sensing.

In summary, detection of extrinsic and intrinsic viral nucleic acids by somatic and hematopoietic cells is an important part of the innate immune system. The type I IFN induced by these two distinct mechanisms is critical to initiate a productive immune response and clearance of viral infections. However, the discrimination of self and non-self nucleic acids by the innate immune system is imperfect.

Innate Immune Detection and Autoimmunity

Despite the importance of type I IFN in clearing viral infections, improper activation of this potent cytokine has been shown to mediate autoimmune pathogenesis. Several autoimmune diseases, such as inflammation of the salivary glands in Sjogren's syndrome, skin inflammation in psoriasis and systemic inflammation in systemic lupus erythematosus (SLE), are associated with elevated levels of type I IFN [142]. The elevated levels of type I IFN observed in SLE results from improper detection of self-derived nucleic acids by TLRs [143]. Three distinct tolerance mechanisms exist to prevent TLR recognition of self nucleic acid. The first mechanism of nucleic tolerance is the localization of nucleic acid-specific pattern recognition receptors to endosomal compartments where they are more likely to encounter viral, and less likely to encounter host, nucleic acids [144, 145]. Additionally, extracellular RNases and DNases degrade host nucleic acid after cell death, but not nucleic acid hidden within viral or bacterial pathogens, to assure detection of pathogenic nucleic acid only after localization to endosomal compartments [146, 147]. Lastly, pathogen nucleic acids slightly differ from host nucleic acids. Bacterial and viral DNA genomes contain unmethylated CPG motifs that are recognized by TLR9, while viral RNA contains 5’ triphosphate moieties and/or double-stranded structures that are detected by TLR7. Host nucleic acids either lack or mask these motifs and therefore should not activate these receptors. However, these tolerance mechanisms are imperfect. Self-nucleic acid detection by

TLR7 and TLR9 on pDCs and autoreactive B cells has been shown to contribute to the pathogenicity of SLE [143]. SLE is characterized by aberrant TLR9 activation on pDCs by circulating self-DNA/ nucleoprotein/ antibody containing complexes [148-150]. pDCs are also activated by TLR7 through self-RNA/ small nuclear ribonucleoprotein/ antibody containing complexes during SLE [151]. Both TLR7 and TLR9 activation on pDCs results in potent type I IFN induction that contributes to autoimmune pathology. In addition, TLR7 and TLR9 activation by immune complexes on autoreactive B cells contributes to autoantibody production [152]. Despite their well-known roles in exacerbating SLE, the events that initiate self-nucleic acid detection by TLR7 and TLR9 are unknown. Interestingly, mutations in the extracellular DNase I cause an IFN-mediated lupus-like disease in mice and some human SLE patients have DNase I mutations, suggesting that sensing of self-nucleic acid from apoptotic material through TLR7/TLR9 can initiate autoimmune disease [147, 153]. All of these examples of IFN induction mediated by nucleic acid detection are examples of the exogenous detection of nucleic acid by specialized immune cells (pDCs). There is only one known example of endogenous nucleic acid detection leading autoimmune disease.

Loss of function mutations in 3' repair exonuclease 1 (Trex1) were identified as causative agents of a type I IFN-mediated autoimmune disease called Aicardi-Goutières syndrome (AGS) [154]. This severe autoimmune disease is normally identified in infancy with symptoms that include elevated type I IFN and cerebrospinal leukocyte infiltrates [155-157]. Interestingly, Trex1 is an IFN-inducible 3 to 5' intracellular DNA exonuclease that metabolizes single-stranded DNA substrates, but it was not understood how mutations in this enzyme caused AGS. Using a murine model of Trex1 deficiency, *Stetson et al.* found that Trex1 was a necessary negative regulator of the interferon stimulatory DNA pathway required to metabolize

endogenous DNA from reactivated retroviruses and retrotransposons. Mice deficient in Trex1 accumulated cytosolic DNA from retroelements and transposons that activated a potent type I IFN response mediated by the ISD pathway. Aberrant activation of the ISD pathway led to lethal inflammatory myocarditis that could be rescued by crossing Trex1 deficient mice to STING, IRF3, IFNAR and RAG2 deficient mice, indicating that the pathology of AGS is mediated by the immune system [158, 159]. In addition to Trex1, mutations in ribonuclease H2 (RNaseH2) and sterile alpha motif domain and HD domain–containing protein 1 (SAMHD1) were also identified as causative agents of AGS [160, 161]. RNaseH2 is an RNA endonuclease that degrades the RNA strand of RNA:DNA hybrids that occur from a multitude of cellular processes such as cellular DNA replication and endogenous retroelements replication intermediates. SAMHD1 is a triphosphohydrolase that converts deoxynucleoside triphosphates into deoxynucleoside diphosphates and was shown to play a role in restriction of HIV infection [162]. However, the role of these two proteins in AGS is not known, although they are potentially involved in regulation of endogenous retroelements and negatively regulating the ISD pathway [163]. These results revealed a novel, cell intrinsic mechanism for initiation of autoimmunity caused by chronic activation of the ISD pathway.

In summary, all organisms from single-celled bacteria all the way to vertebrates have developed ways to combat viral pathogens. Vertebrates have evolved complex multicellular immune systems composed of innate and adaptive immune cells that participate in host defense to viral infections. The development of the antiviral cytokine interferon by the vertebrate immune system allowed the development and regulation of a complex set of antiviral genes induced upon interferon signaling. Pattern recognition receptors that recognize viral nucleic acids couple viral pathogen-associated molecular pattern detection to type I IFN induction. There are

two distinct types of pattern recognition receptors. First, endosomal receptors expressed on specialized innate immune cells detect exogenously derived viral nucleic acids without actually being infected by the virus. In addition, there are intracellular sensors expressed broadly in most cell types that detect endogenously generated viral nucleic acid and allow any infected cell to alert the innate immune system to viral infection by type I IFN production. Among these intracellular nucleic acid sensing pathways, the components of the RNA sensing pathway are well defined, while the sensor and adaptors of the DNA sensing pathways remain elusive. Recently Trex1 was identified as a negative regulator of the DNA sensing pathway and deficiency in Trex1 led to constitutive activation of the ISD pathway and lethal autoimmune disease.

Although several candidate sensor proteins have been proposed, our understanding of the ISD pathway is still incomplete. Several important unanswered questions remain: What are the proximal signaling components, including the sensor(s), of the ISD pathway? Our knowledge of the ISD pathway has been complicated by the differing methods used to study intracellular DNA sensing. The ISD pathway has been studied in immortalized cell lines often utilizing poly dA:dT instead of DNA oligos. As discussed in great depth later, we discovered that immortalized cells do not respond robustly to cytosolic DNA, and therefore, the ISD pathway must be studied in primary cells. In addition, the use of poly dA:dT induces both ISD- and RNA Pol III-dependent type I IFN. This makes data generated using poly dA:dT as an "ISD ligand" difficult to interpret. Our goal was to develop new methods to identify the components of the ISD pathway. We utilized the novel Trex1 deficiency model of autoimmunity to identify novel Trex1 interacting proteins by Trex1 pulldown and yeast 2-hybrid that may play a role in the ISD pathway. In addition, we used a direct approach to identify ISD sensors by utilizing microarray, cytoplasmic

DNA pulldowns and characterization of the AIM2-like receptor (ALR) family of proteins implicated in ISD sensing. Here, we identified novel Trex1 interacting proteins, novel cytosolic DNA binding proteins, STING-dependent ligands, five novel ALR genes in mice and characterized the functions of all known murine and human ALRs. Our analyses reveal functional redundancy among ISD receptors and extensive evolutionary diversity that is recurrently recreated in mammalian genomes.

Section 2: Trex1 Interacting Proteins in Interferon Stimulatory DNA Sensing

Introduction

Mammalian cells are capable of detecting nucleic acids produced by infecting viruses and initiating a type I IFN response by both extrinsic and intrinsic mechanisms. Extrinsic nucleic acid detection involves specialized immune cells such as macrophages and dendritic cells. Toll-like receptors (TLRs) are expressed on the surface of these specialized cells and recognize a variety of bacterially and virally derived ligands including nucleic acids. TLR3 and 7 specifically recognize viral RNA ligands while TLR9 recognizes hypomethylated CPG DNA ligands to initiate a type I IFN response. Importantly, these specialized cells and receptors only recognize nucleic acid ligands made outside of the detecting cell. The intrinsic nucleic acid ligand sensing pathway is shared among all mammalian cells. The sensing of cytosolic RNA ligands has been well characterized with RIG-I and MDA-5 acting as the primary sensors to activate type I IFN through the mitochondrial adaptor protein MAVS in a TBK-1and IRF-3/IRF-7 dependent manner [95-98, 100, 101, 105, 106, 164, 165]. The intrinsic type I IFN response to cytosolic DNA ligands termed the interferon stimulatory DNA (ISD) pathway is dependent on STING, TBK-1 and IRF-3, but little else is known about the sensors and/or adaptor proteins involved upstream of transcription factor activation [113, 114, 118].

The extrinsic toll-like receptors and intrinsic cytosolic sensors are the primary ways type I IFN is produced in response to viral pathogens. However, host cells also expresses potential nucleic acid ligands. Defective discrimination between self and virally-derived nucleic acid ligands or clearance of self-generated nucleic acids can cause type I IFN mediated autoimmunity. For example, DNase I mutations in humans are associated with lupus erythematosus and DNase I

deficiency in mice can cause a lupus-like disease [147, 153]. For the most part, defects in the extrinsic nucleic acid ligand sensing pathway have been shown to mediate type I IFN autoimmunity. Toll-like receptors are necessary for autoreactive B cell activation and proliferation in response to chromatin immune complexes [166]. TLR7 and TLR9 are necessary for autoantibody production in a mouse lupus model, and TLR-7 duplication in mice correlates with development of autoimmunity [152, 167, 168]. In addition, TLRs on plasmacytoid dendritic cells (pDC) in human psoriasis and liver macrophages in a DNase II-deficient mouse model are activated by extracellular DNA and promote autoimmunity [169-171].

The cases discussed above are examples of autoimmunity mediated by the extrinsic pathway of nucleic acid sensing where specialized responding cells sense exogenous nucleic acid produced by other cells. There is only one known autoimmune disease model in which the accumulation of endogenous nucleic acid ligands activates a cytosolic sensor(s) to cause disease. Defects in the exonuclease Trex1 lead to type I IFN dependent autoimmune disease. Trex1 was initially characterized as the most abundant mammalian single stranded 3' → 5' DNA exonuclease and was thought to play a role in base excision repair due to its homology to other DNA editing enzymes and weak base excision activity [172-174]. Trex1 was first implicated in autoimmune disease by genetic studies in patients showing that homozygous loss of function mutations in Trex1 cause Aicardi-Goutières syndrome (AGS) [154]. AGS is a severe autoimmune disease usually identified in infancy. AGS patients suffer from a variety of symptoms that mimic those of a congenitally acquired viral infection and include encephalitis, intracranial calcifications, lymphocyte infiltrates, type I IFN in cerebrospinal fluid, and demyelination of neurons that results in permanent psychomotor retardation although no AGS-associated virus has been identified [155-157]. Distinct mutations in Trex1 cause a variety of

other autoimmune disorders. Some mutations in Trex1 cause a mild lupus-like disease called familial chilblain lupus that is characterized by lesions on the extremities, and Trex1 mutations are found more frequently in systemic lupus erythematosus patients than the general population [175-178]. A C-terminal frameshift mutation in Trex1 that results in the mislocalization of Trex1 protein causes retinal vasculopathy with cerebral leukodystrophy (RVCL) that presents in middle age with symptoms that include loss of vision, stroke and dementia [179]. Lastly, Trex1 deficiency in mice causes a different pathology than in human disease and results in lethal autoimmune inflammatory myocarditis [180]. Despite extensive data illustrating a role for Trex1 in multiple autoimmune diseases, the mechanism of Trex1-mediated autoimmune disease was only recently identified.

Stetson et al. identified Trex1 as an IFN-inducible DNA binding protein in immunoprecipitations of cytosolic biotinylated DNA ligands [159]. They also showed that the lethal autoimmune inflammatory myocarditis seen in Trex1-deficient mice could be rescued by crossing to IRF-3, STING, IFNR-1 and RAG-2-deficient mice [158, 159]. These data imply that the autoimmune disease of Trex1-deficient mice is initiated intrinsically through an unknown cytosolic receptor that signals through the endoplasmic reticulum adaptor protein STING to activate IRF-3 and a type I IFN response that is responsible for the improper activation of lymphocytes. Trex1-deficient murine embryonic fibroblasts and human AGS fibroblasts had previously been shown to accumulate cytosolic single stranded DNA [181]. *Stetson et al.* discovered that this accumulating cytoplasmic DNA isolated from Trex1-deficient mouse hearts originated from reactivated endogenous L1 and LTR retrotransposons. Under normal conditions, Trex1 metabolizes retroelement DNA, but in the absence of Trex1 this accumulating DNA activates the ISD pathway [159]. These data firm establish Trex1 as a negative regulator of the

ISD pathway. Confirming these results, Trex1 has subsequently been shown to inhibit innate immune detection of human immunodeficiency virus (HIV) and facilitate integration by metabolizing the viral genome before detection by the ISD pathway [182, 183].

In addition, these results revealed a novel, cell intrinsic mechanism for initiation of autoimmunity caused by chronic activation of the ISD pathway. However, one fundamental, unanswered question remains: What are the proximal signaling components of the ISD pathway? Our goal was to develop new methods to address this intriguing question using the novel Trex1deficiency model of autoimmunity. In order to elucidate the proteins involved in ISD signaling, we developed a Tandem Affinity Purification strategy and a yeast 2-hybrid screen to identify Trex1 interacting proteins and test them for relevance to the ISD pathway. Since the events responsible for detecting accumulation of cytoplasmic DNA and initiating autoimmunity are precisely the same events essential for inducing protective immunity to infectious virus, this model system allowed us to define basic principles of nucleic acid detection and host defense against viruses.

Materials and Methods

Constructs

Wild-type murine Trex1 and Trex1 catalytic (Trex1-Cat) mutant containing both H195A and D200N amino acid base changes were cloned into a BioEase 6xHis MSCV-IRES-GFP vector. To created BioEase 6xHis MSCV-IRES-GFP, we modified the MSCV-IRES-GFP vector to contain an N-terminal endogenous biotinylation site (BioEase) followed by two tobacco etch virus (TEV) protease cleavage sites and a 6 repeating histidine (6xHis) tag. Murine Trex1, Trex1 catalytic mutant (Trex1-Cat) and C-terminal truncated Trex1 catalytic mutant (Trex1-T) were all

subcloned into pEGFP-N1 to create Trex1-eGFP fusion proteins. Mouse IFITM3 was cloned into pCDNA3.1 with a C-terminal V5-tag and pCMV-HA to express a N-terminal HA-tag. Murine Brd7, Btf3l4, Caf1a, Pias1, Rfc2, Uba2 were cloned into pCMV-HA to express a N-terminal HA-tag.

Cells and Trex1 reconstituted cell lines

HEK 293T cells were purchased from Invitrogen and HeLa cells were purchased from American Type Culture Collection. Wild-type and Trex1 deficient mouse embryonic fibroblasts (MEF) were generated by Barnes et al as described [181]. For Trex1 deficient MEF stable reconstitution, murine leukemia virus (MLV) pseudotyped retrovirus was generated by transfection of 2.5 x 10^6 HEK 293T cells in 10-cm plates with 10μg BioEase 6xHis MSCV-IRES-GFP Trex1 and Trex1-Cat constructs and 10μg of pCL-Eco for 48 hours. 0.5 x 10^6 Trex1 deficient MEFs were transduced with HEK 293T supernatants for 48 hours and split into 10-cm plates to grow to confluence. Transduction efficiency was assayed by resuspending 0.1 x 10^6 MEFs in 1% fetal calf serum (HyClone) PBS for flow cytometry on a FACScan 2 (Becton Dickinson) for GFP expression.

Luciferase reporter assays

0.15 x 10^4 HEK 293T cells were plated in 24-well plates and transfected with 25 ng ISRE-luciferase reporter plasmid alone or together with 160 and 400 ng and 1μg of IFITM3 pCDNA3.1-V5 or IFITM3 pCMV-HA or 1 μg TBK1 as a control. Luciferase activity was determined 24 hours later using the Luciferase Reporter Assay System (Promega) according to the manufacturer's instructions and read using a Centro LB 960 Luminometer (Berthold Technologies).

Immunofluorescence

For IFITM3, 3 x10^4 HeLa cells plated on glass coverslips in 24-well plates were transfected with 500 ng IFITM3 pCDNA3.1-V5. For Trex1 interacting proteins, 100- 1000 ng Trex1-eGFP and Trex1-Cat-eGFP alone or together with 100- 1000ng IFITM3-HA, Brd7-HA, Btfl34-HA, Caf1a-HA, Pias1-HA, Rfc2-HA or Uba2-HA. Cells were fixed 24 hours later with 2% paraformaldehyde in PBS for 15 minutes, permeabilized for 15 minutes with 0.1% Triton-X 100 PBS, and blocked in 2% FCS PBS for 1 hour at room temperature. Cells were first stained with anti–HA-Tag (6E2) mouse antibody (Cell Signaling Technology; 1:50), anti-V5 mouse antibody (Invitrogen; 1:500) or anti-calnexin rabbit antibody (BD Pharmigen; 1:200) for 1 hour at room temperature. Anti-mouse Alexa Fluor 488 (Invitrogen; 1:200), anti–mouse Alexa Fluor 568 (Invitrogen; 1:200) and/or anti-rabbit Alexa Fluor 568 (Invitrogen; 1:200) were used to counterstain for 1 hour each at room temperature. Coverslips were mounted with VectaShield Hardmount with DAPI (Vector Laboratories) and visualized with a Leica SP1 confocal microscope (Leica). Images were taken with the 40X oil objective in the 488 and 568 dichroic channels using Leica Confocal Software (Leica) and pseudocolored in Photoshop CS4 (Adobe).

Immunoprecipitation

For Trex1 interacting proteins, 0.3 x 10^6 HEK 293T plated in 6-well plates were transfected with 2.5 μg BioEase 6xHis Trex1 together with 2.5 μg of Brd7 pCMV-HA, Btf3l4 pCMV-HA, Caf1a pCMV-HA, Pias1 pCMV-HA, Rfc2 pCMV-HA or Uba2 pCMV-HA for 24 hours. Cells were lysed in buffer containing 0.2% Triton X-100, 50mM HEPES pH 7.5, 150 mM NaCl, 1 mM DTT and 1X protease Inhibitor Cocktail (Roche) and Trex1 was immunoprecipitated using 50 μL streptavidin-coupled Dynabeads (Invitrogen) for 1 hour at 4°C. Protein was eluted from streptavidin beads by incubation at 100°C with 1X sample loading buffer (25mM Tris pH 6.8, 2% SDS, 10% glycerol, 0.012% bromophenol blue and 0.2% β-mercaptoethanol), loaded onto an

denaturing SDS gel and probed with anti-Trex1 mouse antibody (BD Transduction Laboratories; 1:2,500) or anti–HA-Tag (6E2) mouse antibody (Cell Signaling Technology; 1:2,500) followed by goat anti-mouse HRP (Jackson Immuno Research; 1:5,000) and visualized by ECL substrate (Thermo Scientific).

Tandem affinity purification

BioEase 6xHis MSCV-IRES-GFP wild-type Trex1, Trex1-Cat and Empty transduced Trex1 deficient MEFs were plated at 2.5 x 10^6 cells per 10-cm plate and treated with 0 mM, 0.2 mM, 1 mM and 5 mM biotin 24 hours prior to harvest. Three plates were pooled, washed and lysed in buffer containing 0.2% Triton X-100, 50mM HEPES pH 7.0, 150mM NaCl, 1mM DTT and 1X Protease Inhibitor Cocktail (Roche). Biotinylated Trex1 was isolated using streptavidin-coupled Dynabeads (Invitrogen) for 30 minutes to 16 hours at 4°C. Isolated Trex1 was removed from the Dynabeads by incubation with 100 units tobacco etch virus (TEV) protease (Promega) for 1-16 hours at 4°C in buffer without protease inhibitors. A second round of purification was performed utilizing the 6X histidine-tag on the resulting supernatant for 1 hour at 4°C with 4 mg of Talon Dynabeads (Invitrogen) in lysis buffer lacking protease inhibitors and DTT. Protein was eluted from Talon beads by incubation at 100°C with sample loading buffer (25mM Tris pH 6.8, 2% SDS, 10% glycerol, 0.012% bromophenol blue and 0.2% β-mercaptoethanol), loaded onto an denaturing SDS gel and probed with anti-Trex1 at 1: 2,500 (BD Transduction Laboratories) followed by goat anti-mouse HRP (Jackson Immuno Research) at 1: 5,000 and visualized by ECL substrate (Thermo Scientific). For mass spectrometry of Trex1 interacting proteins, samples were run on a denaturing SDS gel (Invitrogen), stained with GelCode Blue Coomassie (Thermo Scientific) for 20 minutes, and bands unique to Trex1 purification were excised and sent to the Stanford University Mass Spectrometry Core (Palo Alto, CA) for identification.

Yeast 2-hybrid screen by Protein Links

Trex1 catalytic mutant lacking the C terminal 79 amino acids was cloned into pCWX200 to create a Tet repressor (Tet-R) fusion protein that was used as bait. Ura3 gene expression under the control of the Tet operon was suppressed in the presence of Trex1-Tet-R bait. 10 million cDNA library B42 (the transcriptional activator of the Tet operon) fusion clones generated from mouse spleen were screened with Trex1-Tet-R bait for growth in uracil-free media. Gal4 expression under the control of the Tet operon was used to validate positives by screening for growth in media containing lactose.

Results

Tandem affinity purification of Trex1 revealed candidate ISD signaling components

Stetson et al previously discovered that Trex1 metabolizes single stranded DNA ligands that accumulate in the cytoplasm and is an essential negative regulator of the ISD pathway [159]. Based on these findings, we hypothesized that Trex1 is part of a protein complex containing signaling components relevant to the ISD pathway. Therefore, we used Trex1 to biochemically purify and characterize these proximal signaling components. We developed a tandem affinity purification (TAP)-tagged form of Trex1 to enable rapid purification of Trex1 and its associated proteins. The tag consists of an *in vivo* biotinylation sequence and a 6X histidine tag separated by a Tobacco etch virus (TEV) cysteine protease cleavage site (Fig. 2.1A). The biotinylation site allowed a high affinity first round of purification by Streptavidin beads, followed by release of the protein from the beads with TEV protease, and then a subsequent second purification step using the histidine tag for increased purity. We retrovirally-reconstituted a Trex1-deficient mouse embryonic fibroblast (MEF) cell line with the TAP-tagged Trex1. Cell lines were

generated using both TAP-tagged WT Trex1 and a catalytically inactive mutant of Trex1 that is able to bind but not excise DNA (Fig. 2.1B). We then defined the conditions necessary for Trex1 purification (Fig. 2.1C). Moreover, we discovered that addition of excess biotin to the media before purification dramatically improves *in vivo* biotinylation, and therefore, recovery of the Trex1 protein (Fig. 2.1D).

After optimizing purification conditions, we utilized this scalable system to purify large quantities of Trex1 protein. Several Trex1 interacting proteins were identified by mass spectrometry using this approach (Fig. 2.2A). One of these proteins interferon-inducible transmembrane protein 3 (IFITM3) may be involved in the signaling events of the ISD pathway. The IFITM family containing IFITM1, 2, 3 and 5 was first characterized as a family of type I IFN inducible proteins containing two conserved transmembrane domains [184-187]. Since their discovery this family has been implicated in a variety of cellular functions from cancer, germ cell growth and development, bone mineralization and cell signaling [188-193]. Since a clear role for the IFITM proteins had not been well described, we tested IFITM3 for a role in the ISD pathway.

Immunofluorescence of IFITM3 in HeLa cells showed punctate staining that may indicate localization to intracellular vesicles. However, Trex1 has not been shown to localize to intracellular vesicles and co-staining of IFITM3 with calnexin (an endoplasmic reticulum marker that we used as a Trex1 proxy) did not show co-localization (Fig. 2.2B). We next wanted to determine if overexpression of IFITM3 was sufficient to induce type I IFN through the ISD pathway so we overexpressed IFITM3 in HEK 293T cells with an interferon-stimulated response element (ISRE) reporter. We found that IFITM3 did not induce ISRE-luciferase at any concentration of plasmid suggesting that IFITM3 is not sufficient to drive type I IFN (Fig. 2.2C). However, these data are preliminary and need further confirmation in the IFITM3 knockout mice

or by siRNA mediated knockdown of IFITM3.

Yeast 2-hybrid screen identified six novel Trex1 interacting proteins

Because the initial studies utilizing mouse embryonic fibroblasts (MEFs) reconstituted with the TAP tagged form of Trex1 did not result in the recovery of many bands to identify by mass spectrometry, we initiated a yeast 2-hybrid screen through Protein Links to identify potential Trex1 interacting proteins. Protein Links utilizes a Tet Operon system where bait is fused to the Tet repressor (Tet-R) and interactions with B42 fusion proteins drives expression of the URA3 gene and subsequent growth in uracil-free media. We designed a catalytically inactive and C-terminally truncated form of Trex1 to use as Tet-R fused bait in this screen to prevent cytotoxicity due to overexpression of Trex1 and to allow nuclear localization since Trex1 is normally found on the endoplasmic reticulum (Fig. 2.3A). Trex1 bait was screened against a B42 fusion protein library generated from mouse spleen and potential interacting proteins were screened for both growth in uracil free media and LacZ production for secondary validation (Fig. 2.3B).

Six Trex1 interacting proteins were identified by yeast 2-hybrid screen (Fig. 2.3C,D). Out of 17 positive clones, bromodomain containing protein 7 (Brd7) and chromatin assembly factor 1 subunit a (Chaf1a) were both identified as Trex1 interacting proteins by one clone each. Basic transcription factor 3 like-4 (Btf3l4), protein inhibitor of activated STAT 1 (Pias1), replication factor C subunit 2 (RFC2) and SUMO activating enzyme 2 (Sae2) were recognized as Trex1 interacting proteins by 4, 2, 4 and 5 clones respectively. Relatively little is known about these proteins especially in the context of viral infection. Chaf1a and Rfc2 are both necessary components for DNA replication. Chaf1a (also known as p150) is the largest subunit in a 3-subunit protein that is acts as a histone chaperone and is essential for depositing histone 3 and 4

on newly synthesized DNA [194-197]. Rfc2 (also known as p40) also is a member of a 5-subunit protein complex and it functions to load the proliferating cell nuclear antigen (PCNA) sliding clamp onto template-primer junctions on replicating strands of DNA [198].

Interestingly, both Pias1 and Sae2 are involved in the sumoylation pathway. Sumoylation is a protein conjugation system similar to ubiquitination, but the end result is a SUMO moiety attached to the target protein instead of ubiquitin (reviewed in [199]). Sae2 acts as an E1 SUMO ligase in a heterodimeric complex with Sae1 to activate SUMO proteins and transfer them to the E2 SUMO ligase, Ubc9 [200]. Pias1 is an E3 SUMO ligase that transfers the SUMO onto it final target and has been shown to sumoylate p53 to activate p53-dependent gene transcription and STAT-1 to down regulate the type I IFN response during viral infection [201-205]. Lastly, both Brd7 and Btf3l4 are transcription factors with unknown functions. Brd7 is a tumor suppressor, transcriptional repressor and cell growth inhibitor that has been identified in some breast cancer tumor cell lines [206, 207]. Little to nothing is known about Btf3l4 besides its predicted function as a transcription factor.

Interestingly, all six interacting proteins were nuclear proteins. This was unexpected since Trex1 localizes to the endoplasmic reticulum (ER) [159]. The ER membrane and the nuclear envelope are contiguous, so we used immunofluorescence to see if Trex1 localized to the inner nuclear membrane (INM) in addition to the ER. We found that Trex1 localized to the INM since it co-localized with the INM marker, emerin (Fig. 2.4A). We also observed that overexpression of Trex1-GFP resulted in the disruption of normal emerin localization from the INM to vesicular structures in nearly 80% of cells (Fig. 2.4B). However, Trex1-GFP still co-localized with emerin despite the disturbance of the INM providing further evidence that Trex1 is located on the same cellular structure as emerin (Fig. 2.4A).

With this new information about the localization of Trex1, we next wanted to confirm that the identification of the interacting proteins was not an artifact of the nuclear localization of the C-terminal truncated form of Trex1 that we used in the yeast 2-hybrid assay. To confirm Trex1 interactions with these proteins, we used the full length, wild-type form of Trex1 along with the cloned form of each Trex1 interacting protein expressing a hemaglutinin (HA) tag for co-immunoprecipitations. We confirmed Trex1 interaction with three (Brd7, Chaf1a and Rfc2) out of six interacting proteins (Fig. 2.5A). The three unconfirmed interacting proteins were expressed at extremely low levels within the cell and resulted in inadequate amounts of protein to immunoprecipitate with Trex1.

In an alternative approach, we used immunofluorescence to look for co-localization of the Trex1 interacting proteins with Trex1. All of the Trex1 interacting proteins identified by the yeast 2-hybrid screen (except for Rfc2) localized to the nucleus, and we observed partial co-localization of wild type Trex1 with all of the interacting proteins (Fig. 2.5B). The truncated Trex1 catalytic mutant (Trex1-Cat-T) altered the localization of Trex1 from the INM and ER to the entire nucleus and increased the co-localization of the interacting proteins with Trex-1. This suggests that under normal conditions only a small proportion of Trex1 on the INM interacts with these nuclear proteins and that without proper ER/INM localization Trex1 is capable of interacting with larger pools of the Trex1 interacting proteins.

Discussion

The signaling components of the interferon-stimulatory DNA (ISD) pathway still remain largely unknown. The discovery of Trex1 as a negative regulator of the ISD pathway presented a new model system in which to study the ISD pathway. We reconstituted the Trex1 knockout

MEF cell line with a TAP-tagged form of Trex1, and using this new method identified a novel Trex1 interacting protein, IFITM3. Using co-immunoprecipitations, we confirmed the interaction between IFITM3 and Trex1. However, immunofluorescence and ISRE reporter assay did not show a definitive role for IFITM3 in the ISD pathway.

Despite the initial identification and subsequent confirmation of interaction of Trex1 and IFITM3, we were unable to determine a role for IFITM3 in the ISD pathway. There could be several reasons for this. The tandem affinity purification of IFITM3 with Trex1 could be an artifact of the new method we developed, and under physiological conditions Trex1 may not actually interact with IFITM3. Addition of the endogenous biotinylation site or 6 histidine repeat tag on Trex1, the streptavidin or Talon Dynabeads used in the purification, or the disruption of normal localization pattern of either or both Trex1 and IFITM3 during cell lysis may have contributed to an artificial interaction.

Alternatively, IFITM3 and Trex1 may interact, but their interaction may be important for other cellular functions distinct from ISD pathway activation of type I IFN. Since our initial characterization of IFITM3, new data has emerged showing a role for IFITM3 in preventing RNA (influenza, West Nile, dengue) and DNA (HIV-1 and VSV) virus infections by preventing viral membrane fusion in the late endosomes by increasing acidification of these compartments [208-211]. These studies confirmed the vesicular localization of IFITM3 we observed by immunofluorescence, but ascribe a completely distinct role to IFITM3 other than ISD pathway activation. If the interaction between IFITM3 and Trex1 occurs *in vivo*, it is possible that IFITM3 could bring Trex1 to the endosome during viral infection to metabolize DNA virus genomes to restrict infection.

Lastly, IFITM3 may play a role in the ISD pathway, but we have not done the definitive experiments. To address this, we could return to the ISRE-luciferase assays in the presence of the ISD adaptor protein STING, since the murine AIM2-like receptors investigated in Section 3 only activated ISRE-luciferase when co-transfected with this adaptor protein. We could also test the IFITM family knockout mouse for deficiencies in type I IFN activation with DNA viruses and transfected DNA ligands. Any potential phenotype could be further substantiated with rescue by reconstitution of the knockouts with IFITM3.

In an alternative approach to tandem affinity purification to identify Trex1 interacting proteins, we initiated a yeast 2-hybrid screen using Trex1 as bait. The screen identified 6 novel Trex1 interacting proteins: Brd7, Btf3l4, Chaf1a, Pias1, Rfc2 and Sae2. We were able to confirm the interaction of Trex1 and these proteins by immunoprecipitation and immunofluorescence. All of these Trex1 interacting proteins were localized to the nucleus, prompting us to re-evaluate the localization of Trex1 and discover that in addition to the ER, Trex1 is also localized to the inner nuclear membrane.

We reasoned that Trex1 and its interacting proteins are involved in sumoylation since both Pias1 and Sae2 have well described roles as E3 and E1 sumo ligases. However, the potential target(s) of sumoylation by Trex1 interacting proteins remained unclear. We hypothesized that proliferating cell nuclear antigen (PCNA) is a target of Trex1-associated sumoylation. Previous work has shown PCNA sumoylation during S-phase to prevent unwanted recombination events and allow sufficient time for DNA damage repair [212, 213]. Additionally, Chaf1a and the RFC complex, which includes Rfc2, are both capable of interacting directly with PCNA [214-216]. The last line of evidence linking Trex1 to PCNA is the ability of the nuclease RNaseH2, which cleaves the RNA strand in RNA:DNA hybrids, to bind PCNA [217, 218]. Interestingly,

mutations in RNaseH2 also cause Aicardi-Goutières syndrome (AGS) identical to mutations in Trex1, therefore placing Trex1 and RNaseH2 in the same retroelement metabolizing pathway [155, 160]. Since RNaseH2 interacts with PCNA, we reasoned that Trex1 may indirectly interact with PCNA through Chaf1a and/or Rfc2 and that the anti-transposition/retroviral response may utilize PCNA as a scaffold. We also hypothesized that association of Trex1 with PCNA would allow Pias1 and Sae2 to associate with and modify PCNA via sumoylation.

To test the role of Trex1 interacting proteins in PCNA sumoylation, we overexpressed Pias1 and Sae2 and looked for PCNA size change that indicates sumoylation by western blot. Additionally, we tested for the addition of a HA-tagged SUMO-1 to PCNA upon overexpression of Pias1 and Sae2. We also performed experiments to determine if Trex1 interacts indirectly with PCNA by overexpressing Trex1 with Chaf1a or Rfc2, immunoprecipitating Trex1 and blotting for PCNA. However, there was no difference in PCNA sumoylation or Trex1 interaction with PCNA under the conditions tested (data not shown). However, we did not exhaustively test the role of Trex1, Pias1, Sae2, Chaf1a and Rfc2 in PCNA sumoylation. Overexpression of Trex1 interacting proteins may not induce a potent enough response to reveal a deficiency in PCNA sumoylation or interaction of PCNA with Trex1. siRNA mediated knockdown of Pias1 and/or Sae2 and looking in Trex1-deficient cells could potentially reveal a defect in PCNA sumoylation. Immunofluorescence of Trex1 and PCNA with Pias1, Sae2, Chaf1a or Rfc2 could also address whether these proteins co-localize. Additionally, the conditions where Trex1-dependent PCNA sumoylation may occur and the overall purpose of such a modification are also unclear. PCNA could be differentially sumoylated during DNA virus infections, retrotransposition events or upon simulating a viral infection with treatment of ISD ligand or IFN-β. Another distinct

possibility is that Trex1 interacting proteins are involved in sumoylating nuclear proteins other than PCNA.

Another candidate nuclear protein for Trex1-mediated sumoylation is promyelocytic leukemia (PML) protein. PML was first identified in chromosomal translocations that fused it to the retinoic acid receptor alpha gene to cause acute promyelocytic leukemia [219]. PML forms SUMO-1 dependent nuclear bodies under cellular stress [220-226]. Several DNA viruses including Epstein-Barr, varicella-zoster and cytomegalovirus all associate with and disrupt the formation of PML nuclear bodies [227-232]. However, the precise function of PML in viral infections is still not understood. Since PML is sumoylated during viral infections and Trex1 interacting proteins Pias1 and Sae2 are capable of sumoylating nuclear proteins, we hypothesized that Pias1 and Sae2-dependent sumoylation of PML may be important for either ISD sensing or inhibiting retrotransposition events.

We tested PML-deficient MEFs for a defect in type I IFN activation after DNA ligand stimulation. However, we did not observe a significant difference between PML knockouts and the wild type control indicating that PML is not part of the ISD pathway (data not shown). In addition, immunofluorescence of PML with overexpressed Trex1, Pias1 and Sae2 did not show co-localization even with type I IFN treatment (data not shown). Trex1 interacting protein-mediated PML sumoylation may be involved in inhibition of retrotransposition events, but we did not exhaustively tested this hypothesis. PML sumoylation by western size shift or addition of HA-tagged SUMO-1 upon overexpression or knockdown of Pias1, Sae2 and Trex1 could be tested to determine if PML sumoylation is dependent on Pias1 and/or Sae2. Additionally, Pias1 and/or Sae2 siRNA knockdowns could be tested for differential PML sumoylation patterns. Since PML does not play an essential role in ISD sensing, as discussed above with PCNA, the

conditions in which Trex1 interacting proteins might sumoylate PML are not well understood. Instead of ISD sensing, PML sumoylation could be important for retrotransposition events. Experiments to address the genomic integration capabilities of long terminal repeat (LTR) and intracisternal A-particle (IAP) transposons in PML knockout cells as described in [159] could determine if PML plays a role in retrotransposition events. But despite being an attractive candidate for mediating anti-viral and -retrotransposition responses in the nucleus, PML may actually not be involved in either of these responses in a Trex1-dependent manner.

Despite confirming the interactions of Brd7, Btf3l4, Chaf1a, Pias1, Rfc2 and Sae2 with Trex1 by immunoprecipitation and immunofluorescence, we were unable to identify the specific functions of these Trex1 interacting proteins in sumoylation of PCNA or PML. However, we did not specifically address the role of each of the Trex1 interacting protein in ISD sensing or inhibition of retrotransposition. To definitively test this, siRNA retroviral knockdowns of each Trex1 interacting protein could be developed and tested for deficiencies in type I IFN induction after DNA ligand treatment and for increased ability of LTR and IAP transposons to integrate in the retrotransposition assay developed by *Stetson et al.* [159]. This approach is possible for 4 out of the 6 Trex1 interacting proteins. The role of Rfc2 and Chaf1a in ISD sensing and retrotransposition is more difficult to determine since are both required for normal DNA replication genomic stability and knockdown would presumably cause cytotoxicity.

In closing, it is important to consider that the Trex1 interacting protein approach assumed Trex1 would interact with members of the ISD or anti-retrotransposition pathways. It is possible that Trex1 does not recruit or interact with ISD sensor or adaptor proteins and its localization to the ER and INM places Trex1 in an optimal cellular location to metabolize ISD substrates. Therefore, the novel Trex1 interacting proteins we identified IFITM3, Brd7, Btf3l4, Chaf1a,

Pias1, Rfc2 and Sae2 may be involved in normal DNA metabolism and not directly part of the anti-viral or anti-transposon responses. However, due to the strong evidence that Trex1 is a negative regulator of the ISD pathway and essential metabolizer of endogenous retroelements that upon accumulation cause autoimmunity, it is likely that these Trex1 interacting proteins are involved in an unknown component of either the ISD pathway or response against retroelements.

Section 2 Figures

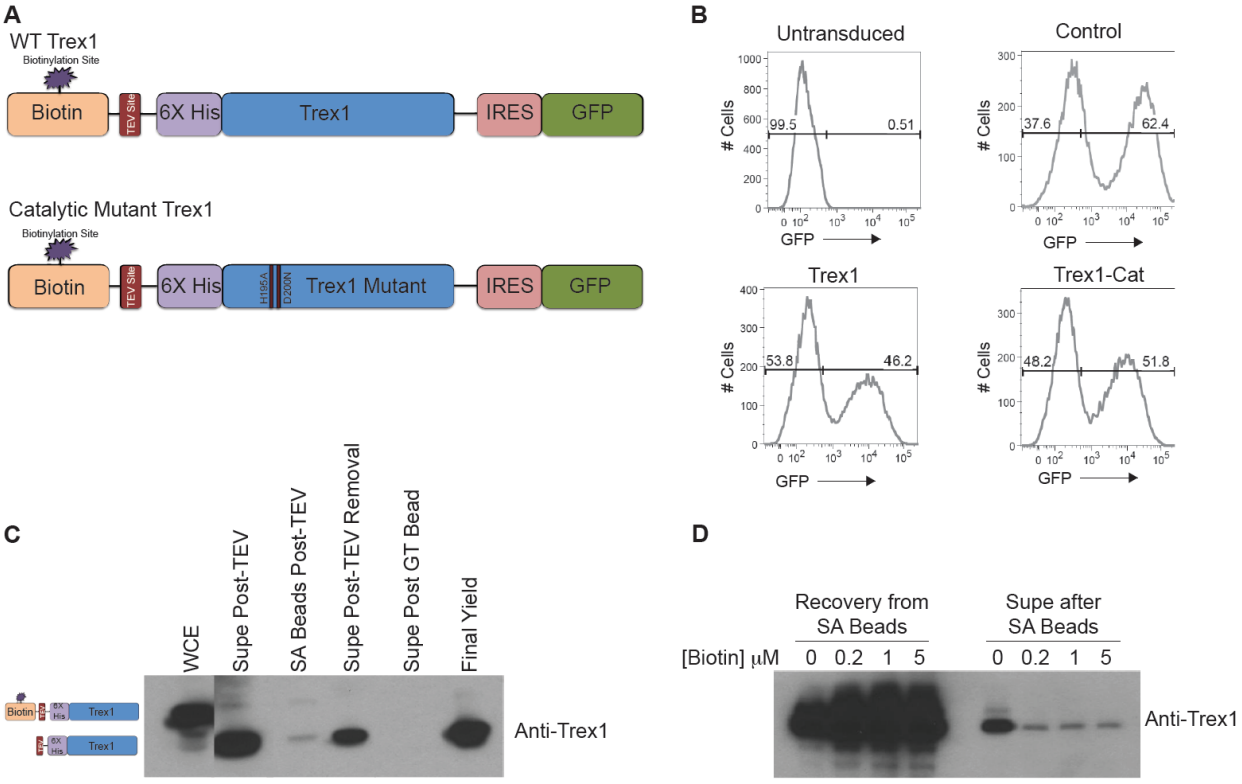


Figure 2.1 Tandem affinity purification method of BioEase 6xHis Trex1 constructs. **(A)** Wild type and catalytically inactive Trex1 mutant proteins were cloned into an MSCV vector containing an N-terminal tag composed of a biotinylation site and a repeating 6X histidine sequence separated by a TEV protease cleavage site. (**B)** Reconstitution of Trex1 KO MEFs. Trex1 KO mouse embryonic fibroblasts were retrovirally reconstituted with MSCV-IRES-GFP constructs containing no Trex1 protein (Empty MSCV), murine Trex1 (Trex1 MSCV), or the catalytically inactive mutant of murine Trex1 (Trex1- Cat MSCV). Reconstitution efficiency was evaluated by FACS analysis of the percent of cells expressing GFP as compared to non-transduced controls. **(C)** TAP-Tagged Trex1 Purification. Trex1 from whole cell extracts prepared from TAP-tagged Trex1 reconstituted MEFs are first purified using Streptavidin (SA) beads. SA purification is followed by an overnight TEV protease cleavage step. TEV is removed using Glutathione Beads before 6X histidine purification by Talon Beads. Final Trex1 protein is eluted from the Talon beads using 200mM imidazole. (**D)** Trex1 protein recovery increases on biotin addition. Increasing amounts of Biotin were added to the media of TAP- tagged Trex1 reconstituted MEFs 24 hours prior to harvest. Even at the lowest concentration of biotin, TAP-Tagged Trex1 recovery is increased compared to no biotin controls.

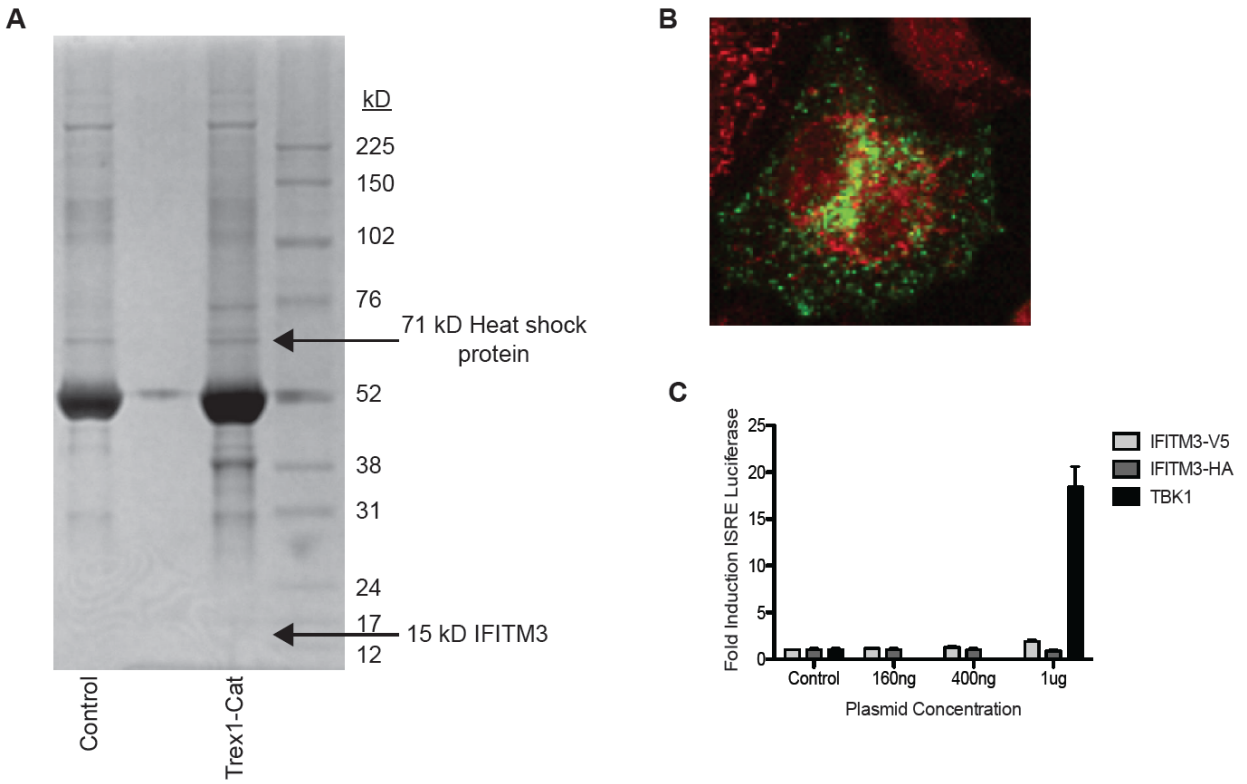
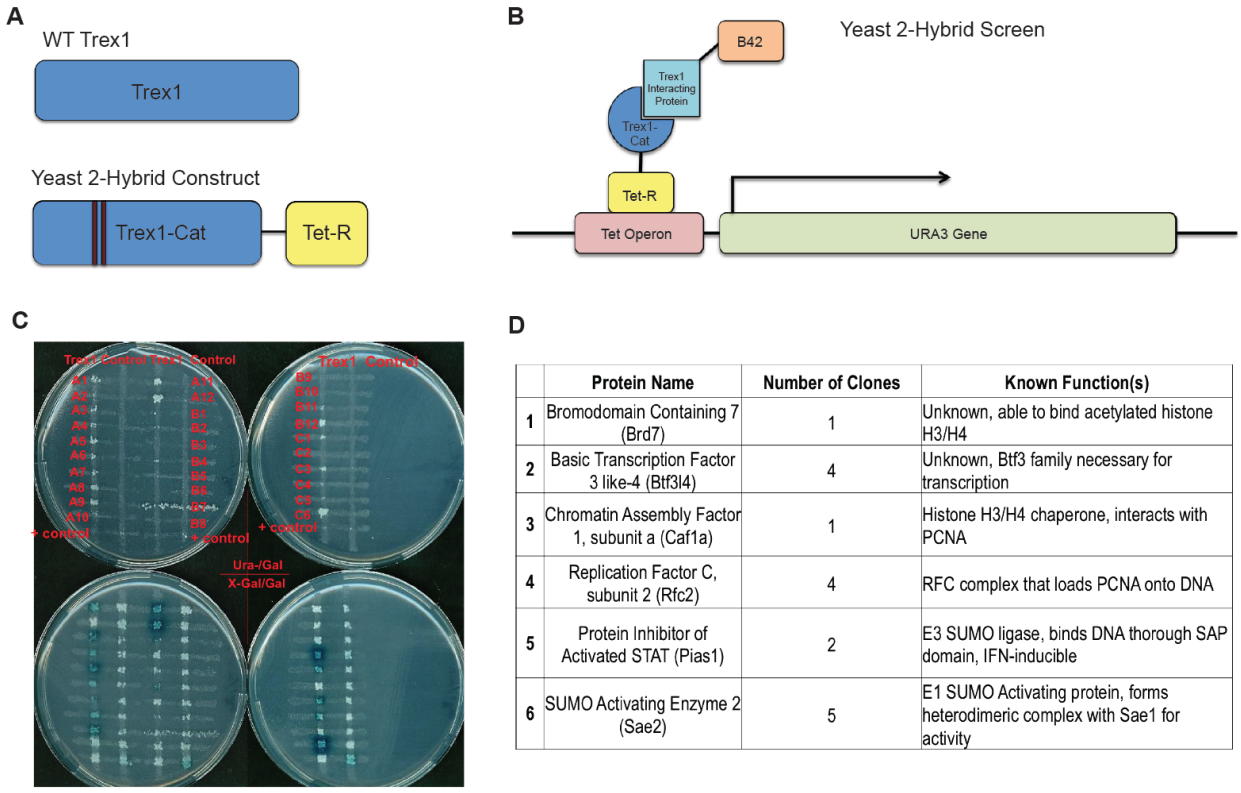


Figure 2.2 Identification of IFITM3 as a Trex1 interacting protein by tandem affinity purification. (A) Trex1 deficient MEFs reconstituted with Empty BioEase 6xHis MSCV (control) and Trex1 catalytic mutant BioEase 6xHis MSCV (Trex1-Cat) were used for tandem affinity purification and isolated Trex1 and associated proteins were size separated on a denaturing SDS gel and stained with coomassie blue. Two distinct bands were seen in the Trex1-Cat as compared to control, excised and sent for mass spectrometry analysis. Two unique proteins were identified as IFITM3 at 15 kD and heat shock protein at 71 kD. **(B)** Immunofluorescence of IFITM3-HA (green) and calnexin (red) did not show localization of IFITM3 to the endoplasmic reticulum (ER). IFITM3 appears instead in vesicular compartments. **(C)** ISRE-luciferase reporter assay with IFITM3-HA and IFITM3-V5 constructs did not show specific induction of ISRE-luciferase as compared to TBK1 control.



	Protein Name	Number of Clones	Known Function(s)
1	Bromodomain Containing 7 (Brd7)	1	Unknown, able to bind acetylated histone H3/H4
2	Basic Transcription Factor 3 like-4 (Btf3l4)	4	Unknown, Btf3 family necessary for transcription
3	Chromatin Assembly Factor 1, subunit a (Caf1a)	1	Histone H3/H4 chaperone, interacts with PCNA
4	Replication Factor C, subunit 2 (Rfc2)	4	RFC complex that loads PCNA onto DNA
5	Protein Inhibitor of Activated STAT (Pias1)	2	E3 SUMO ligase, binds DNA thorough SAP domain, IFN-inducible
6	SUMO Activating Enzyme 2 (Sae2)	5	E1 SUMO Activating protein, forms heterodimeric complex with Sae1 for activity

Figure 2.3 Identification of 6 novel Trex1 interacting proteins by yeast 2-hybrid screen. **(A)** Schematic of wild type Trex1 as compared to the 79 amino acid C-terminal truncated Trex1 catalytic mutant (H195A and D200N) Tet repressor (Tet-R) fusion protein used as bait in the yeast 2-hybrid screen. **(B)** Yeast 2-hybrid screen by Protein Links. The URA3 under the control of the Tet operon is repressed by Trex1-Tet-R bait without an interacting partner. A cDNA library generated from mouse spleen identified Trex1 interacting proteins by allowing growth in uracil-free media with the close proximity of the B42 transcriptional activator of the Tet operon. **(C)** Validation of 17 Trex1 interacting clones by growth on uracil-free media and LacZ production. **(D)** Mapping of yeast 2-hybrid clones revealed 6 novel Trex1 interacting proteins Brd7, Btf3l4, Chaf1a, Rfc2, Pias1 and Sae2.

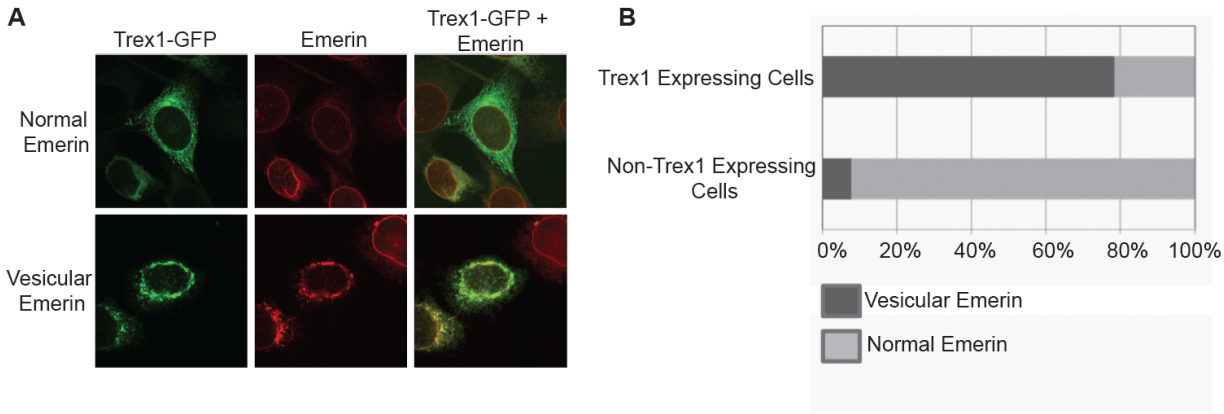


Figure 2.4 Inner nuclear membrane localization of Trex1. (A) Overexpression of Trex1-eGFP (green) and endogenous emerin (red) show co-localization of Trex1 and emerin on the inner nuclear membrane (INM). Two distinct staining patterns were observed with Trex1 overexpression. In some cases Trex1 stained the INM and endoplasmic reticulum (ER) and co-localized with emerin only in the INM (Normal). Alternatively, Trex1 overexpression disrupted the INM and Trex1 and emerin co-localized to vesicular structures (Vesicular). **(B)** Staining pattern in (A) quantified to illustrate overexpression of Trex1 causes disruption of the INM and vesicular emerin localization. Trex1 expressing cells n= 101, Non-Trex1 expressing cells n= 118.

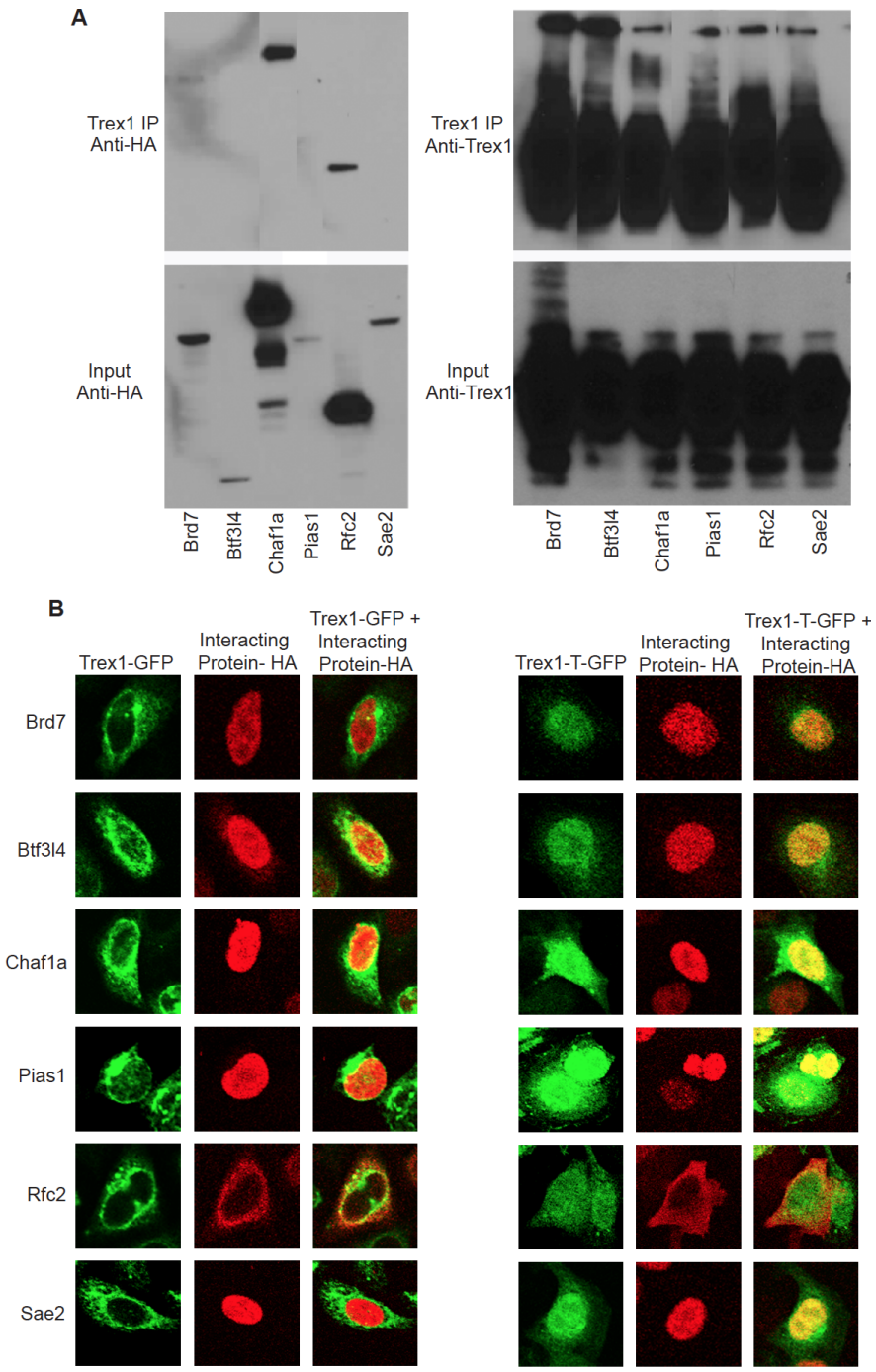


Figure 2.5 Confirmation of yeast 2-hybrid Trex1 interacting proteins. (A) Co-immunoprecipitation utilizing streptavidin Dynabeads of wild type BioEase 6xHis Trex1 with HA-tagged Trex1 interacting proteins. Brd7, Chaf1a and Rfc2 immunoprecipitated with Trex1. **(B)** Immunofluorescence of wild type Trex1-eGFP (green) with HA-tagged interacting proteins (red) shows partial co-localization with all Trex1 interacting proteins. C-terminal truncated Trex1-eGFP (Trex1-T-GFP) altered the normal ER/INM localization of Trex1, but showed more co-localization with Trex1 interacting proteins.

Section 3: Characterization of the Interferon Stimulatory DNA Pathway

Introduction

As previously discussed, innate and adaptive antiviral responses are initiated by detection of nucleic acids by two distinct mechanisms. Toll-like receptor (TLRs) are expressed on the surface and endosomes of innate immune cells such as macrophages and dendritic cells and activate inflammatory cytokine responses including type I IFN that alert the adaptive immune system to infection. TLR3 and 7 are responsible for detecting viral RNA ligands while TLR-9 recognizes CPG-DNA motifs (reviewed in [46]). Importantly, this system utilizes specialized cells to detect nucleic acid ligands made extrinsically in other infected cells.

In addition to this extrinsic pathway of nucleic acid detection initiated by specialized cell types, all mammalian cells are capable of detecting nucleic acids produced intrinsically. Cytosolic nucleic acid receptors detect nucleic acids derived from viral genomes and invasive bacteria in infected cells. All viruses contain a genome made of either RNA or DNA, and innate immune sensors for both of these nucleic acids are broadly expressed and poised to detect infection. The RIG-I and MDA5 RNA helicases are the main intracellular receptors for viral 5' triphosphate and long, double stranded RNA ligands respectively [101, 104, 164, 233]. These receptors signal activation of the type I IFN-mediated antiviral response through the mitochondrial adaptor protein MAVS in a TBK-1 and IRF3-dependent manner [95-98, 234-238].

In addition, detection of intracellular DNA activates at least three signaling pathways (reviewed in [239]). First, the interferon-stimulatory DNA (ISD) pathway is a sequence-independent antiviral response triggered by detection of DNA that activates type I IFN gene expression through the adaptor protein STING [114, 118, 159]. Second, detection of A/T-rich

DNA in the form of the synthetic polymer poly dA:dT activates the innate immune response through both STING as well as through transcription of this DNA by cellular RNA polymerase III (Pol III) into RNA that activates the RIG-I pathway [120-123]. Interestingly, cyclic dinucleotides (bacterial second messenger molecules) trigger type I IFN through activation of STING by direct binding to STING [131, 240]. Finally, intracellular DNA activates an ASC/Caspase-1-dependent signaling complex known as the inflammasome that mediates processing and secretion of interleukin-1β (IL-1β) [241].

Identifying the receptors that trigger these intracellular DNA sensing pathways has been the subject of intense research. In the case of the RNA Pol III/RIG-I pathway and cyclic dinucleotide sensing, the mechanisms of detection are known, although the relevance of these pathways in host defense remains to be established. AIM2 was recently identified as a key activator of the ASC inflammasome [126-129], and AIM2-deficient mice validate its role in IL-1β production in response to numerous viruses and intracellular bacteria [130, 242]. Many receptors for the STING-dependent ISD pathway have been proposed, including DAI/ZBP-1 [132] and DDX41 in dendritic cells [134]. However, DAI-deficient mice have a normal ISD pathway [133], and the role of DDX41 in other cell types is unknown. Recently, the AIM2-like receptors (ALRs) human IFI16/mouse IFI204 were implicated in the STING-dependent IFN response, and IFI16 was also shown to activate the inflammasome in response to Kaposi's Sarcoma Herpes Virus [135, 137]. Some of the differences in results in these studies may reflect the use of different activating ligands such as poly dA:dT which activates not only the ISD pathway, but the RNA PolIII/RIG-I pathway as discussed above and various primary and immortalized cell types which do not respond robustly to intracellular DNA ligands (Fig. 3.5). The lack of straightforward information regarding ISD ligand specificity therefore, necessitated a

clear definition of STING-activating ligands Despite the lack of clarity in defining the ISD sensors, it is becoming clear that the STING-dependent ISD pathway is required for the IFN response to numerous DNA viruses, retroelements and retroviruses, intracellular bacteria, and parasites [118, 120, 159, 183].

The recent identification of ALRs that activate the inflammasome (AIM2, IFI16) and the STING-dependent IFN response (IFI16, IFI204) suggests that this family of receptors may play important roles in both pathways. The ALRs are a family of IFN-inducible genes, and most known ALRs contain a Pyrin domain that mediates protein-protein interactions and a HIN domain that can bind directly to DNA, thus making them ideal candidates for innate immune receptors of DNA [136, 243]. Eight ALRs have been previously annotated in mice, and four in humans, but the functions of most of these ALRs remain unknown.

We took three different experimental approaches to definitively identify the sensors and/or proximal signaling components of the ISD pathway. First, we used microarray analysis to identify differences in gene expression after intracellular DNA and RNA stimulation. We hypothesized that ISD sensor proteins are specifically upregulated in DNA stimulated cells. In a parallel approach to microarray, we initiated a screen of cytosolic DNA binding proteins in an attempt to identify potential sensors by mass spectrometry. Lastly, we characterized the entire ALR family of proteins as a family of potential DNA sensors. Here, we identified microarray profiles of DNA and RNA stimulated cells, discovered three novel cytosolic DNA binding proteins by DNA immunoprecipitation, identified STING-dependent ligands in multiple cell types, discovered five novel ALRs in mice, analyzed the phylogeny of mammalian ALRs, and characterized the functions of all known murine and human ALRs. Our analyses reveal functional redundancy among ISD receptors and extensive evolutionary diversity that is

recreated in mammalian genomes.

Materials and Methods

Cells and mice

C57BL/6 mice were purchased from Jackson Laboratories. Tmem173-/- mice were generated as described in [114], backcrossed to C57BL/6 mice, and screened by genome-wide single nucleotide polymorphism analysis until they were >99.8% C57BL/6. Mavs-/- mice were generated on a C57BL/6 background as described in [158] and Ifnar1-/- mice backcrossed >10 times to C57BL/6 were kindly provided by Murali-Krishna Kaja (Emory University). MyD88-/-Ifnar1-/- bone marrow was kindly provided by Ruslan Medzhitov (Yale University). Bone marrow derived macrophages (BMDM), bone marrow dendritic cells (BMDCs), and primary, early passage mouse embryonic fibroblasts (MEF) were generated and cultured according to standard techniques. WT controls for all experiments were age-matched C57BL/6 mice. All animals were maintained in accordance with guidelines of the University of Washington Institutional Animal Care and Use Committee. HEK 293T were purchased from Invitrogen and HeLa cells were purchased from ATCC. HEK 293T reporter cells stably expressing ASC, Caspase 1 or ASC/Caspase1 were generated by stable transfection and tested for expression of the relevant proteins.

Cell treatments and analysis

ISD oligonucleotides (sense sequence 5' TACAGATCTACTAGTGATCTATGACTGATCTGT ACATGATCTACA 3') were annealed as described in [113] and RIG-I ligand was in vitro transcribed as previously described by Saito et al. [244]. BMDM and BMDCs were plated at a density of 0.7 x 10^6 and MEF at 0.2 x 10^6 per well in 12 well plates for transfection. For transfections, 5 **μ**g calf thymus genomic DNA (Sigma), 5 **μ**g ISD, 5 **μ**g poly dA:dT (Invitrogen),

2.5 μg cyclic-di-AMP (BioLog) and 1 μg RIG-I ligand were complexed with Lipofectamine 2000 (Invitrogen) at a ratio of 1 μg nucleic acid to 1 μL lipid. 30 μg/mL DMXAA (Sigma) and 10 ng/mL LPS (Sigma) were added directly to the culture media. For quantitative RT-PCR of IFN-β or ALR mRNA, cells were harvested into RNA-Bee (Teltest, Friendswood, TX). RNA was reverse transcribed with Superscript III (Invitrogen), and cDNAs were used for PCR with EVA Green reagents (Biorad) on a Biorad CFX96 Real-Time System. The abundance of each cytokine mRNA was normalized to HPRT expression and compared to untreated cells to calculate the relative induction.

Bioinformatic analysis of mammalian ALRs

All bioinformatic and evolutionary analyses were performed by Janet M. Young and Harmit S. Malik (Fred Hutchinson Cancer Research Center, Seattle, WA). We used queries to the genes that flank the ALR locus in mammals, CADM3 (cell adhesion molecule 3) and SPTA1 (spectrin, alpha, erythrocytic 1) to identify the syntenic region in other mammalian genomes via the UCSC Genome Browser [245]. The species and genome sequence versions we used are as follows: Mus musculus (mouse), mm9, July 2007; Rattus norvegicus (rat), rn4, Nov. 2004; Homo sapiens (human), hg19, Feb. 2009; Equus caballus (horse), equCab2, Sept. 2007; Canis familiaris (dog), canFam2, May 2005; Bos taurus (cow), bosTau6, Nov. 2009; Sus scrofa (pig), susScro2, Nov. 2009. We refined the annotation and description of the olfactory receptor genes using minor modifications to a previously described bioinformatics pipeline [246]. To annotate ALRs in the genomes of species other than human and mouse we considered multiple lines of evidence. We first used amino acid sequences of the human and mouse ALRs we identified as queries in tblastn searches of each genome assembly. We parsed the tblastn output using a custom Bioperl script [247], and displayed the locations of matching sequence regions as a custom track on the

UCSC browser, keeping even weak matches to ALRs. In some cases, scrutiny of the tblastn search results revealed that a stop codon or frameshift interrupted some ALR coding fragments; we annotated such ALRs as pseudogenes (data not shown). We also performed a similarity search using just the Pyrin and HIN domains alone, allowing us to more easily visualize domain organization. Pyrin domain matches were often found very close in the genome to matches to both exons of the HIN domain. Indeed, if we use the criteria of identifying all Pyrin domains found within 1 MB of HIN domains, we did not find any putative ALR genes outside the ALR locus. However, within the syntenic ALR loci, we also considered pyrin-domain or HIN-domain matches alone as candidate genes (see main text). We next examined the coordinates of mapped expressed sequence tags (ESTs), experimentally determined mRNAs and known genes. We used those mapped sequences to define ALR intron-exon boundaries, considering any sequences that overlapped the tblastn matches described above. We next compared the results of our curated analyses to the "ab initio" gene predictions from gene-finding algorithms like Genscan reported in the Ensembl database. We performed blastp (protein-protein) searches of the Ensembl dataset using the full-length human and mouse ALRs as queries. We integrated information from the three sources described above (tBLASTN searches, EST analysis, Ensembl predictions) to make a combined ALR gene prediction. EST or mRNA sequences provided the most high-confidence delineation of intron-exon boundaries: for many ALRs, many ESTs were available covering most or all of the gene. Most predicted ALRs were typical of the human and mouse ALRs we characterized in detail experimentally, containing 6-8 coding exons whose predicted protein sequence contains both a pyrin and a HIN domain. We compared our EST-based predictions with Ensembl predictions where they overlapped and found that some ALRs appear well-predicted by ENSEMBL's pipeline. However, some of ENSEMBL's predictions include a large

number of tiny exons, which is likely a misprediction of exon/intron boundaries. We therefore only used Ensembl's gene predictions for regions of ALR genes where EST/mRNA information was partially or entirely missing. One caveat of our analyses is that although the mouse and human genome assemblies are of very high quality in the ALR locus, the other assemblies are still in draft form and contain some gaps in the region. We would expect most ALRs to be at least partially represented in the draft assemblies, but very recently duplicated genes are unlikely to be resolved and might be represented by just a single sequence. Following up on our initial observations that full length AIM2 was missing in some genomes, we performed additional analyses. To find direct evidence of an AIM2 pseudogene, we used the nucleotide sequence of the AIM2 ortholog from horse (the most closely related species that contains a functional AIM2) as query in a blastn search of the cow, dog and pig genome assemblies described above, as well as in searches of lower coverage genome assemblies of dolphin (Tursiops truncatus, turTru1, 2.6x coverage), sheep (Ovis aries, oviAri1, Feb. 2010, ~3x coverage) and cat (Felis catus, felCat4, Dec. 2008, ~2.8x coverage). We examined all hits that clearly matched functional AIM2 sequences better than any other ALR and thus comprise partial AIM2 orthologs. We found cow, sheep and dolphin sequences that match the second, third, and fourth coding exons of AIM2, dog and cat sequences that match the third coding exon of AIM2 (the first part of the HIN domain), and a pig sequence that matches the second coding exon of AIM2 (between the pyrin and HIN domains). We did not find full-length copies of AIM2 in any of these genomes. Blastx searches revealed that each species' AIM2-like sequences contained frameshifts and/or premature stop codons. These results together indicate AIM2 is a degenerate, unprocessed pseudogene in cow, dog and pig (and related genomes). We used TimeTree (www.timetree.org) to estimate divergence times between mammalian species and to date the common events that led to

pseudogenization of AIM2 in some species. Given that large portions of the AIM2 gene appear to have been deleted in each of these genomes, it is not possible to determine whether the two lineages lost AIM2 function independently, or whether their common ancestor had already experienced AIM2 inactivation.

DNA immunoprecipitations

Wild type C57BL/6 bone marrow macrophages were plated at a density of 2.0 x 10^6 cells per 15-cm plate and treated with either 0 U or 25 U recombinant murine IFN-β (R&D Systems) for 24 hours prior to harvest. Five plates were pooled, washed and lysed in standard lysis buffer (50 mM HEPES pH 7.4, 100 mM NaCl, 0.1% Triton X-100, 5% Glycerol, 1 mM DTT and 1X Protease Inhibitor Cocktail (Roche)), EDTA lysis buffer (50 mM HEPES pH 7.4, 100 mM NaCl, 0.1% Triton X-100, 5% Glycerol, 2 mM ETDA, 1 mM DTT and 1X Protease Inhibitor Cocktail) or magnesium and zinc containing lysis buffer (50 mM HEPES pH 7.4, 100 mM NaCl, 0.1% Triton X-100, 5% Glycerol, 2 mM magnesium chloride, 0.2 mM zinc acetate, 1 mM DTT and 1X Protease Inhibitor Cocktail). Reg-80-S/Reg-80-AS and Bio-80-pThio-S/Reg-80 -AS DNA was pre-crosslinked to streptavidin Dynabeads (Invitrogen) for 30 minutes before addition to lysates for 16 hour incubation at 4°C. Samples were washed 5 times in standard lysis buffer before elution of DNA binding proteins from streptavidin beads by incubation at 100°C with sample loading buffer (25mM Tris pH 6.8, 2% SDS, 10% glycerol, 0.012% bromophenol blue and 0.2% β-mercaptoethanol). DNA binding protein lysates were loaded into denaturing SDS gels (Invitrogen) and evaluated by either fixing and staining in 0.1% silver nitrate or GelCode Blue Coomassie (Thermo Scientific) according to manufacturer's instructions. Bands that appeared exclusively in the IFN-β treated macrophages were excised and sent to the Stanford University Mass Spectrometry Core (Palo Alto, CA) for identification.

Immunofluorescence

ALR cDNAs were cloned into pCDNA.3 with a C-terminal hemagglutinin-tag (HA-tag). The open reading frame of murine STING was cloned into pCDNA.3 with a C-terminal FLAG-tag, and murine ASC was subcloned from pCDNA.3 into pEGFP-N1 to create a plasmid expressing an ASC-eGFP fusion protein. 3 x 10^4 HeLa cells plated on glass coverslips in 24 well plates were transfected with 500 ng HA-tagged ALR alone, or together with 500 ng STINGFLAG and/or 500 ng ASC-EGFP. 24 hours later, cells were fixed with 2% paraformaldehyde in PBS for 15 minutes, permeablized for 10 minutes with 0.1% Triton-X 100 PBS, and blocked in 2% FCS PBS for 1 hour at RT. Cells were stained with anti-HA-Tag (6E2) mouse monoclonal antibody (Cell Signaling; 1:50) and anti-DDDDK FLAG rabbit polyclonal antibody (Abcam; 1:200) for 1 hour, washed, and stained with anti-mouse Alexa Fluor 568 (Invitrogen; 1:200), anti-rabbit Alexa Fluor 647 (Invitrogen; 1:200) and then directly conjugated rabbit anti-GFP AlexaFluor 488 (Invitrogen; 1:200) for 1 hour at RT. Coverslips were mounted with SlowFade Gold with DAPI (Invitrogen) and visualized by microscopy with a Nikon A1RSi Scanning Laser Microscope (Nikon). Images were acquired with a Plan Apo VC 60x Oil DIC N2 objective (Nikon) in the 405, 488, 561 and 638 dichroic channels with NIS- Elements Software (Nikon) and pseudocolored with Fiji open source software.

Luciferase and IL-1β reporter assays

1.3 x 10^5 HEK 293T cells in 24 well plates were transfected with 25 ng interferon stimulated response element (ISRE)-luciferase reporter plasmid (Clontech) with 0 ng, 100 ng, 300 ng or 1 **μ**g of pCDNA.3 expressing HA-tagged ALRs, alone or together with 300 ng STING-FLAG expression vector, and then incubated for 24 h. Cells were lysed and luciferase activity was assessed using the Luciferase Reporter Assay System (Promega) according to the manufacturer's

instructions and read using a Centro LB 960 Luminometer (Berthold Technologies). Inflammasome reporter cells were generated by single-cell cloning of HEK 293T cells stably expressing murine ASCFLAG and murine caspase-1. Cells were plated at a density of 1.3 x 10^5 in 24 well plates and transfected for 48 h with 0 ng, 100 ng, 300 ng or 1 μg of HA-tagged ALR expression vectors with 100 ng of pCDNA3.1-pro-IL-1β expression vector. The supernatant was harvested and assayed for IL-1β secretion using a mouse IL-1β ELISA Set (Becton-Dickinson) according to the manufacturer's instructions. ELISA plates were read using an iMark Microplate Reader (BioRad).

Phylogenetic analysis of mammalian ALRs

We separately aligned sequences of each pyrin domain and each HIN domain (sequences outside those domains did not align well) and made phylogenetic trees. For this, we first produced amino acid alignments by obtaining seed alignments of HIN (PF02760) and Pyrin (PF02758) domains from PFAM [248]. We edited PFAM's pyrin alignment to retain sequences from the ALR locus and only one outgroup sequence (NAL12). We used hmmbuild from the HMMER package (http://hmmer.janelia.org, version 3.0) to make a hidden Markov model (HMM) representing each alignment. We then used HMMER's hmmsearch algorithm to identify those domains in each of the predicted ALR protein sequences, specifying the -A option to automatically output domain sequences aligned to the HMM. This amino acid alignment was then used to generate a corresponding nucleotide alignment that retained codon information. This nucleotide alignment was used to generate a PhyML bootstrap tree (www.phylogeny.fr) using either a HKY85 or GTR (general time reversible) model of nucleotide substitutions. Bootstrap values were calculated as a percentage of 1000 trials. Phylogenies were visualized and formatted using the Dendroscope software [249]. All nodes with bootstrap values of less than 50% were collapsed for ease of

presentation.

Microarray

MyD88-/- Ifnar1-/- bone marrow-derived macrophages were plated at a density of 0.75 x 10^6 before no treatment or transfection with 3 μg ISD or 1 μg RIG-I ligand for 4, 8 and 12 hours. Cells were harvested into RNA-Bee (Teltest, Friendswood, TX) for phenol:chloroform extraction followed by RNeasy clean up (Qiagen) of total cellular RNA. 3 μg of each RNA sample was sent to Yale University Keck Biotechnology Resource Laboratory (New Haven, CT) for Affymetrix microarray analysis. Data was normalized by eliminating samples with p values < 0.01 and analyzed by computing fold induction of ISD and RIG-I ligand treatments over untreated and then comparing ISD and RIG-I ligand fold induction values.

Retroviral and lentiviral siRNA knockdowns

IFIT1, IFIT2 and control siRNA contructs were cloned into the pMSCV-LMP plasmid. For IFIT1 and IFIT2 knockdown in wild type and Ifnra1-/- bone marrow derived macrophages, murine leukemia virus (MLV) pseudotyped retrovirus was generated by transfection of 2.5 x 10^6 HEK 293T cells in 10-cm plates with 10μg IFIT1 or IFIT2 pMSCV-LMP knockdown contruct and 10 μg of pCL-Eco for 48 hours. 4.0 x 10^6 bone marrow-derived macrophages were transduced with HEK 293T supernatants for 48 hours and selected to purity for 3 days in 5 μg/mL puromycin (Invitrogen) before plating for ligand treatments.. Transduction efficiency was assayed by resuspending 0.1 x 10^6 macrophages in 1% fetal calf serum (HyClone) PBS for flow cytometry on a FACScan 2 (Becton Dickinson) for GFP expression. Knockdown efficiency was evaluated in whole macrophage cell extracts treated with CT DNA or left untreated for 4 hours by western blotting with anti-IFIT1 goat polyclonal antibody (Santa Cruz; 1:500) and anti-IFIT2 rabbit polyclonal antibody (Abcam; 1:1,000) according to standard techniques.

IFI204 and control siRNA constructs were designed and cloned into the pLKO.puro plasmid. IFI204 and MNDA are targeted by siRNA #1 sequence (sense): 5'CCAAGAGCAATACACCACGAT3'. siRNA #2 exclusively targets IFI204 sequence: 5'GCTAAGGAAGAAGATCACCAT3'. The control siRNA targets eGFP: CAACAAGATGAAGAGCACCAA. Knockdowns were validated by transfecting 2 x 10^5 HEK 293T cells in 12 well plates with 500 ng-1 **μ**g HA-tagged IFI204 expression vector along with 1-2 **μ**g of either control or IFI204-specific pLKO.1 vector. 24 hours later, IF204 knockdown in whole cell extracts was evaluated by western blotting with anti-HA-Tag (6E2) mouse monoclonal antibody (Cell Signaling; 1:2,500) according to standard techniques. For IFI204 knockdown in bone marrow derived macrophages (BMDM) and mouse embryonic fibroblast (MEF), VSV pseudotyped lentivirus for stable transduction was produced by transfecting 2.5 x 10^6 HEK 293T cells in 10-cm plates with 10 **μ**g of pLKO siRNA knockdown construct, 9 **μ**g psPAX-2, and 1 **μ**g pVSV-G for 48 h. 4 x 10^6 BMDM and 2.5 x 10^6 MEF were transduced with HEK 293T viral supernatants, selected with 5 **μ**g/mL puromycin (Invitrogen) for 3 days, and plated for ligand treatments.

Results

Microarray profiles of DNA and RNA ligand stimulations

In order to identify the DNA specific components of the ISD pathway, we compared the gene induction profiles after DNA and RNA stimulation of cells. We reasoned that the sensor(s) and adaptor protein(s) for the ISD pathway would be specifically upregulated in DNA, but not RNA stimulated cells. In order to overcome the contribution of genes induced by TLR3, 7 and 9

after recognition of RNA and DNA ligands respectively and autocrine type IFN signaling, we used primary bone marrow derived macrophages from Myd88 and Ifnra1 double deficient mice.

Surprisingly, we found that the gene induction profiles of RIG-I ligand and ISD stimulated macrophages were nearly identical illustrated by the clustering of genes near the origin of the x- and y- axis when graphed (Fig. 3.1A). IFN-β was the most highly expressed transcript in both stimulations, and of the top 25 genes induced, only 3 differed between DNA and RNA treatments (Fig. 3.1B). When we compared the fold induction ISD over RIG-I ligand we found that 19 histone variants were more highly induced in the ISD stimulated group (Fig. 3.1C). However, the overall signal of these histones during ISD stimulation was low and ranged from 2-30 as compared to the signals of the strongest induced genes at or over 3,000 (Fig. 3.1B,C and data not shown). The finding that a large variety of histone proteins are specifically upregulated in ISD and not RIG-I ligand treated cells is interesting because histones are known to associate with DNA virus genomes (reviewed in [250]). However, there is currently no evidence suggesting histones are involved in the ISD pathway. Since no appreciably differences in gene induction between ISD and RIG-I ligand stimulation were observed by microarray other than histones, it is likely that the sensor/adaptor proteins involved in the ISD pathway are already present and poised to sense infecting DNA viruses under normal conditions, and that stimulation with various ligands does not dramatically upregulate their expression. Without specific differences in gene induction in ISD versus RIG-I ligand stimulation to indicate potential contributors to the ISD pathway, we next developed a system to directly isolate and identify DNA binding proteins.

DNA immunoprecipitation identifies novel DNA binding proteins

In a complimentary approach to the microarray, we used cytosolic DNA pulldowns to

further identify the proximal components of the ISD pathway. However, there are several inherent challenges we considered in searching for the sensor/adaptor proteins of the ISD pathway. First, previous research in the field focused on the response to the synthetic DNA polymer Poly dA:dT which is not the same as the response to natural DNA ligands. Poly dA:dT can simultaneously active the ISD pathway and the RNA sensing pathway via transcription by RNA Polymerase III into a RIG-I ligand (discussed later in Fig. 3.5B,C and [121-123]). Cell lines are also commonly used to screen for the ISD sensor(s), but we discovered that transformed cells are unable to robustly respond to transfected DNA (Fig. 3.2). Therefore, it is important to use primary cells to screen for the ISD components. With these challenges in mind, we designed a method that utilizes biotinylated DNA oligos added directly to primary macrophage cell lysates and immunoprecipitation with streptavidin beads to identify potential ISD sensors.

Initial experiments using regular phosphodiester DNA oligos and standard lysis buffer did not recover many cytosolic DNA binding proteins (Fig. 3.3A, lane 2). We found an increase in the number of DNA interacting proteins recovered from lysates of primary macrophages using a synthetic form of double stranded DNA that contains one strand of regular phosphodiester DNA and a strand of phosphorothioate DNA (Fig. 3.3A, lanes 3-10). Interestingly, this synthetic phosphodiester/phosphothiol DNA also enhanced IFN-β induction upon transfection as compared to regular phosphodiester double stranded DNA (Fig. 3.3B). We further optimized DNA pull down conditions in primary cells and found that lysis buffer containing EDTA, magnesium, zinc, or both magnesium and zinc combination changed the bands we observed by silver staining (Fig. 3.3A, lanes 3-10). We reasoned that since many DNA binding proteins are dependent on divalent cations for their structure and/or function that including magnesium and zinc and excluding EDTA in our lysis buffer would allow us to identify novel DNA binding

proteins. We also optimized the recovery of cytosolic DNA binding proteins by pre-treating macrophages with IFN-β for 24 hours prior to lysis since other sensor molecules such as RIG-I, MDA5, TLR3, and AIM2 are all IFN-inducible. Using this approach we performed a large scale pulldown of DNA binding proteins in primary macrophages treated with and without IFN-β. Several proteins were enriched in the IFN-β pretreated macrophages and identified by mass spectrometry as Interferon Induced Protein with Tetratricopeptide Repeats 1 (IFIT1) and 3'-5'Oligoadenylate Synthetase 3 (OAS3) (Fig. 3.3C).

OAS3 is a member of the 3'-5' Oligoadenylate Synthetase family of type I IFN inducible proteins that includes OAS1, OAS2 and OAS3 [251-256]. OAS proteins are activated by direct binding of RNA virus genomes and subsequently direct the catalysis of ATP into 2-5A molecules, which activate the endoribonuclease RNaseL that in turn is thought to degrade viral RNA [257-260]. The 2-5A products generated from OAS family proteins have been shown to protect cells from a variety of picornavirus and retrovirus infections [261-264]. IFIT1 however, does not have a well-described role in viral infections. IFIT1 (also known as ISG56) was identified as one of the first type I IFN inducible genes [265, 266]. IFIT1 is part of a family of four genes in humans that includes IFIT1 (ISG56), IFIT2 (ISG54), IFIT3 (ISG60) and IFIT5 (ISG58) and only IFIT1 (ISG56), IFIT2 (ISG54) and IFIT3 (ISG49) in mice [253, 267-274]. Since their discovery the IFIT family has been implicated in a myriad of anti-viral functions. IFIT1 and IFIT2 have been shown to inhibit global cellular translation initiation by binding host Eukaryotic initiation factor 3 (EIF3) to inhibit viral replication, and IFIT1 directly inhibits Human Papilloma Virus (HPV) replication by binding virally-encoded E1 helicase [275-280]. Interestingly, IFIT1 was also shown to negatively regulate type I IFN in response to Sendai virus infection by binding STING [281]. Since IFIT1 does not have a single well-defined anti-viral

activity, we developed a system to evaluate the role of IFIT1 and related family member IFIT2 in ISD sensing.

Evaluation of IFIT1 and IFIT2 as ISD sensors

With the identification of candidate IFIT proteins, we next developed a small interfering RNA (siRNA) knockdown system to determine their relevance to the ISD pathway. We first designed siRNA target sequences to target IFIT1 and IFIT2 mRNA and cloned them into a LMP retroviral vector to express siRNAs to target messenger RNAs in IFIT1 and IFIT2, a puromycin selection marker and GFP and optimized the transduction conditions in an immortalized MEF cell line. We next transferred this LMP siRNA knockdown system to primary macrophages. To alleviate the IFN-inducible nature of IFIT1 and IFIT2 after ISD stimulation, we used Ifnrα1-deficient bone marrow derived macrophages. After transduction and puromycin selection, over 97% of Ifnrα1 deficient macrophages expressed the knockdown constructs whereas selection in the wild type macrophages was less efficient (Fig. 3.4A). We also verified knockdown after calf thymus DNA stimulation by western blot and found similarly sufficient knockdown that was resistant to induction by ligand stimulation (Fig. 3.4B). We then stimulated the IFIT1 and IFIT2 knockdown macrophages with a panel of DNA ligands and compared the IFN-β response to calf thymus DNA (CT), interferon stimulatory DNA (ISD) and in vitro transcribed RIG-I ligand by quantitative real-time PCR. However, we did not observe a significant difference in the IFN-β response to CT or ISD DNA in IFIT1 or IFIT2 knockdown macrophages as compared to siRNA knockdown control macrophages (Fig. 3.4C). We also observed a decrease in the IFN-β response to RIG-I ligand after IFIT1 and IFIT2 knockdown, which we did not expect if IFIT1 or IFIT2 was exclusively involved in ISD sensing.

These results suggested that IFIT1 and IFIT2 may positively regulate IFN-β induction

during antiviral responses in a non-ISD specific manner. However, it remains to be determined whether the IFIT family regulates the antiviral IFN-β response through the ISD and/or RIG-I ligand pathways directly, by interacting with sensor(s), adaptor(s) and/or transcription factor(s), or indirectly potentially by regulating gene expression of proteins necessary for type I IFN induction. With these negative results from the DNA pulldown, we next decided to focus on further characterizing the stimulator of IFN genes (STING) protein whose discovery prompted a re-evaluation of ISD specific ligands.

Defining the role of STING and ligand specificity of the ISD pathway

Recently STING was identified as a necessary endoplasmic reticulum adaptor protein component of both the ISD and RIG-I ligand pathways [114, 118]. While not the sensor of interferon stimulatory DNA, STING has been shown to directly bind and activate type I IFN in response to cyclic dinucleotides [131]. Additionally, the innate immunity field utilizes numerous ligands to trigger the DNA-activated antiviral response in various cell types, often interchangeably, but the pathways activated by each ligand have not been clearly defined. We therefore performed a comprehensive analysis of these and additional ligands in primary bone-marrow derived macrophages and dendritic cells, and in murine embryonic fibroblasts (MEF), to compare directly their stimulatory potency and evaluate their requirements for STING and MAVS.

Consistent with previous reports, we found that calf thymus DNA and annealed, 45 base-pair ISD oligos potently activated IFN-β expression in macrophages and MEFs [113, 114]. This antiviral response was reduced by over 99.8% in STING (Tmem173)-deficient cells, demonstrating that these two DNA preparations represent “pure” ISD ligands (Fig. 3.5A). As shown previously shown, cyclic di-AMP triggered an IFN-β response that was also entirely

STING-dependent (Fig. 3.5A) [131, 282]. Additionally, we found that the chemotherapeutic agent and nucleoside analog DMXAA, which activates type I IFN expression, is a STING activating ligand (Fig. 3.5A) [283]. In contrast, the IFN-stimulating activity of the DNA polymer poly dA:dT was largely intact in STING-deficient macrophages and MEFs (Fig. 3.5A), consistent with its ability to trigger both the ISD pathway and the RNA polymerase III/RIG-I pathway [121, 122]. Indeed, the antiviral response to poly dA:dT was reduced by over 99% only in macrophages deficient in both STING and MAVS, and not in macrophages lacking STING alone or MAVS alone (Fig. 3.5B). We obtained identical results in bone marrow-derived dendritic cells (Fig. 3.5C), demonstrating substantial redundancy in these cell types between the ISD-STING pathway and the Pol III/RIG-I pathway in response to poly dA:dT. As expected, an in vitro-transcribed triphosphate RIG-I ligand activated MAVS-dependent IFN production, and LPS-induced IFN required neither STING nor MAVS (Fig. 3.5) [165]. Together, these data provide a comparative analysis of IFN-inducing ligands in macrophages, DCs, and MEFs, and genetically define "pure" activators of the STING-dependent antiviral response. Importantly, we formally demonstrate that ISD and poly dA:dT cannot be used interchangeably as ligands for the DNA-activated antiviral response, and that data implicating specific candidate receptors in the IFN response to poly dA:dT must be interpreted in light of the two pathways triggered by this polymer. Additionally, this data unambiguously identified STING as an essential adaptor protein exclusive to the ISD pathway. With this new knowledge of STING-dependent ligand specificity, we took a directed approach to identify components of the ISD pathway.

Evolutionary diversity among mammalian AIM2-like receptors

We next wanted to re-examine the role of the AIM2-like receptor (ALR) family of proteins in ISD sensing. We initiated our analysis of mammalian ALRs by focusing first on

ALRs that had been previously annotated in mouse (eight genes) and human (four genes). We also performed an in-depth characterization of the ALR gene family in C57Bl/6 mice using homology searches. Our analyses yielded five novel genes, bringing the total number of ALR genes within this mouse locus to thirteen (Fig. 3.6A and Table 3.1). These five newly identified mouse genes are located between the Aim2 and Ifi204 genes, and all contain a Pyrin and/or a HIN domain that is diagnostic of the ALR gene family. We named the genes based on domain organization in an effort to distinguish them from the other currently known murine ALRs. The newly identified genes are: Pyblhin-C, which contains Pyrin and HIN domains; Pyhin-B, which is highly similar to the known Pyhin-1 ORF but has a truncated HIN domain lacking the last 21 amino acids; Pyr-A (Pydc4), which contains a Pyrin domain but no HIN domain; Pyr-rv1 (Pydc3), which contains a Pyrin domain followed by a three-exon in-frame fusion of three distinct sequences from a murine endogenous retrovirus B2 (ERVB2) open reading frame; and Pyhin-A, which contains Pyrin and HIN domains (Fig. 3.6A and Table 3.1). Additionally, we found an orphan HIN domain adjacent to the Pyrin domain of Pyr-rv1 that is most similar to the HIN of Pyblhin-C, but we have not yet confirmed that this HIN domain is expressed or linked to the Pyr-rv1 Pyrin. These five novel genes reveal an even higher complexity within the murine ALR gene family than previously appreciated.

To systematically address the evolutionary underpinnings of the ALR complexity that we have uncovered, we performed genomic characterization of the ALR loci across seven mammalian species that represent approximately 95 million years of evolutionary divergence. In each of these species, we found a single, continuous ALR locus, flanked by the Cadm3 gene on one end and by a number of olfactory receptors and the Spta1 gene on the other end (Fig. 3.6A). Remarkably, the size and makeup of the ALR locus is dramatically different among these

species. The murine ALR locus, with 13 genes distributed across 573 kb, is the most expansive among the species we examined (Fig. 3.6A). In contrast, horses have six ALR genes, and the rat and human ALR loci each contain four genes. Interestingly, dog and pig each have two ALR genes and cows only have a single ALR gene (Fig. 3.6A). Perhaps the most surprising finding of our analysis is that the canonical inflammasome activator, Aim2, is no longer functional in the cow, sheep, dolphin, pig, cat and dog genomes because of the acquisition of deletions, frameshifts and stop codons (Fig. 3.6B,C). The shared frameshift mutation between the cow, sheep and dolphin genomes and the two shared frameshifts between the dog and cat genomes reveal that these two mammalian orders lost their Aim2 gene at least 55 million years ago, although it was present in the last common ancestor of eutherian mammals (Fig. 3.6C).

We used the sequences of these ALR open reading frames to construct separate phylogenetic trees for their Pyrin and HIN domains based on an alignment of their nucleotide sequences. These analyses reveal extensive evolutionary diversity within this gene family across species, as well as a number of specific features that have important implications for the biology of ALRs. First, the phylogenetic trees confirm that mouse, rat, human, and horse each have an orthologous Aim2 gene based on gene location, orientation, and homology throughout both Pyrin and HIN domains, and validate Aim2's absence from dog, pig, and cow (Fig. 3.6 D,E). Second, with the exception of the four-member Aim2 "clade", the Pyrin and HIN trees reveal an almost complete lack of orthology among full-length ALR genes in any pairwise comparison between any mammalian species (Fig. 3.6D,E) including the closely related species rat and mouse. For example, human IFI16 and mouse Ifi204 are considered orthologs because they both contain a single Pyrin domain and two HIN domains [135]. However, the phylogeny reveals that these two genes are not orthologous (each protein domain from IFI16 and IFI204 is more closely related to

other genes). In fact, among the mammalian species we analyzed, there are no other Pyrin-containing ALRs that have two HIN domains, revealing that this domain organization is fortuitous rather than evolutionarily preserved from a common ancestor. Indeed, with the exception of AIM2 and "ALR gene 1" in pig and cow, no ALR is more closely related across its full length to an ALR in another species than it is to an ALR(s) within the same species. The reason for this lack of orthology can be inferred from the numerous species-specific expansions (highlighted in yellow) where protein domains of an ALR are much more closely related to ALRs from the same species, rather than to 'orthologs' from other species. Whether due to independent gene duplications or recurrent gene conversions across the ALR genes, these patterns point to highly rampant species-specific specialization of ALRs.

Third, even within the same gene, the phylogeny of the Pyrin and HIN domains is often not congruent. While in some instances, like in 'ALR gene 1' of pigs and cows, there is clearly a common history for both Pyrin and HIN domains, in most other instances, the Pyrin and HIN domains fall in completely disparate parts of the phylogeny (Fig. 3.6D,E). This strongly suggests that recombination may have led not only to species-specific expansions of ALR genes, but also scrambled the existing genes into novel combinations of Pyrin and HIN domains.

This evolutionary analysis reveals a remarkable plasticity in mammalian ALR genes, with no single ALR gene preserved among all mammals and little orthology preserved across species. Instead, the ALR genes have undergone extensive, species-specific diversification, suggesting the existence of strong evolutionary pressures that have shaped ALR sequences and functions throughout the mammalian lineage. These data highlight the difficulty in applying conventional, sequence-based inference of ALR functions across even closely related mammalian species. For instance, the dramatic difference in the number and sequences of mouse

and rat ALRs is due to two factors: gene expansions in mouse of three ancestral rodent ALRs, and independent reassortment of the Pyrin and HIN domains that create additional diversity. Importantly, the three human ALRs (aside from AIM2) are not represented within these three rodent-specific clusters (Fig. 3.6D,E). This completely revised view of ALR evolution challenges some of the interpretations of previous cross-species comparisons that have been made with the assumption of orthology, and necessitates the application of unbiased assays to test the function of each ALR within defined experimental conditions.

Most murine ALRs are abundantly expressed and all are inducible

We focused on the murine ALR gene family for a more comprehensive analysis, with the hypothesis that characterizing all of these genes might reveal functional correlations to place within the evolutionary perspective characterized in Figure 3.6. We developed and validated quantitative RT-PCR primers that specifically and efficiently amplify each of the 13 murine ALR ORFs (Fig. 3.7 and Table 3.2). We determined the mRNA expression levels of each ALR in unstimulated bone marrow-derived macrophages and MEFs (Fig. 3.8A). We found that 11 of the 13 ALRs are abundantly expressed, the five most abundant being PYHIN-A, MNDAL, PYHIN-1, IFI203, and PYBLHIN-C (Fig. 3.8A). In contrast, we detected very little PYHIN-B or IFI202b mRNA in resting cells (Fig. 3.8A). We determined the contribution of tonic type I IFN receptor signaling to the basal expression levels of each ALR mRNA by comparing wild-type and Ifnar1-/- macrophages. Interestingly, PYHIN-A and AIM2 expression was largely unaffected in Ifnar1-/- macrophages, whereas the basal expression of all other ALRs was reduced by at least 90% in the absence of IFN signaling (Fig. 3.8A). We next quantitated the inducibility of each murine ALR in response to type I IFN treatment, activation of the ISD pathway, or activation of RIG-I. Consistent with findings reported for the eight previously described ALRs, we found that

all thirteen murine ALRs are IFN-inducible genes, as evidenced by increased expression levels six hours after treatment with recombinant IFN-β in WT macrophages but not Ifnar1-/- macrophages (Fig. 3.8B). The level of IFN-inducibility was inversely correlated with basal expression level: the most abundant ALRs were the least inducible, and the least abundant ALRs were the most inducible (Fig. 3.8B). Interestingly, we found that ALR expression is dramatically increased after stimulation with ISD or RIG-I ligands, even in the absence of IFN signaling (Fig. 3.8C). Together, these data provide a direct comparison of expression of all thirteen murine ALRs and reveal two modes of ALR inducibility, one activated by IFNs and one activated by nucleic acid ligands.

IFI204 is not a non-redundant sensor of DNA

Murine IFI204 was previously implicated as an antiviral sensor of DNA, based on proposed homology to human IFI16 and a ~60% decrease in DNA-activated type I IFN production by murine cells in which IFI204 was knocked down [135]. However, our identification of novel ALRs, several of which are abundantly expressed and highly homologous to IFI204 in their Pyrin or HIN domains, along with the lack of true orthology between IFI16 and IFI204, raised the possibility that IFI204 itself might not be a unique DNA sensor. We therefore designed lentiviral siRNAs targeting IFI204, taking into account the high sequence homology between IFI204 and other ALRs. We identified two different siRNAs that robustly reduced expression of an epitope-tagged IFI204 protein (Fig. 3.9A), and verified knockdown of the corresponding endogenous IFI204 mRNAs in primary BMDMs and MEFs (Fig. 3.9B). One of these siRNAs also targeted MNDA, whereas the other was unique for IFI204 and no other murine ALR (Fig. 3.9A,B). We transduced primary BMDMs and MEFs with these lentiviral IFI204 siRNAs or a control lentiviral siRNA targeting GFP, selected the transduced cells to

>99% purity in puromycin, and stimulated them with the entire panel of ISD and non-ISD ligands defined in Figure 1. Four hours after treatment, we measured IFN-β induction by quantitative RT-PCR. We found that IFI204 knockdown did not reproducibly or specifically impair the ISD pathway in either BMDMs or MEFs (Fig. 3.9C). We obtained identical results with stable IFI204 knockdown in Ifnar1-/- BMDMs, which respond robustly to ISD despite reduced basal expression levels of most ALRs. These data suggest that while IFI204 may contribute to the antiviral response, it is not uniquely required in either BMDMs or MEFs for full activation of the ISD pathway.

Murine ALRs have discrete patterns of localization in the presence of STING and ASC

The sequence homology among the Pyrin and HIN domains within the murine ALR family and the evolutionary diversity of ALRs among mammalian species led us to develop unbiased assays to quantitatively compare the localization and function of each ALR. We generated expression vectors encoding haemagglutinin (HA) epitope-tagged versions of all thirteen murine ALRs and examined their localization after transfection into HeLa cells, which do not have a functional ISD pathway. We found that with the exception of AIM2, which is a cytosolic protein [128], the other 12 murine ALRs localized exclusively to the nucleus when expressed alone (Fig. 3.10A, left two columns). We next determined the localization of each ALR upon co-expression with either STING or ASC, which are required for DNA-activated IFN production and inflammasome triggering, respectively. Interestingly, most of the 13 ALRs colocalized with STING, either partially or entirely, when overexpressed with STING alone (Fig. 3.11A). Similarly, all 13 ALRs colocalized with ASC when overexpressed with ASC alone, and again, the extent of colocalization varied from partial to complete (Fig. 3.11B). Thus, all ALRs relocalize when expressed with relevant adaptor proteins, and co-expression of an ALR with a

single adaptor protein reveals indiscriminate colocalization. We therefore expressed each ALR together with both STING and ASC to more closely approximate the expression of all three proteins in macrophages and other cell types. These experiments revealed three predominant localization patterns of murine ALRs that were not apparent when each ALR was expressed alone or with a single adaptor protein. The first pattern, exemplified most clearly by AIM2, MNDA, and MNDAL, was a strong colocalization with puncta of ASC and minimal colocalization with STING (Fig. 3.10A). The second pattern, most evident with PYHIN-B, PYR-A, PYHIN-1, PYHIN-A, IFI204, IFI203, and IFI205, was a predominant colocalization with STING in structures that resemble previously defined STING-positive ER-Golgi compartments [114], together with a concomitant recruitment of ASC to these areas of ALR-STING colocalization. The remaining three ALRs (PYBLHIN-C, PYR-RV1, and IFI202B) localized to ASC puncta with detectable but not complete colocalization of STING. We confirmed the specificity of these interactions for two of the novel ALRs by coimmunoprecipitation in cells expressing STING and ASC. In agreement with the microscopy data, we found that MNDA interacted with ASC but not STING (Fig. 3.10B), and that PyrA interacted with both STING and ASC, with the ASC interaction significantly weaker than the STING interaction (Fig. 3.10C). Together, these data reveal that murine ALRs can be categorized based on their ability to be recruited to ASC, STING, or both.

Novel murine and human ALR activators of STING-IFNs and the ASC-inflammasome

We developed simple assays to evaluate and compare the functional capabilities of all murine and human ALRs. We found that overexpression of each murine and human ALR alone did not activate an interferon-stimulated response element (ISRE)-luciferase reporter in HEK 293T cells (data not shown). We therefore included co-expression of STING at a concentration

that was insufficient to activate the ISRE-luciferase reporter on its own (Fig. 3.12A). Interestingly, we found that several murine ALRs robustly activate STING-dependent IFN production. The strongest activator of STING was the Pyrin-only protein PYR-A, and among the full-length ALRs we observed strongest IFN induction by PYHIN-A and IFI203 (Fig. 3.12A). IFI202b, which has two HINs but no Pyrin domain, also activated a robust STING-dependent IFN response, suggesting that a Pyrin domain is not required for this activity in cis. The remaining ALRs exhibited weak or minimal STING-dependent activity, as exemplified by AIM2, consistent with the intact ISD pathway in AIM2-deficient murine cells [128].

We next generated HEK 293T cells that stably express ASC and Caspase-1 for evaluation of inflammasome activation by murine ALRs. We cotransfected plasmids encoding each ALR with an expression vector for full length murine IL-1β and measured inflammasome activation using an ELISA for secreted IL-1β. ALR-activated IL-1β secretion was undetectable in the parental HEK 293T cells or in cells stably expressing ASC alone or Caspase-1 alone, validating the specificity of this assay (data not shown). As expected, we found that overexpression of murine AIM2 potently activated IL-1β secretion (Fig. 3.12B). Interestingly, we identified two additional murine ALRs – MNDA and MNDAL - as novel and robust activators of the ASC inflammasome (Fig. 3.12B). The remaining ten murine ALRs exhibited minimal or undetectable IL-1β processing activity in this assay. Thus, three murine ALRs are capable of activating the inflammasome.

We plotted the ability of each murine ALR to activate STING-dependent IFN production versus ASC-dependent IL-1β processing, thus creating an activity map of the entire murine ALR family. We found largely nonoverlapping activities of each murine ALR in these assays, such that individual ALRs robustly activate either STING-IFN (PYR-A, PYHIN-A, IFI203, PYHIN-

B, IFI202b), ASC-inflammasome (AIM2, MNDA, MNDAL), or neither (Fig. 3.12C). Importantly, the ability of an individual ALR to activate STING or ASC closely correlated with its colocalization with these adaptors (Fig. 3.10), validating the predictive value of these cell biological and functional assays.

We next performed the same ISRE and IL-1β assays with the four human ALRs and uncovered novel activities. Interestingly, whereas murine AIM2 activates ASC but not STING, human AIM2 robustly activates both pathways, making it the only ALR among the 17 tested that has dual activity (Fig. 3.12D). Moreover, we found that human PYHIN-1 is a novel, robust, and specific activator of STING-dependent IFN production (Fig. 3.12D). In contrast, human IFI16 and MNDA did not strongly activate either STING or ASC in these assays, although IFI16 showed detectable activation of STING-IFN and ASC-inflammasome, consistent with previous reports [137] [135]. These data provide an unbiased functional characterization of all currently known murine and human ALRs, revealing several novel activators of STING and ASC.

Discussion

The innate immune response triggered by detection of intracellular DNA is important for host defense and causes a number of specific human autoimmune diseases, but the receptors that activate this response have remained elusive. We present here three distinct approaches used to identify the intracellular DNA sensor and components of the ISD pathway. Microarray analysis of DNA and RNA stimulated primary cells, DNA immunoprecipitation and DNA binding protein analysis and, evolutionary and functional characterization of a family of DNA receptors have resulted in a better understanding of the ISD pathway.

We found very few differences between gene induction of ISD and RIG-I ligand

stimulated MyD88-/- Ifnra1-/- macrophages by microarray. Of the top 25 genes most highly induced in ISD and RIG-I ligand stimulations only 3 of the 25 genes were exclusive to either stimulus (Fig. 3.1A,B). This was surprising since we reasoned that a different set of genes would be activated during the initial sensing of DNA as compared to RNA and that we would be able to use differential gene induction patterns to direct our search for candidate ISD sensors or adaptors. However, the cytosolic DNA and RNA sensing pathway are both dependent on the kinase TBK-1 and transcription factor IRF-3 for induction of IFN-β and this may be the reason for similar gene induction profiles [113] [95-98, 123, 164]. The most highly induced ISD-specific genes were histone variants (Fig. 3.1C). However, the overall induction of these histone variants in ISD stimulated macrophages was relatively low at around 2 fold over untreated as compared to the thousands of fold induction of IFN-β making them unlikely candidates for ISD sensors or adaptor proteins. Despite their lower induction after DNA stimulation, histone variants could be involved in anti-viral responses since many DNA virus genomes including simian virus 40 (SV40), adenovirus, herpes simplex virus 1 (HSV-1), Epstein-Barr virus (EBV), Hepatitis B virus (HBV) and Kaposi's Sarcoma herpes virus (KSHV) are all associated with chromatin upon infection (reviewed in [250]). For example, both adenovirus and HSV-1 genomes are assembled into nucleosomes during lytic infections, and HSV-1 uses activating histone H3 Lys14 acetylation and Lys4 methylation marks to promote viral gene expression [284-286]. Histones are poised to interact with DNA virus genomes, but it has yet to be shown that they are directly involved in anti-viral responses. It is possible that histones could serve as scaffolding for anti-viral molecules or that they permanently silence viral gene expression to prevent replication. Future work studying the role of histones in an anti-viral context could provide evidence for these speculations.

Our DNA immunoprecipitation approach identified two novel DNA binding proteins Interferon Induced Protein with Tetratricopeptide Repeats 1 (IFIT1) and 3'-5'Oligoadenylate Synthetase 3 (OAS3) in the lysates of IFN treated macrophages (Fig. 3.3C). In addition, we developed a siRNA knockdown system to evaluate the role of IFIT1 and its closely related family member IFIT2 as candidate ISD sensors or adaptor molecules. Despite this novel approach IFIT1 and IFIT2 knockdown did not affect the ability of primary macrophages to activate an IFN response in a specific manner (Fig. 3.4C). Upon IFIT1 and IFIT2 knockdown, we saw a non-specific decrease in type I IFN to both DNA and RNA ligands indicating that it may regulate IFN expression in a non-ISD specific manner. The results of our microarray showing IFIT1 upregulation in both DNA and RNA stimulated macrophages also indicate that IFIT1 does not have an ISD specific function (Fig. 3.1B). Another possibility is functional redundancy between the IFIT family members as we found with the AIM2-like receptors (Fig. 3.12). It is possible that the three IFIT proteins in mice (IFIT1, IFIT2 and IFIT3) have similar functions and knockdown of one IFIT is insufficient to reveal an IFN-β phenotype after DNA ligand stimulation. This could be addressed by developing siRNA knockdowns to target conserved regions in all three murine IFITs or by using a combination of IFIT siRNAs to knockdown all IFITs. Lastly, the identification of IFIT1 as an ISD binding protein may have been an artifact of our DNA immunoprecipitation assay. Since we immunoprecipitated DNA interacting proteins from lysates instead of transfecting and recovering DNA during normal physiological conditions, we may have identified non-specific nucleic acid binding proteins. We also used artificial phosphorothioate /phosphodiester DNA as a ligand that may have also increased non-specific protein recovery since its secondary structure is unknown. Indeed among its many ascribed functions discussed previously IFIT1 was recently shown to bind to 5'

triphosphate double stranded RNA[287] indicating that IFIT1 may actually be binding to our artificial DNA ligand because of secondary structure. In closing, our microarray analysis and DNA immunoprecipitation assays yielded some preliminary insights and several candidate ISD sensor/adaptor proteins.

The directed approach we took to identifying cytosolic DNA sensors by charactering the AIM2-like receptor family of proteins resulted in substantive data that increased our understanding of the ISD pathway components. Our results validate the principal findings of the first studies that identified specific ALRs as important components of the DNA-activated inflammasome [126-129] and the STING-dependent ISD pathway [135], and we identify several novel activators of these two pathways among human and mouse ALRs. Together, these data provide an evolutionary foundation and a functional framework for understanding the role of ALRs in host defense.

Our phylogenetic analysis of ALRs reveals a number of interesting features of this gene family, with important evolutionary and biological implications. On a practical level, the lack of preserved orthology between human and mouse ALRs (AIM2 excepted) means that a new nomenclature will need to be established for these genes, especially for those ALRs (MNDA and PYHIN-1) that have the same name in mouse and human but are neither evolutionary nor functional orthologs (Figs. 3.6 and 3.12). Interestingly, our functional analysis reveals that the two previously defined functions of STING-IFNs and ASC-inflammasome are preserved among the repertoire of murine and human ALRs despite this lack of sequence orthology at the level of individual genes. Moreover, we found that there are multiple ALRs capable of activating STING-IFNs and multiple ALRs capable of activating the ASC-inflammasome (Fig. 3.12). We propose that this functional redundancy, together with the remarkable diversity of

ALRs among even closely related species, reflects strong evolutionary pressures placed on the ALR gene family by pathogens (likely DNA viruses). Importantly, these selective pressures are pervasive throughout the mammalian lineage, and no single ALR is conserved as a functional gene in all mammals. The expansion of the murine ALR locus is the most dramatic among the mammals we examined, and our preliminary analysis of this locus in other strains of mice suggests that the numbers of ALR genes and pseudogenes differ among strains (data not shown). Such strain-specific differences in ALR gene composition might explain the contribution of ALRs to lupus susceptibility in certain inbred mice [288]. This species-specific diversification is unusual among candidate innate immune response genes, but not unprecedented. For example, the genes encoding the NOD-like receptors with pyrin domains (NLRPs; [289]), the APOBEC3 family of cytidine deaminases [290], and the immunity-related GTPases (IRGs; [291]) are all dramatically different in composition and number between humans and rodents.

We characterized the localization of all murine ALRs, in the absence and presence of STING and ASC, revealing that 12 of the 13 ALRs are nuclear proteins but move to localize with the adaptor proteins upon overexpression-induced signaling (Fig. 3.10). The first studies that identified the DNA-activated antiviral response suggested that detection of DNA likely occurs in the cytosol, thus offering a simple compartment-based explanation for how these pathways distinguish foreign from self DNA. However, it is important to note that nearly all DNA viruses, including those that have been implicated as ALR activators [135, 137], replicate in the nucleus. Thus, the steady-state nuclear localization of nearly all murine and human ALRs may position these receptors to respond to specific replication intermediates of DNA viruses and retroelements (Fig. 3.10; [128]). How these receptors distinguish microbial DNA from host genomic DNA, and how they move to the cytosol to form signaling complexes with STING

and/or ASC are important questions for future study. We observed this movement from the nucleus to form signaling complexes with STING or ASC for nearly all murine ALRs, and the localization patterns correlated strongly with the functional assays (Figs. 3.10 and 3.12). We found that the STING-activating ALRs recruited ASC to STING, but the ASC activators did not recruit STING to ASC (Fig. 3.10). This finding hints at a potential explanation for why ASC-deficient cells have a hyper-responsive ISD pathway [128], yet STING-deficient cells have a normal if reduced DNA-activated inflammasome response (Fig. 3.13). We suggest that this inducible recruitment of ASC to STING during ISD pathway activation is a negative regulatory event, the cell biology of which will be an interesting area for future research.

Our phylogenetic analysis of ALR diversity and the lack of known functions for most ALRs prompted us to develop simple, unbiased assays to determine the functional capabilities of each ALR and compare these functions among all of the ALRs in mice and humans. Using these assays, we uncovered novel ALRs that activate STING-dependent IFNs and the ASC inflammasome. Among the murine ALRs that potently activate STING, we found one Pyrin-only protein (PYR-A), one HIN-only protein (IFI202b), and two additional full-length ALRs (PYHIN-A, IFI203). These data reveal considerable functional redundancy among the ALR gene family in mice and suggest that both Pyrin and HIN domains can independently contribute to STING activation. Interestingly, for the three full-length murine ALRs that strongly activate STING defined here and elsewhere (PYHIN-A, IFI203, IFI204 [135]), their HIN domains cluster together on the phylogenetic tree but their Pyrin domains do not (Fig. 3.6C). We found two novel activators of the ASC inflammasome among murine ALRs (MNDA and MNDAL). However, AIM2-deficient mouse cells are profoundly impaired for the DNA-activated inflammasome [130], so neither MNDA nor MNDAL are functionally redundant with AIM2, at least for the

ligands and microbes used in these studies. We suggest that the nuclear MNDA and MNDAL receptors might respond to unique ligands or to a specific class of viruses that is not sensed by AIM2.

Additionally, we revealed novel functions for human ALRs. Of the 17 murine and human ALRs examined, human AIM2 is the only ALR capable of activating both IFNs and the inflammasome (Fig. 3.12). In contrast, the AIM2 ortholog in mouse (a rare example of clear sequence orthology in this family) activates the inflammasome but not IFNs. We propose that other, currently uncharacterized ALRs might have dual functionality, and that the loss of AIM2 in numerous mammalian species (Fig. 3.6) might be compensated for by other ALRs that activate the inflammasome, thus preserving this important innate immune function. Interestingly, we found that human PYHIN-1, an ALR with no previously described function, potently activates STING-dependent IFN production (Fig. 3.12). This potential antiviral function of PYHIN-1 provides a mechanistic basis for further study of the recently identified human PYHIN1 polymorphisms associated with asthma in African Americans [292]. Finally, we found several ALRs that had little detectable activation of either STING-IFNs or ASC-inflammasome. These ALRs may activate a response distinct from the ones we measured, they could require additional cofactors for activation, or they could be negative regulators of the other ALRs. Since several ALRs are co-expressed in cells (Fig. 3.8), the net outcome of activating multiple ALRs at once may be different from the activation of individual ALRs in isolation [95]. Overall, while we cannot infer in vivo function or ligand specificity from these studies, we emphasize that the utility of these assays lies in the direct, unbiased comparison of all ALRs, thus creating a map of ALR function that will guide future efforts to understand these receptors.

We found, using stable lentiviral knockdowns in primary macrophages and MEFs, that

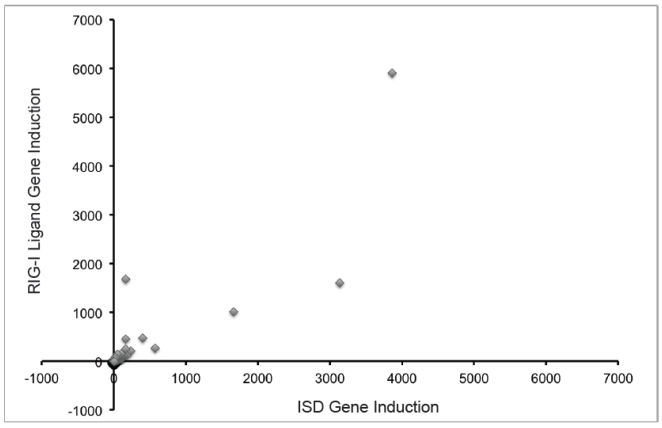
murine IFI204 is not a unique activator of the ISD pathway (Fig. 3.9). Moreover, we found that lentiviral knockdown of each of the murine ALRs individually had no significant effect on ISD sensing (data not shown). However, these knockdown siRNAs also did not efficiently reduce ALR mRNA levels as compared to control siRNAs indicating that knockdowns are an insufficient way to study the contribution of ALRs to type I IFN activation. Although others have shown a contribution of murine IFI204 to the DNA-activated IFN response [135], knockdown of IFI204 in these studies did not completely abolish ISD sensing, suggesting that other receptor(s) might contribute to this pathway. In light of our identification of several novel ALRs in mice that are closely related to IFI204 and our data revealing additional ALRs that activate STING, we suggest that multiple ALRs might be responsive to ISD ligands, similar to the way that transfected poly I:C can activate both RIG-I and MDA5 [104]. Based on these findings, complete characterization of the ALR gene family will require the generation of mouse models deficient in all ALRs.

In closing, our findings reveal the evolutionary diversity and functions of mammalian ALRs, and offer a number of possible explanations for why the dissection of ISD sensors has remained elusive despite several years of intense efforts. First, the use of "pure" ISD ligands that uniquely trigger the STING-dependent antiviral response is important because certain DNA polymers trigger two largely redundant pathways (Fig. 3.5). Second, the ALR gene family has little preserved orthology among mammals, complicating efforts to understand ALR function without the clear phylogenetic picture of ALR evolution that we provide here (Fig. 3.6). Third, individual ALRs are functionally inert when overexpressed on their own in the cell lines used for these assays (Fig. 3.12), thus rendering these receptors "invisible" to the conventional expression cloning approaches used to identify RIG-I/MDA5 [164], MAVS [96], and STING [114]. Finally,

we show that multiple mouse and human ALRs are competent for STING activation (Fig. 3.12), suggesting a functional redundancy among these ALRs that may complicate genetic approaches and RNAi-based screens to identify ISD sensors. Together, the data presented here create a framework for investigating the role of ALRs in the antiviral response.

Section 3 Figures

A



B

ISD Inducible Genes

Rank	Gene Name	Value
1	Ifnb1	3860.4
2	Cxcl10	3135.3
3	Tyki	1660.0
4	Rsad2	569.9
5	Irg1	398.0
6	Ifit3	233.6
7	Vcam1	166.9
8	Ifna2	164.8
9	Cxcl2	140.2
10	LOC667370	117.7
11	Ifit2	115.2
12	Oasl2	114.6
13	Cxcl1	111.1
14	Traf1	100.6
15	Peli1	93.1
16	Ccl5	87.7
17	Olfr503	61.5
18	Gpr109a	61.0
19	Tnf	59.7
20	Ccrl2	55.2
21	Cd69	52.0
22	A130038J17	50.1
23	Gbp2	49.6
24	Zdhhc18	42.9
25	Socs3	36.6

RIG-I Ligand Inducible Genes

Rank	Gene Name	Value
1	Ifnb1	5907.0
2	Ifna2	1686.0
3	Cxcl10	1609.5
4	Tyki	1008.8
5	Irg1	475.3
6	Vcam1	451.8
7	Rsad2	267.0
8	Ifit3	207.0
9	Cxcl1	158.5
10	Traf1	156.9
11	Gbp2	141.9
12	Cxcl2	129.9
13	Ifit2	110.4
14	Ccl5	93.7
15	Zdhhc18	91.5
16	LOC667370	84.8
17	Egr1	82.9
18	Socs3	68.4
19	Tnf	66.2
20	Cd69	60.1
21	Gbp3	59.1
22	Gpr109a	58.7
23	Oasl2	55.2
24	Ifna5	48.3
25	Peli1	47.6

C

ISD Specific Genes

Gene Name	ISD Fold Induction	RIG-I Ligand Fold Induction	Fold ISD over RIG-I Ligand
Gls	4.6	0.2	21.5
Hist1h4k	7.6	0.8	9.8
Ddc8	2.6	0.3	7.7
Hist1h4c	11.9	1.7	7.0
Hist1h4k	5.5	0.9	6.0
Hist1h4m	14.8	2.8	5.2
Hist1h2bb	2.3	0.5	4.6
Cpsf2	2.5	0.7	3.4
Egr3	14.1	4.2	3.3
Rfxank	2.2	0.7	3.3
Hist1h4j	5.0	1.6	3.2
Cpne4	4.1	1.3	3.0
Hist1h4f	6.0	2.0	3.0
Hist1h1b	7.0	2.3	3.0
Morn3	2.7	0.9	3.0
Rgs2	2.7	0.9	2.9
Zcchc2	3.2	1.1	2.9
Tnfrsf12a	2.4	0.8	2.9
Hist1h4a	4.8	1.7	2.9
Zeb1	6.4	2.3	2.8
Syt12	8.1	2.8	2.8
Elf3	2.2	0.8	2.8
Ctnnb1	2.7	1.0	2.7
Phlda3	2.2	0.8	2.7
Ass1	5.1	1.9	2.7
Rtn4	3.6	1.3	2.7
Mef2b	2.3	0.9	2.6
Hist1h2bg	29.9	11.4	2.6
Hist1h3a	4.0	1.6	2.6
Hist1h3h	4.8	1.9	2.5
Hist1h3e	3.5	1.4	2.5
Hist1h4f	3.7	1.5	2.4
Hist1h4f	3.6	1.6	2.3
Hist1h3f	3.4	1.5	2.3
Hist1h4m	3.0	1.4	2.2
Hist2h3c1	2.2	1.0	2.1
Hist2h2be	3.1	1.5	2.1

Figure 3.1 Gene induction in ISD and RIG-I ligand stimulated macrophages. (A) Graphical representation of gene induction in Myd88-/- Ifnrar1 -/- macrophages stimulated for 4 hours with ISD oligo or in vitro transcribed RIG-I ligand. Each point represents a transcript. **(B)** Tables of the top 25 transcripts upregulated in ISD (left) and RIG-I ligand (right) stimulated macrophages. Of the top 25 genes, only 3 differ between ISD and RIG-I ligand stimulation (red). (C) Top ISD-specific genes were identified by calculating fold induction ISD and RIG-I ligand stimulations over untreated and then dividing the ISD-stimulated fold induction over RIG-I ligand fold induction. The top 19 ISD-specific genes were histone variants (red).

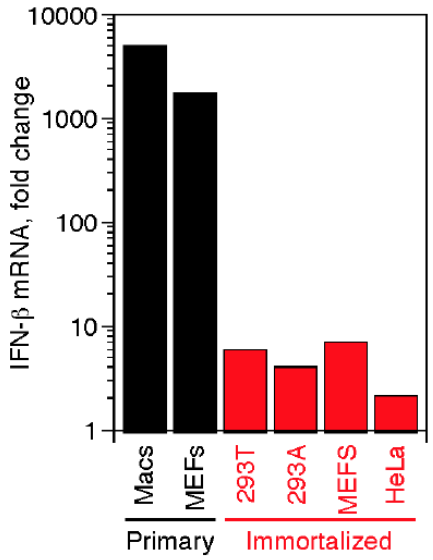


Figure 3.2 Immortalized cells do not robustly response to intracellular DNA. Primary C57BL/6 bone marrow-derived macrophages and primary low passage mouse embryonic fibroblasts (MEFs, black) and immortalized 293T, 293A, MEFs and HeLa cell lines (red) were transfected with 1 μg calf thymus DNA for 4 hours and IFN-β mRNA induction was assessed by quantitative real-time PCR. Primary cells responded over 100 fold more strongly to transfected DNA than any immortalized cell line tested.

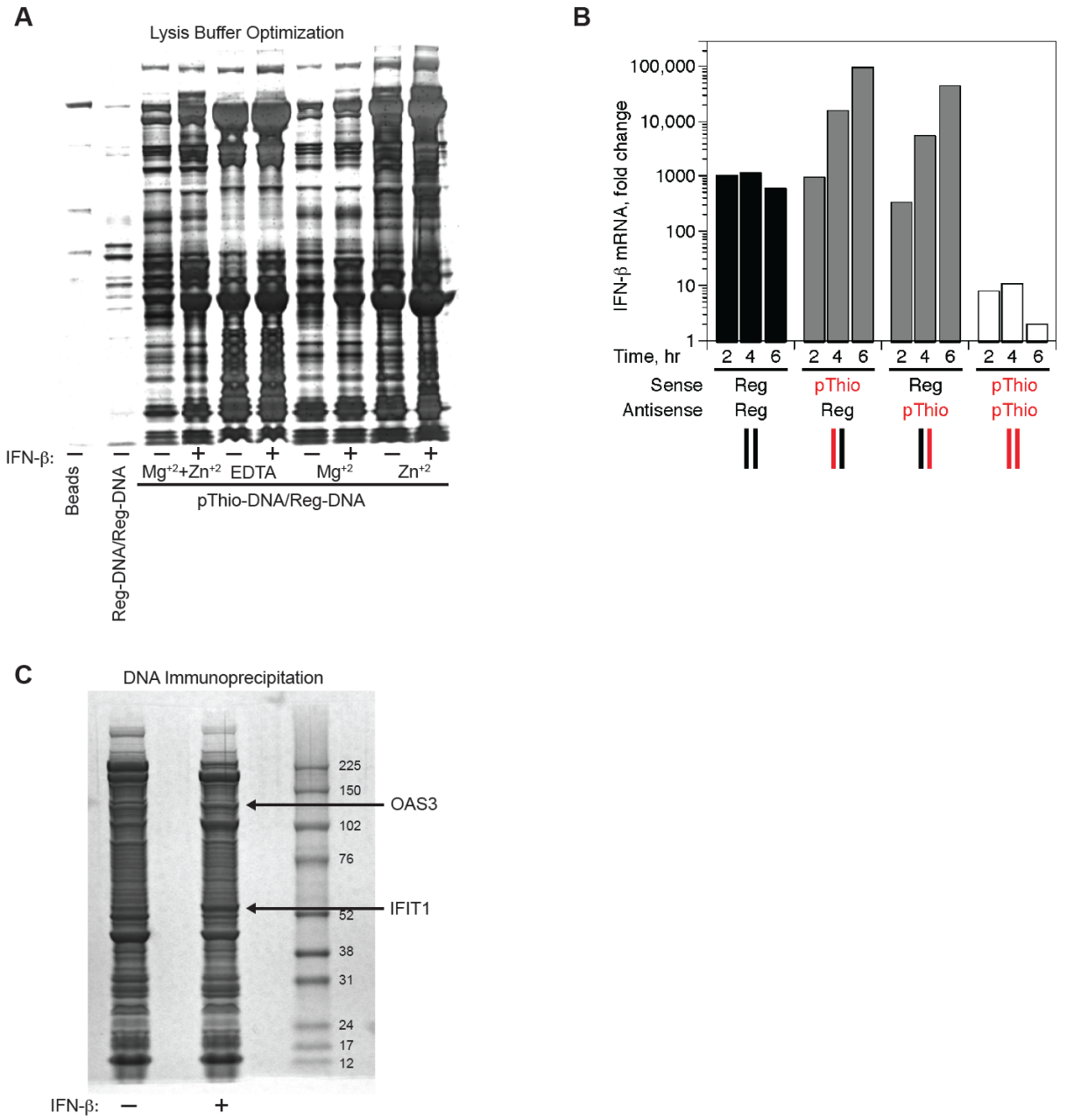


Figure 3.3 DNA immunoprecipitations identify two novel DNA binding proteins. (A) Lysis buffer optimization of bone marrow-derived macrophages pre-treated with 25U IFN-β for 24 hours before lysis in buffer containing either 2 mM EDTA, 2 mM magnesium (Mg^{+2}), 0.2 mM zinc (Zn^{+2}), or both magnesium and zinc. (B) Bone marrow-derived macrophages were transfected with regular DNA, regular/phosphothiol DNA, or phosphothiol DNA for 4 hours and IFN-β production was assayed. Regular/phosphothioate DNA hybrids induced the greatest IFN-β production. **(C)** Primary bone marrow derived macrophages were treated and probed as described in (A) using lysis buffer containing both magnesium and zinc. After size separation and coomassie blue stain bands that were enriched in IFN-β pre-treated macrophages were excised and identified by mass spectrometry as OAS3 and IFIT1.

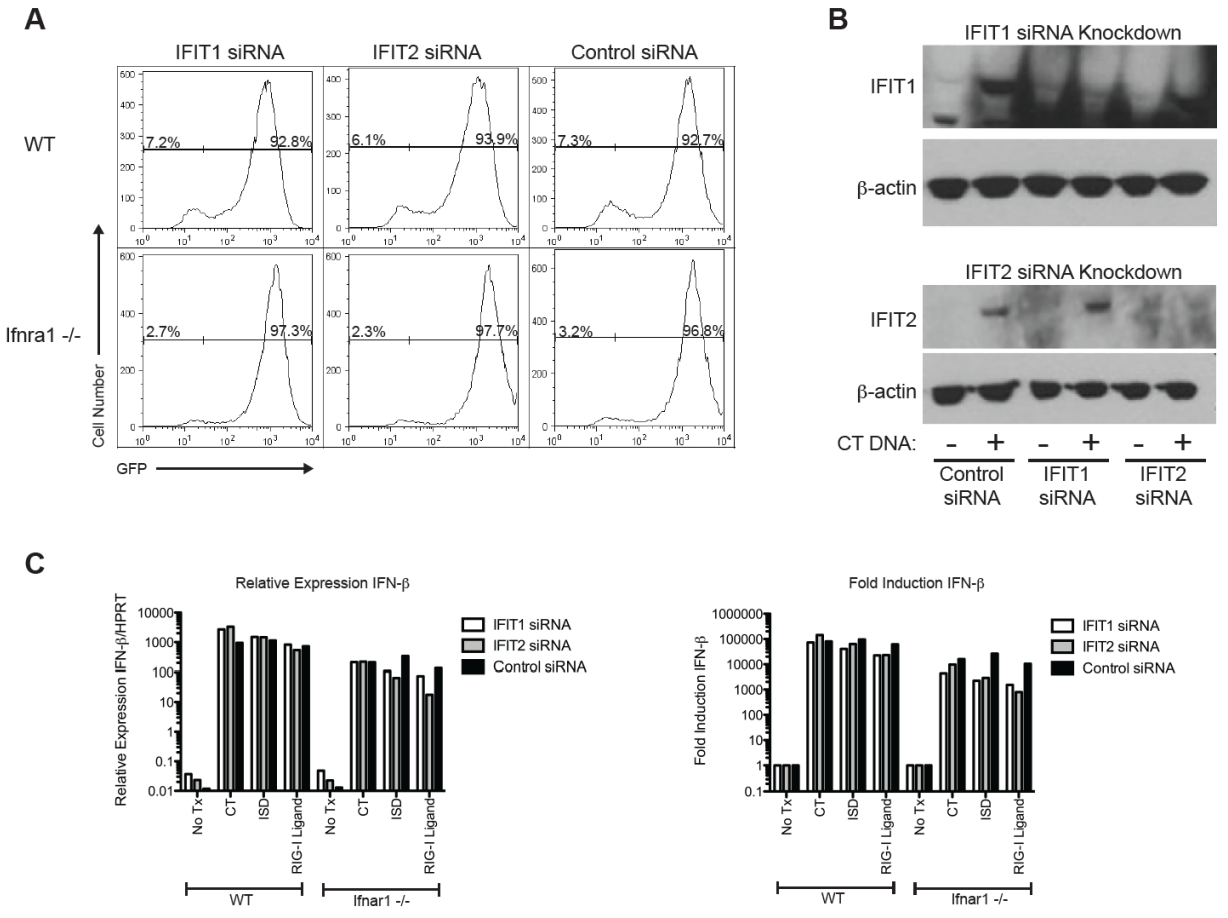


Figure 3.4 IFIT1 and IFIT2 knockdown in primary Ifnar1-/- macrophages. (A) Flow cytometry of GFP expression after puromycin selection of IFIT1, IFIT2 and Control siRNA transduction shows over 97% purity of Ifnra1-/- transduced macrophages as compared to ~93% purity observed in wild type C57BL/6 macrophages. **(B)** Western blot analysis of Ifnra1-/- knockdown lysates after 4 hours of no treatment or CT DNA stimulation. Both IFIT1 and IFIT2 were absent in steady state conditions and remained resistant to upregulation after CT DNA treatment as compared to control siRNA knockdown. **(C)** Quantitative real-time PCR of relative expression to HPRT (left) and fold induction (right) IFN-β mRNA 4 hours after ISD DNA, CT DNA and RIG-I ligand stimulation. IFN-β production is slightly decreased in both IFIT1 and IFIT2 knockdown in response to all stimulations.

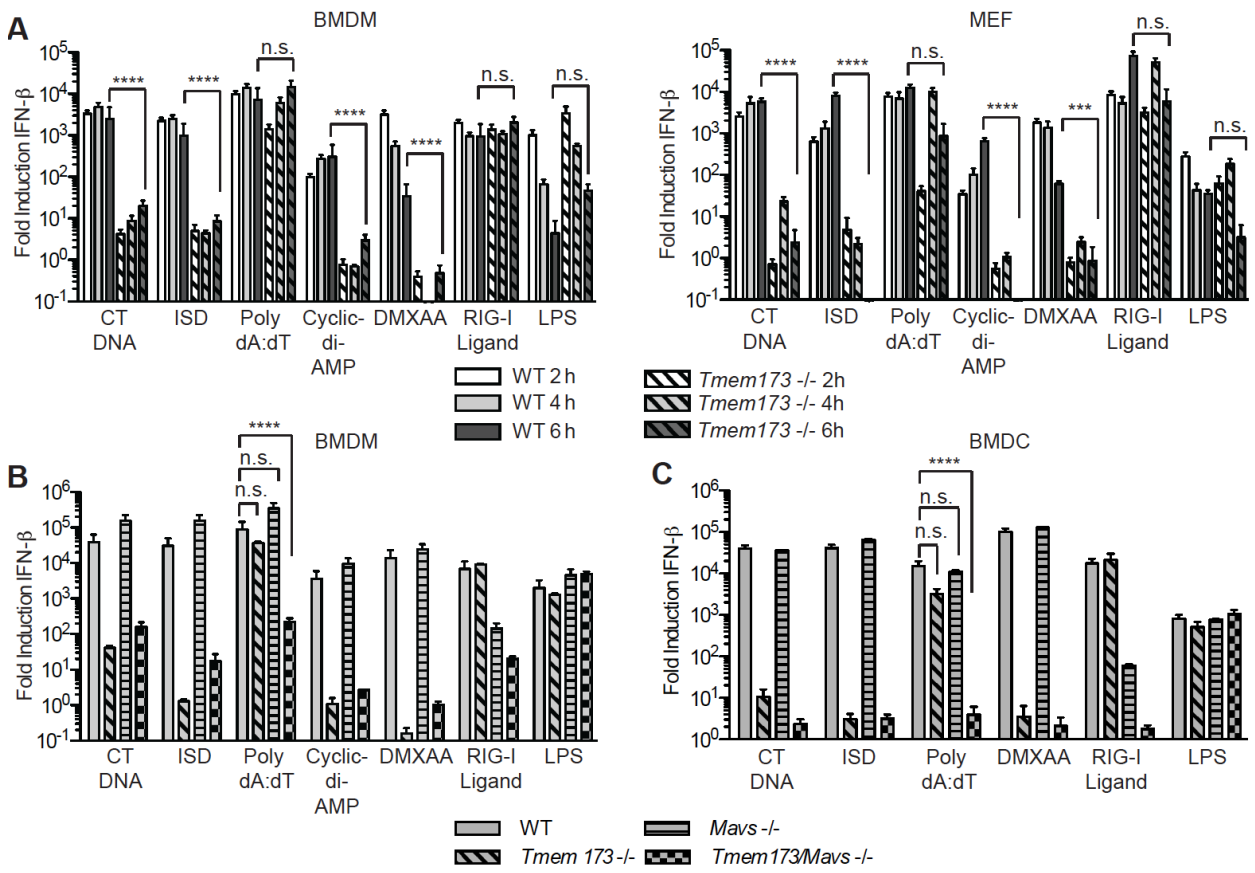


Figure 3.5 STING is required for sensing of natural DNA, cyclic dinucleotides and DMXAA. (A) Quantitative RT-PCR analysis of IFN-b mRNA expression in wild-type (WT) or Tmem173-/- bone marrow derived macrophages (BMDM) (left panel) and primary, early passage mouse embryonic fibroblasts (MEFs) (right panel); cells were treated with the indicated ligands and harvested for analysis every 2 hours. **(B)** WT, Tmem173-/-, Mavs-/-, and Tmem173-/- Mavs-/- bone marrow derived macrophages were stimulated with the indicated ligands for 4 hours. **(C)** Bone marrow dendritic cells (BMDCs) of the indicated genotypes were generated and stimulated with the indicated ligands for 4 hours before harvest and quantitative RT-PCR analysis. Data are representative of 2-3 experiments, each with triplicate treatments for every time point (error bars, s.e.m.). Statistical analysis was performed on the 6 hour timepoint (A) and 4 hour timepoint (B, C) with a 2-way ANOVA test. ****$p < 0.0001$; ***$p < 0.001$; n.s. = not statistically significant ($p > 0.05$). Significant values in panels B and C are identical to those found in Fig. 1 A and are omitted for figure clarity.

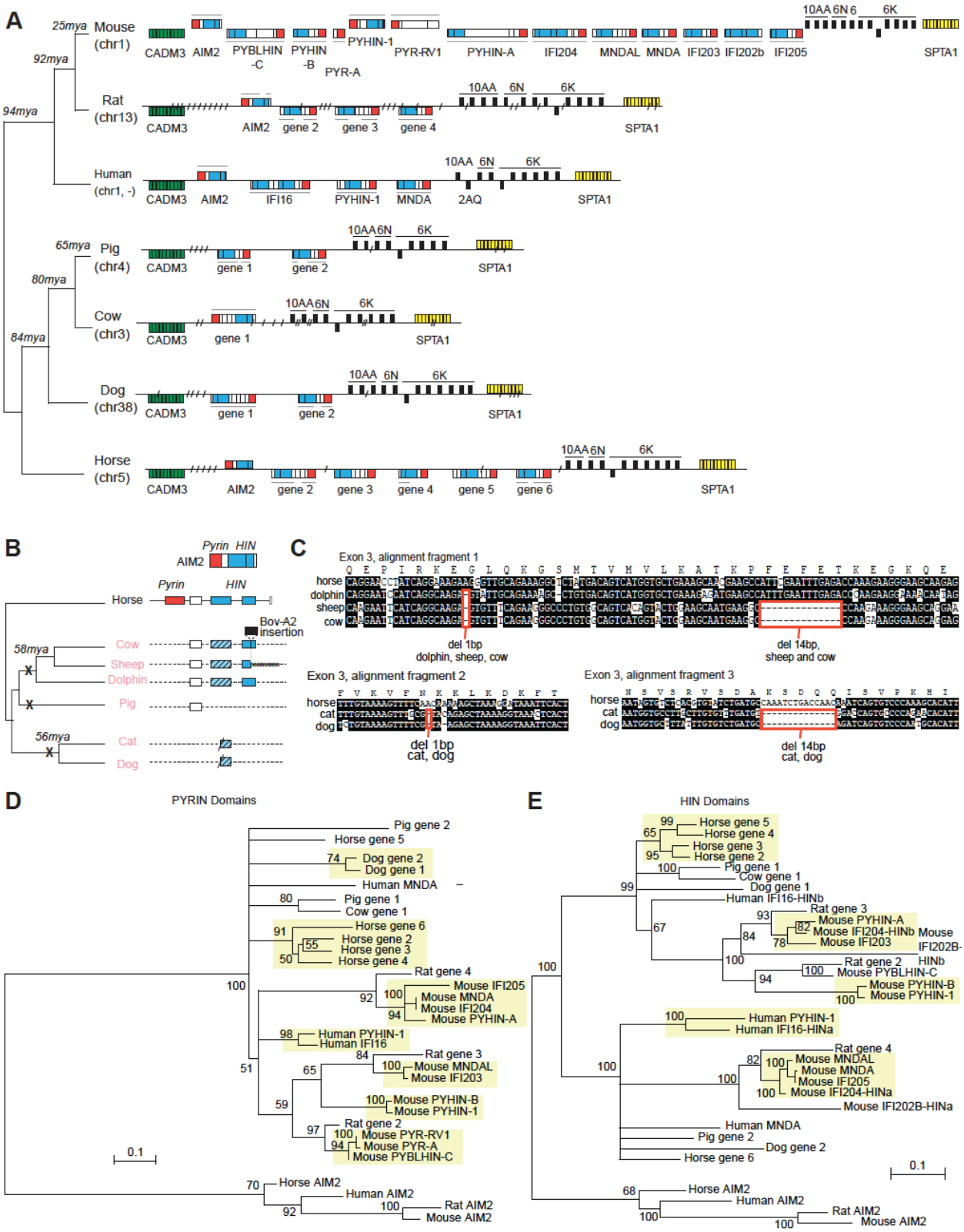

A
25mya
92mya
94mya
65mya
80mya
84mya
58mya
56mya
Mouse (chr1)
Rat (chr13)
Human (chr1, -)
Pig (chr4)
Cow (chr3)
Dog (chr38)
Horse (chr5)
CADM3
AIM2
PYBLHIN -C
PYHIN -B
PYHIN-1
PYR-RV1
PYR-A
PYHIN-A
IFI204
MNDAL
MNDA
IFI203
IFI202b
IFI205
10AA 6N 6
6K
SPTA1
gene 1
gene 2
gene 3
gene 4
gene 5
gene 6
IFI16
2AQ
B
Pyrin
HIN
AIM2
Bov-A2 insertion
Horse
Cow
Sheep
Dolphin
Pig
Cat
Dog
C
Exon 3, alignment fragment 1
horse
dolphin
sheep
cow
del 1bp
dolphin, sheep, cow
del 14bp,
sheep and cow
Exon 3, alignment fragment 2
cat
del 1bp
cat, dog
Exon 3, alignment fragment 3
del 14bp
cat, dog
D
PYRIN Domains
Pig gene 2
Horse gene 5
Dog gene 2
Dog gene 1
Human MNDA
Pig gene 1
Cow gene 1
Horse gene 6
Horse gene 2
Horse gene 3
Horse gene 4
Rat gene 4
Mouse IFI205
Mouse MNDA
Mouse IFI204
Mouse PYHIN-A
Human PYHIN-1
Human IFI16
Rat gene 3
Mouse MNDAL
Mouse IFI203
Mouse PYHIN-B
Mouse PYHIN-1
Rat gene 2
Mouse PYR-RV1
Mouse PYR-A
Mouse PYBLHIN-C
0.1
Horse AIM2
Human AIM2
Rat AIM2
Mouse AIM2
E
HIN Domains
Horse gene 5
Horse gene 4
Horse gene 3
Horse gene 2
Pig gene 1
Cow gene 1
Dog gene 1
Human IFI16-HINb
Rat gene 3
Mouse PYHIN-A
Mouse IFI204-HINb
Mouse IFI203
Mouse IFI202B-HINb
Rat gene 2
Mouse PYBLHIN-C
Mouse PYHIN-B
Mouse PYHIN-1
Human PYHIN-1
Human IFI16-HINa
Rat gene 4
Mouse MNDAL
Mouse MNDA
Mouse IFI205
Mouse IFI204-HINa
Mouse IFI202B-HINa
Human MNDA
Pig gene 2
Dog gene 2
Horse gene 6
0.1
Horse AIM2
Human AIM2
Rat AIM2
Mouse AIM2

Figure 3.6 Evolutionary relationships of the ALR gene family in mammals. (A) Genomic loci containing ALRs in selected mammalian species (loci not drawn to scale). Boxes represent coding exons, with those encoding Pyrin domains in red and HIN domains in blue. Introns are omitted for clarity. ALRs from species other than mouse and human are arbitrarily named. Gray horizontal lines show which exon-exon boundaries are supported by mRNA or EST evidence; other boundaries are predicted computationally. Short diagonal lines indicate gaps in the genome assemblies. Flanking genes CADM3 (green) and SPTA1 (yellow) have numerous exons that are not all presented here. Olfactory receptor genes and pseudogenes (black boxes) are also found at one end of the locus and are labeled according to their family and subfamily (e.g. 10AA = family 10, subfamily AA) using the HORDE system. A species tree (not to scale) is shown on the left along with the approximate divergence times. The position of horse on the tree is uncertain i [293]; we have depicted it in the position that is best supported by current evidence. Chromosomal locations of the locus are in parentheses, with a "-" symbol denoting that the human locus is depicted as flipped with respect to its numbered chromosomal orientation. **(B)** Using the horse AIM2 as a query in blast searches, we detected fragments of AIM2 pseudogenes in the genomes of cow, sheep and dolphin, whereas pig Aim2 appears to have only retained a recognizable exon 2. Cat and dog genomes only have a partial exon 3. This allows us to date the loss of AIM2 to at least 55 million years ago. **(C)** Three selected portions of a multi-species AIM2 alignment showing inactivating mutations shared by more than one species (red boxes). Highlighted bases are conserved in >50% of the sequences shown. These shared mutations show that the AIM2 gene was inactivated in the carnivore lineage before cat and dog diverged (~56 mya) and in at least some species of the Cetartiodactyla lineage since before dolphins, sheep and cattle diverged (~58 mya). **(D, E)** Phylogenetic trees of the indicated mammalian ALRs based on Pyrin domains **(D)** and HIN domains **(E)**. Species-specific ALR gene expansions are indicated in yellow boxes, with clear evidence of horse-, human- and mouse-specific gene expansions.

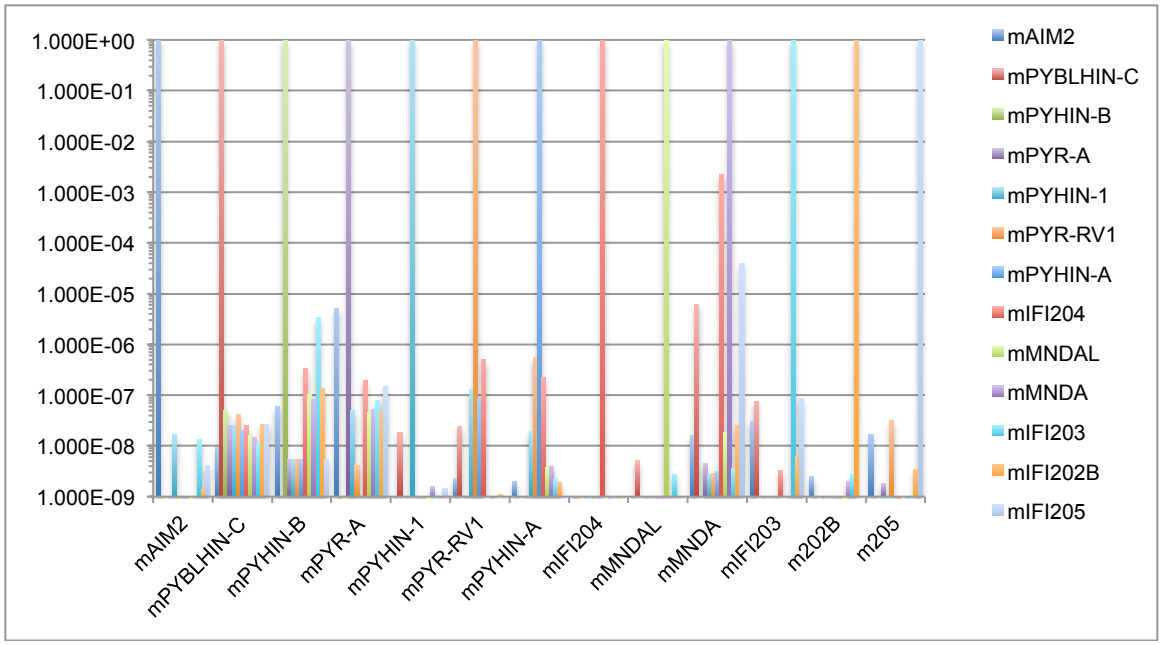


Figure 3.7 Specificity of quantitative RT-PCR amplification of the murine ALR cDNAs. 1 ng of plasmid containing each ALR cDNA was amplified with primers specific for each ALR. The amplification efficiency (Ct value) of the intended target was set to 1, and the relative efficiency of amplification for all other ALRs with each primer set was compared to that of the intended ALR target sequence. With one exception (MNDA), amplification of off-target ALRs was over five orders of magnitude lower than the intended amplicon. MNDA, which has extremely high similarity to IFI204 and IFI205 within the amplified region, was still amplified 500 times more efficiently than IFI204 and 25,000 times more efficiently than IFI205.

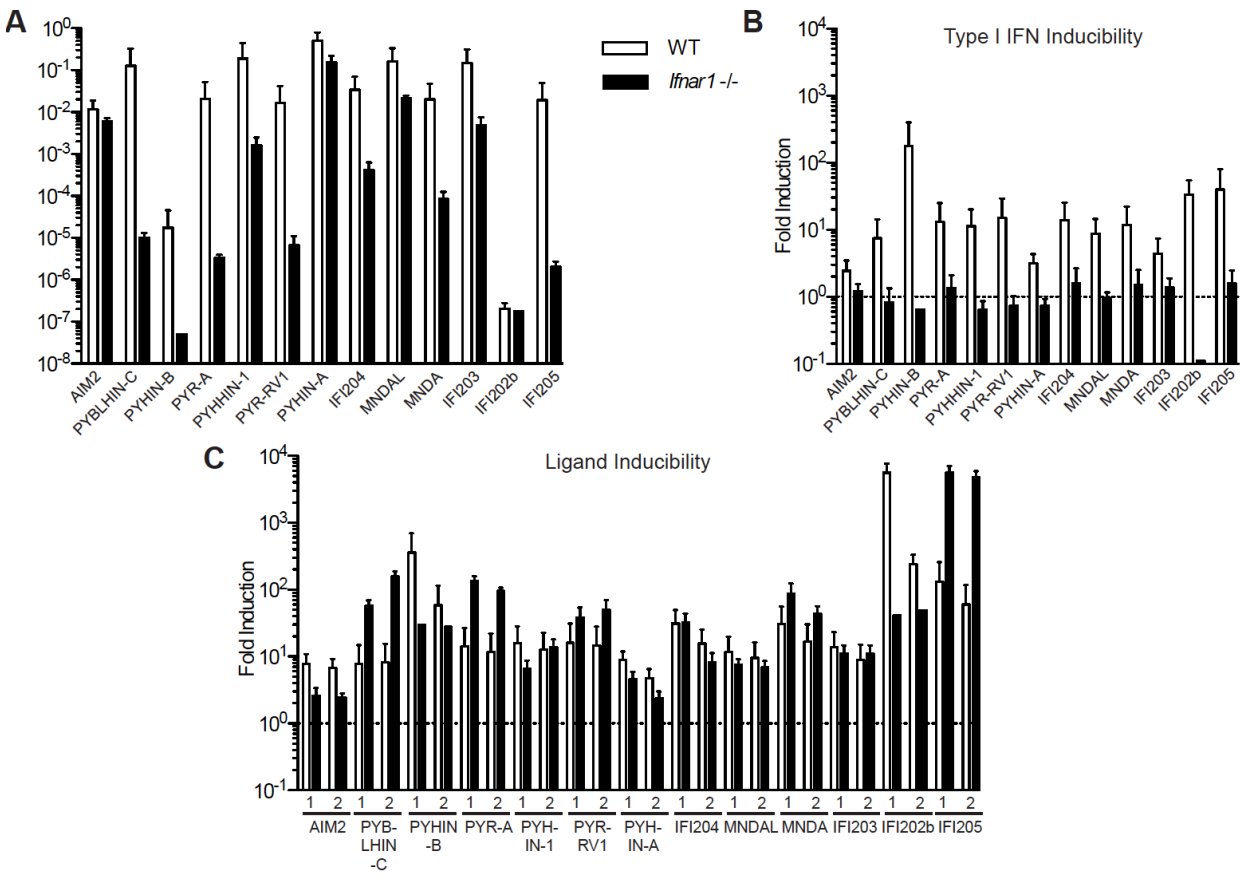


Figure 3.8 Expression and inducibility of murine ALR mRNAs. (A) Basal expression of each ALR mRNA in untreated WT and Ifnar1-/- BMDM, calculated relative to HPRT expression. **(B)** IFN-inducibility of murine ALRs: the basal expression level of each ALR was set to "1", and the fold induction of each ALR mRNA over untreated in WT and Ifnar1-/- BMDM is shown in response to a 6 hour stimulation with 100 U/ml recombinant mouse IFN-β. **(C)** Ligand inducibility: CT DNA (1) or RIG-I Ligand (2). Results are representative of three experiments, with triplicate treatments in each experiment (error bars, (A) s.d., (B,C) s.e.m.).

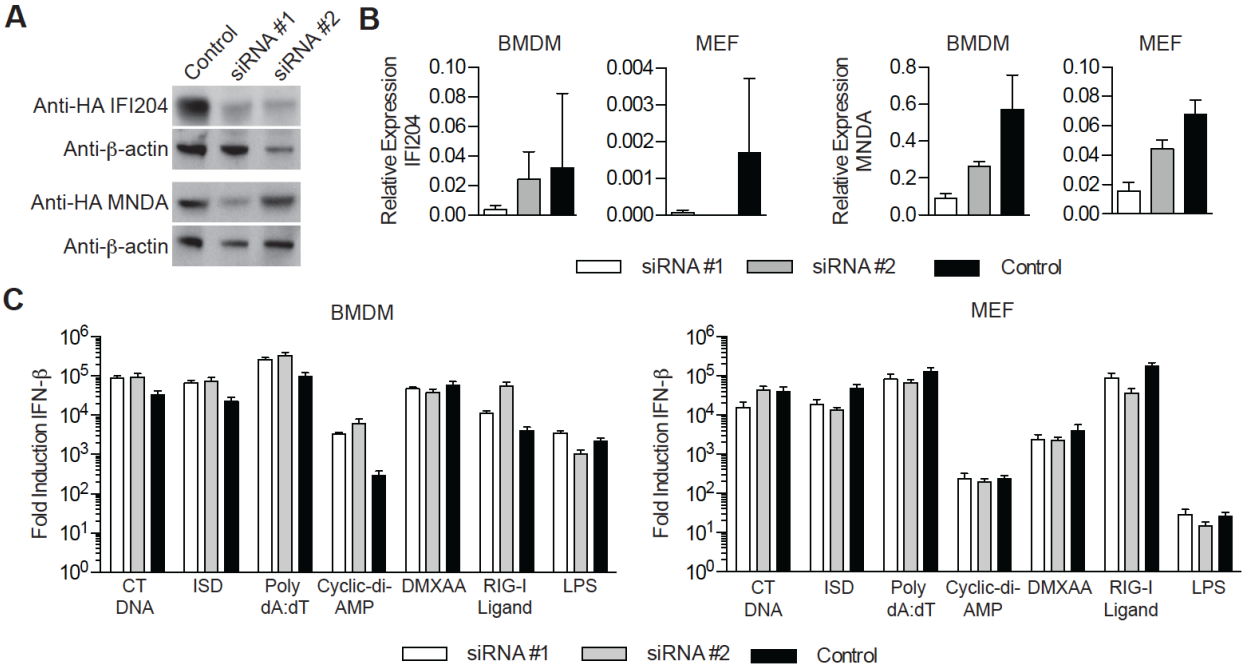


Figure 3.9 Murine IFI204 is not a non-redundant sensor of DNA. (A) Knockdown efficiency in HEK 293T cells transiently transfected with 500 ng hemagglutinin-tagged (HA-tagged) IFI204 and 2 μg siRNA #1, siRNA #2, or control siRNA for 24 hours. IFI204 protein knockdown was assessed by anti-hemagglutin (anti-HA) immunoblot (IB). **(B)** Quantitative real-time PCR of IFI204 mRNA expression in BMDM (left panel) and MEF (right panel) stably expressing the indicated lentiviral siRNA constructs. **(C)** Quantitative RT-PCR analysis of IFN-β mRNA expression in BMDM and MEF stably expressing the indicated lentiviral siRNAs and treated as indicated for four hours before harvest. Results are representative of two independent experiments, each with triplicate treatments for each ligand (error bars, (B) s.d., (C) s.e.m.). Statistical analysis was performed with a 2-way ANOVA test, and none of the stimulated knockdown cells reached statistical significance of $p<0.05$ compared to control cells.

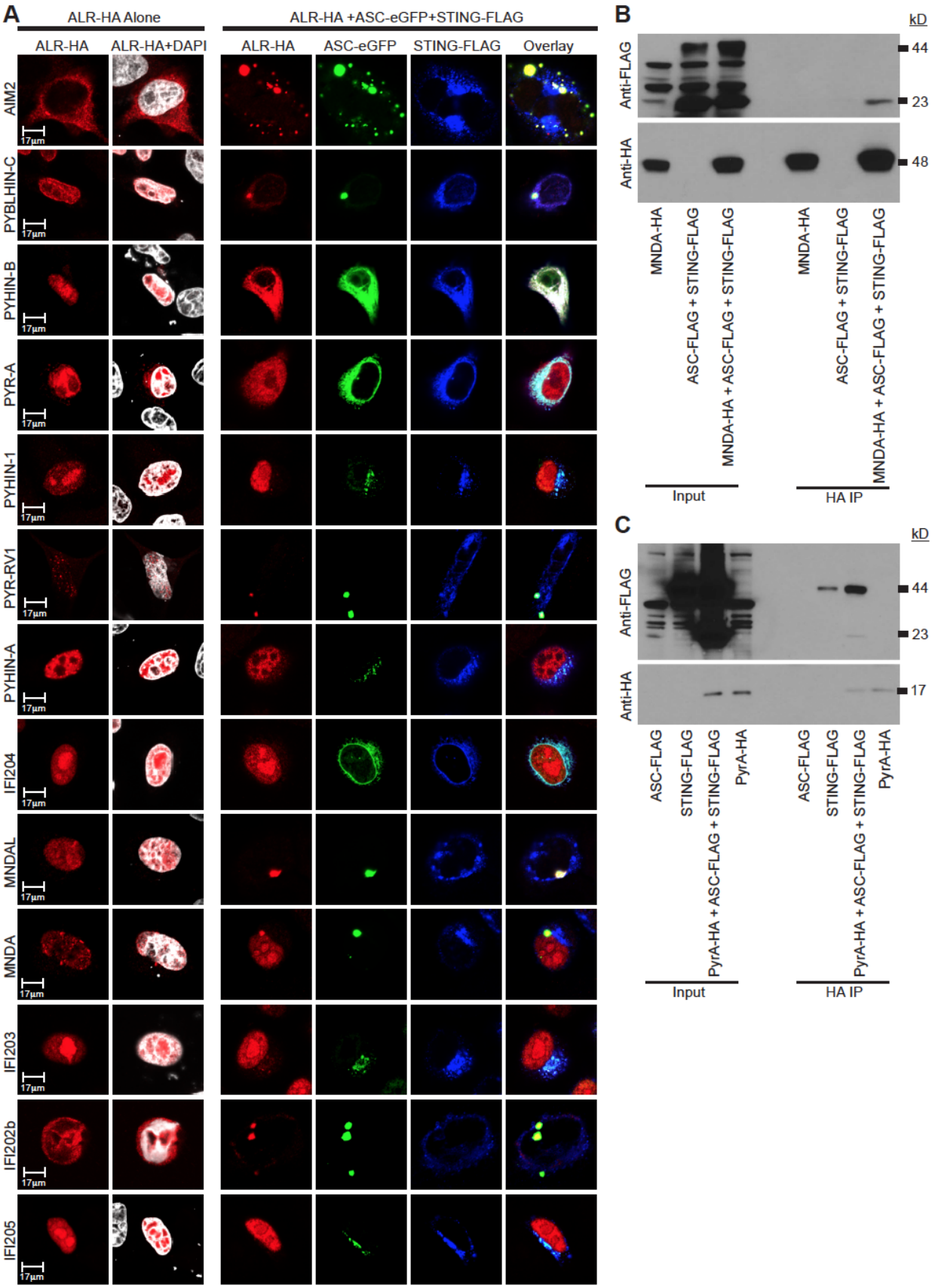

A
ALR-HA Alone
ALR-HA +ASC-eGFP+STING-FLAG
ALR-HA
ALR-HA+DAPI
ALR-HA
ASC-eGFP
STING-FLAG
Overlay
AIM2
PYBLHIN-C
PYHIN-B
PYR-A
PYHIN-1
PYR-RV1
PYHIN-A
IFI204
MNDAL
MNDA
IFI203
IFI202b
IFI205
17μm
B
kD
Anti-FLAG
Anti-HA
44
23
48
MNDA-HA
ASC-FLAG + STING-FLAG
MNDA-HA + ASC-FLAG + STING-FLAG
Input
HA IP
C
kD
Anti-FLAG
Anti-HA
44
23
17
ASC-FLAG
STING-FLAG
PyrA-HA + ASC-FLAG + STING-FLAG
PyrA-HA
Input
HA IP

Figure 3.10 Intracellular localization of murine ALR proteins. (A) 500 ng ALR-HA expression plasmids alone (left two columns) or 500 ng ALR-HA, 500 ng ASC-eGFP and 500 ng STING-FLAG (right columns) were transfected into 3.0 x 10^4 HeLa cells and probed 24 hours post-transfection with anti-HA-ALR (red), anti-eGFP (green), anti-FLAG (blue) and DAPI (grey). **(B)** HA-tagged murine MNDA was immunoprecipitated from control cells and cells expressing STING-FLAG and ASC-FLAG, which can be distinguished based on their size despite the same tag. **(C)** HA-tagged Pyr-A was immunoprecipitated from cells expressing the indicated constructs, followed by blotting for STING-FLAG and ASC-FLAG. All data are representative of 3-5 independent experiments.

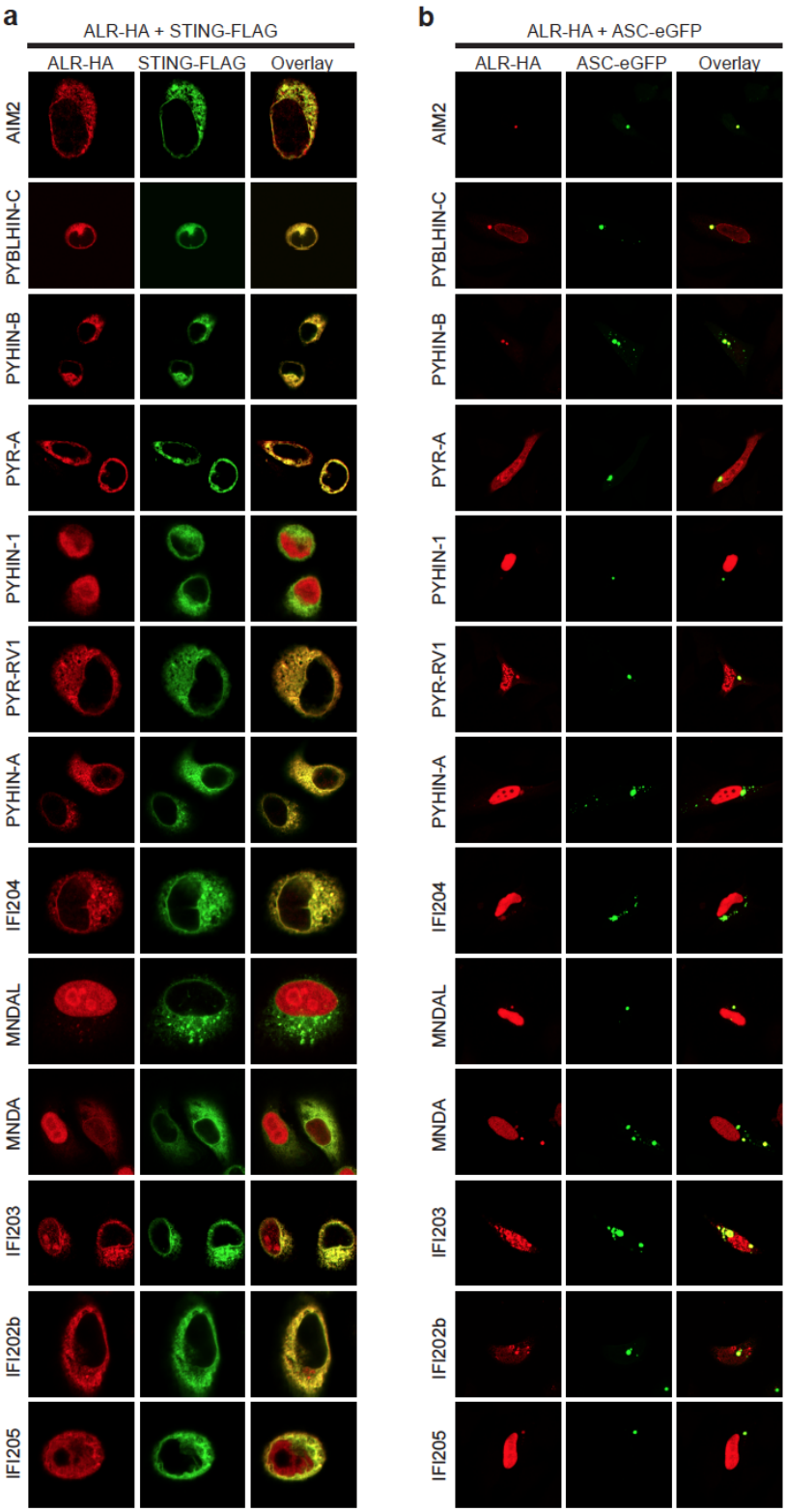


Figure 3.11 Immunofluorescence of ALR with STING or ASC. $3.0x10^4$ HeLa cells were co-transfected with 500ng of HA-tagged mouse ALR protein and 500ng of FLAG-tagged mouse STING (A) or 500 ng mouse ASC-EGFP **(B)** for 24 hours. Cells were fixed, permeabilized and stained with anti-HA-ALR (red) and anti-FLAG (green) or anti-eGFP (green).

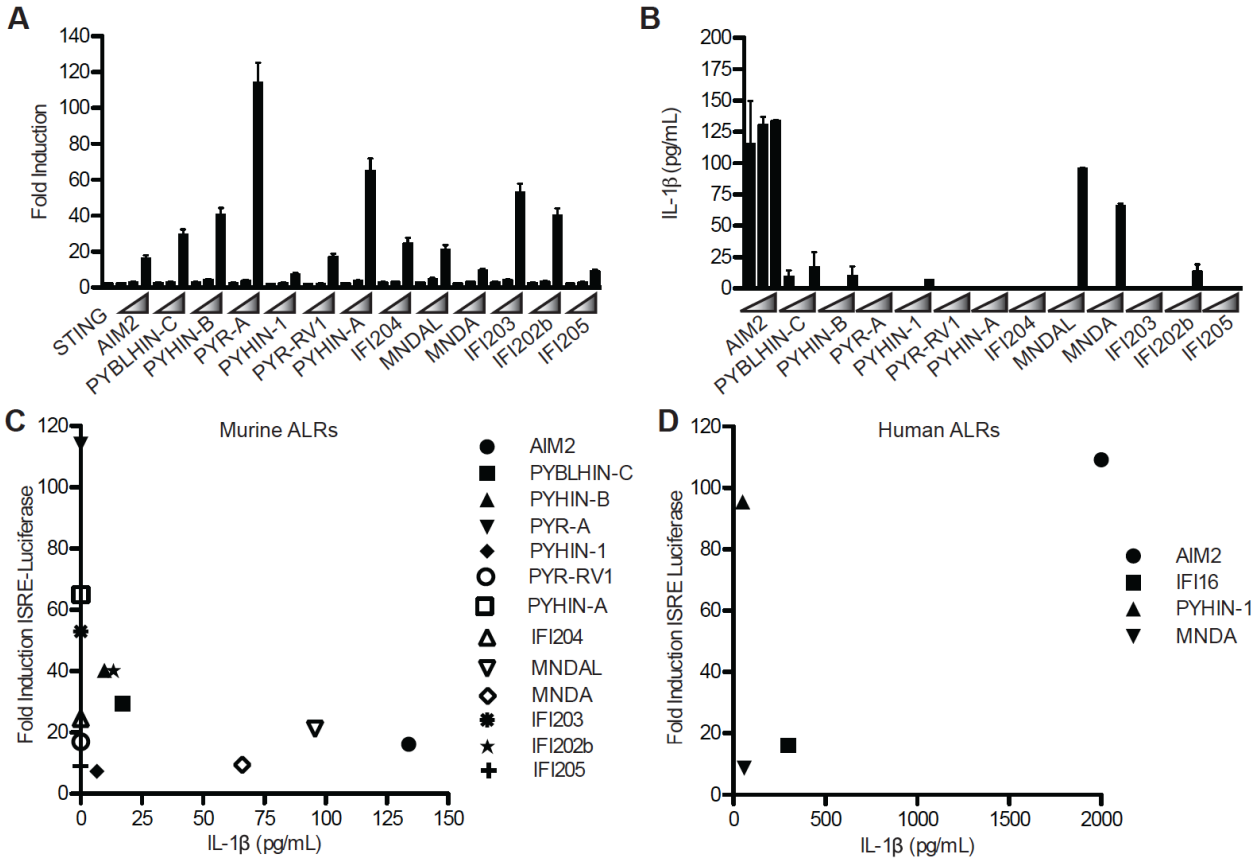


Figure 3.12 Unbiased functional characterization of all mouse and human ALRs. (A) Interferon stimulated response element (ISRE) luciferase reporter assay in HEK 293T cells transiently transfected with ISRE luciferase reporter and increasing amounts of ALR-HA expression vectors with 300 ng STINGFLAG vector, measured 24 h post transfection. Results are expressed as fold induction over ISRE-luciferase alone. **(B)** IL-1β ELISA of HEK 293T cells stably expressing murine ASC-FLAG and murine Caspase-1, transfected with increasing amounts of ALR-HA expression vectors and 100 ng pro-IL-1β vector. **(C)** Plot of ISRE luciferase induction versus IL-1β secretion for each ALR. **(D)** Plot of ISRE-luciferase induction versus IL-1β secretion for each human ALR. Results are representative of 2-5 experiments, with each treatment done in triplicate (error bars, (A) s.e.m., (B) s.d.).

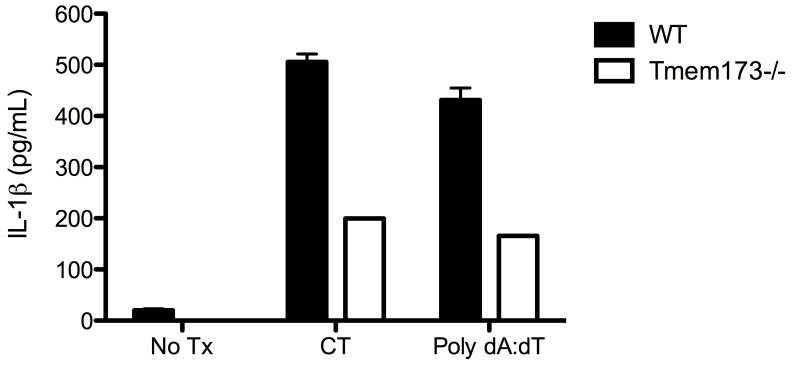


Figure 3.13 STING deficient macrophages do not have an enhanced inflammasome response. Wild type C57BL/6 and STING deficient (Tmem173) bone marrow derived macrophages were pretreated with LPS for 16 hours prior to 5 μg calf thymus (CT) DNA or 5 μg poly dA:dT transfection before supernatant harvest 24 hours post-transfection for IL-1β ELISA. STING deficient cells actually had slight reduced inflammasome responses to DNA ligands.

Section 3 Tables

Table 3.1 Open reading frames of thirteen ALRs in C57BL/6 mice

Gene name: *Aim2*
Other names: Gm1313; Ifi210
Open reading frame:
ATGGAGAGTGAGTACCGGGAAATGCTGTTGTTGACCGGCCTGGACCACATCACGGAGGAA
GAACTGAAACGGTTCAAGTACTTTGCTTTGACTGAGTTTCAGATTGCCAGGAGCACACTCG
ACGTGGCAGATAGGACAGAGTTAGCTGACCACCTGATTCAAAGTGCAGGTGCGGCGTCTG
CAGTGACCAAGGCCATTAATATTTTCCAGAAGTTGAATTATATGCATATTGCAAATGCTCTT
GAAGAGAAAAAGAAAGAAGCTGAACGTAAACTCATGACCAATACAAAGAAGAGAGGAACAC
AGAAGGTAGAAAATAGAAGTCAAGCTGAAAACTGCTCTGCTGCCTCTGCCACCCGCAGTG
ACAATGACTTTAAGGAACAGGCTGCTACAGAAGTCTGTCCTCAAGCTAAGCCTCAGAAGAA
ACAGATGGTGGCAGAACAGGAAGCCATCAGAGAAGATTTACAGAAAGATCCACTTGTTGTC
ACGGTGCTGAAAGCTATAAATCCCTTTGAGTGTGAGACTCAGGAAGGAAGACAAGAGATAT
TTCATGCAACAGTGGCCACGGAGACAGATTTTTTCTTTGTAAAAGTTTTAAACGCACAGTTT
AAAGATAAATTTATCCCAAAGAGGACAATTAAAATATCAAACTACCTTTGGCACAGTAACTT
CATGGAGGTCACCAGTTCCTCAGTTGTGGTTGATGTTGAATCTAACCACGAAGTCCCAAAT
AACGTTGTTAAGAGAGCCAGGGAAACTCCCAGGATTAGTAAACTGAAGATTCAGCCATGTG
GAACAATTGTGAATGGGCTGTTTAAAGTCCAGAAGATAACAGAGGAAAAAGATAGAGTACT
GTATGGTATACATGATAAAACAGGGACAATGGAGGTGTTGGTGCTGGGAAACCCAAGCAA
AACAAAGTGCGAGGAAGGAGACAAGATTAGACTCACGTTCTTTGAGGTGTCAAAAAATGGA
GTGAAAATTCAGTTGAAATCTGGACCTTGTAGCTTTTTTAAGGTTATTAAGGCTGCAAAGCC
AAAAACTGACATGAAAAGTGTGGAGTGA

Gene name: *Pyblhin-C*
Other names: “gamma-interferon-inducible protein 16-like”; Gm4955; Ifi206
Open reading frame:
ATGGTGAATTACTACAAGCAAATTGTTCTTCTGTCTGGATTAGAATATATGAATGACTATAAC
TTTAGAGCACTTAAGTCCTTACTGAACCATGACCTAAAACTGACTAAGAACATGCAAGATGA
TTATGACAGAATTAAGATTGCTGACTTAATGGAAGAAAAGTTCCCAGAGGATGCTGGATTG
AGCAAACTGATAGAAGTATGCGAAGATATTCCAGAACTTGCAGATCATGTTGATATACTTAG
AAAAAAGAAGGAGAAAGTTAAAAATAAAACCAAAATAAAATCTGAATCCAGCCCTCCTCCAC
TTACCTCCAGCCTGATGGAAGCTTGGGAAGTAGAACCAGCTATGGTGACAGCTTCATCAGA
GGAAAGCAAAGATACTATCCCTGAATCACCTGATACCATGACAACACAGTTTTTGGAGGAA
AAGCCTAAGTTTCCATTACTTTCAACAACCAGCACAAGCCAGGCTGAGGGTGAACCCCTGA
CTCCTCAGAGATTTCCAACAACAGCCTCCAGCAGTCTCCAGACACCCCTGGAGCCTCCAG
AAATATCGTCCACCATATTGGCAACATCTCAGGGTTCTTCAGTACCATATTCCACCTGTGAC
AAATCTTCTCGGGTGCCTCCAGTGACAGCCTCCAGCAGCACTCAGACCATCCAGACATGT
CTGGCAACCTCAACACTTCCTTGTAGTCACCAAGCTACTCTGAAGTCTCCAAAAACAGAGC
CCAGCAGTGTCCAGGCCACTCAGATGACTCAAGCAATAAAGGCAAGTGGTCATAATTGTCC
TCAGGTGCCTGCAGCACCAGTATCCAGCAGTTTCATTAAACCTCAGGTGACCCCAGCAATG
TTACTCAGCGGTGTCCAAGCTCCTTTGATGCCTCAAGCAACAGTGCCCAGCTGTGCCCAG
ACCTTTGAAATGAACCCAGCAACAATGACCAATGGTTGTAATAGTCCTCAGATGTCTGCAG
CAACAGTACACAGCAGTTACAGTAACCCTCGGGTGACTCCAGTAACAGTGGCCAGCAGTG
GCCAGAGCTTTCGACTGACCCCATCAAGATGGTCCAGTGGTTTTAATAGTCCTCATGTTTC
TGCAGCAACAATATCCAGCAGTTATAGTAAGCCTCATGTGACTCTAGCAATAGTGCCCCAC
AATGCCCAAACCTTGCAACAGACCCCAGCAGCAATGACCAGTGGTTGTAACAGTCCTCAGA
TGTCTGCAGCAACAATATCCAGCAGTTATAATAACCCTCAGGTAACTCCAGCAACAGTGCC

AAGAAGTGCCCAGTCCCTTTTGATTCCTCAAGTAGCAGTGCCCAGCAGTGCCTTTCAAATG
ATCTTTCAAATGACCCCAGCAGCAATGGCCAATGGCTGTAACAGTCCTCAGATCTCTGCAG
CAACAGTACACAGCAATTACAGTAACCCTCGGGTGACTCCAGCAACAGTGGCCAGCAGTG
ACCAGAGCTATCGACTGACCCCATCAAGATGGTCCAGTGGTTTTAATAGTCCTCATGTTTCT
GCAGCAACAATACCCAGCAGTTACCAAAATTCCCCTCAGGTGTCTTCTGCAACAATGCGCA
GGAGTTTTCCTGCCATGTCCCTGTCTCCAGCAACACCACCCAAGAAACCAAGACTGAAGAA
AGTACCTAAACAACCTTCTGAAGAAAATGGTCACCAGCAAGGTTTCAAACAAGTGATGGTG
TTAAAAGTCACAGAGCCATTCACATATGACATTAGAAAAGACAAAAGAATGTTCCATGCCAC
CGTGGCTACGGAAACTGAATTCTTCAGAGTGAAGGTGTTTGACATAGCCCTCGAGGAAAAG
TTCATCCCCAGGAAAGTCATTGTCATCTCAGATTATATCGGCTACAATGGGTTTCTAGAGAT
ATACACAGCCTCCTGCGTATCAGAGGTGAACGACAGCAATGTGATGAATATCCCAACTTCC
CTGAGACAAAAAGCCAACGCAACACCTAAAATTAGTACTCTTTGTACACAAAGAGCAGGAA
CATTTGTGAATGGAACATTTACAGTATATGAGAAAACAGTGAGGAGTGAATTCATTTACTAT
GGAATAGAAGATAGAACGGGGAGGATGGAAGTGGTGGTCTACGGACAATTTACCAAACTG
TACTGTGAGCCTGGGGATAAGCTTAGCCTCTTCTGCTTTGAACTGAGCTCCAGTTTGGATA
TGTGGCAGCTGAGATCTGTGAGGCACAGTTATATGCAGGTCATCCAGGCTAGGAAGAGAG
ACTATGAACTTCCCAATGTTCATGCAATTATTGAAACAGCTCAAAGCTCACTCTCATGGATC
CCTGAACAACAATGGTAA

Gene name: *Pyhin-B*
Other names: BC094916
Open reading frame:
ATGGTGAATGAATACAAGAGAATTGTTCTTCTGACAGGATTAATGGGTATTAATGACCATGA
TTTTAGAATGGTTAAGTCCTTGTTGAGCAAAGAATTAAAACTAAATAAAATGCAAGATGAATA
CGACAGAGTTAAGATTGCTGATTTGATGGAAGACAAGTTCCCAAAAGATGCTGGAGTGGTC
CAACTGATAAAACTATATAAGCAGATTCCAGGACTTGGAGACATTGCTAATAAACTCAAAAA
TGAGAAGGCAAAAGCTAAAAGGAAAGGGAAAGGGAAACGGAAAACCGCAGCAAAAAGACA
AAGGCAAGAAGAACCCAGTACTTCCCAACCTATGTCCACCACAAATGAAGATGCAGAACCA
GAATCAGGGAGGAGTACACCTGACACACAGGTTGCTCAGTTATCTTTACCAACTGCTTCCC
GAAGGAACCAAGCCATTCAAATTTCTCCAACAATAGCATCCAGCAGTGGTCAGACCAGCAG
CAGATCTTCAGAAACATTACAAAGCATCATTCAGTCCCCCGAAACTCCAACAAGATCATCCA
GCAGGATTCTGGACCCTCCAGTGTCTCCAGGAACAGCATATAGCAGTGCCCAGGCACTTG
GAGTTCTTCTAGCAACACCAGCCAAGAGACAGAGACTGAAAAATGTACCCAAAGAACCTTC
TGAGGAAAATGGTTACCAGCAAGGTTCCAAAAAAGTGATGGTGTTAAAAGTAACAGAACCA
TTTGCATATGACATGAAAGGAGAGAAGATGTTCCATGCCACCGTGGCTACGGAAACAGAAT
TCTTTAGAGTAAAGGTGTTTGATATTGTCCTGAAGGAGAAGTTCATACCAAATAAGGTCCTT
ACCATCTCAAACTATGTTGGCTGCAATGGGTTCATAAATATATACAGTGCATCTAGTGTGTC
AGAGGTAAATGATGGCGAACCAATGAATATCCCACTTTCACTGAGGAAAAGCGCCAATCGA
ACACCTAAAATCAACTATCTTTGCTCAAAGAGAAGAGGAATATTTGTGAATGGAGTGTTTAC
TGTATGTAAGAAAGAAGAAAGGGGGTACTACATTTGCTATGAGATAGGAGATGATACAGGG
ATGATGGAAGTGGAGGTCTACGGACGACTGACCAATATTGCTTGTAATCCTGGAGATAAAC
TTAGACTCATGCTTTGA

Gene name: *Pyr-A*
Other names: Pydc4
Open reading frame:
ATGGTGAATTACTGCAAGCAAATTGTTCTTCTCTCTGGATTAGAATATATGAATGACTATAAC
TTTAGAGCACTTAAGTCCTTACTGAACCATGACCTAAAACTGACTAAGAACATGCAAGATGA
TTATGACAGAATTAACATTGCTGACTTAATGGAAGAAAAGTTCCCAGAGGATGCTGGATTGA
GCAAACTGATAGAAGTATGTGAAGACATTCCAGAACTTGCAGCTCGTGTTGATATACTTAGA

AAAGAGATGGAGAAAGTTAAAAATAAAACCAAAATAAAATCAGAATCCAGCCCTCCTCCACT
TACCTCCAGCCTGATGGAAGCTTGGGAAGTAGAACCAGCTATGGTGACAGCTTCATCAGA
GGTACACCGTCTAAGATCTCACTGA

Gene name: *Pyhin-1*
Other names: Ifi209; Ifix
Open reading frame:
ATGGTGAATGAATACAAGAGAATTGTTCTTCTGACAGGATTAATGGGTATTAATGACCATGA
TTTTAGAATGGTTAAGTCCTTGTTGAGCAAAGAACTAAAACTAAATAGAATGCAAGATCAGT
ACGACAGAGTTAAGATTGCTGATTTGATGGAAGACAAGTTCCCAAAAGATGCTGGAGTGGA
CCAACTGATAAAACTATATAAGCAGATTCCAGGACTTGGAGACATTGCTAATAAACTCAAAA
ATGAGAAGGCAAAAGCTAAAAGGACACGGACAGGGAAACGGAAAACCGCAGCAAAAAGAC
AAAGGCAAGAAGAACCCAGTACTTCCCAACCTATGTCCACCACAAATGAAGATGCAGAACC
AGAATCAGGGAGGAGTACACCTGACACACAGGTTGCTCAGTTATCTTTACCAACTGCTTCC
CGAAGGAACCAAGCCATTCAAATTTCTCCAACAATAGCATCCAGCAGTGGTCAGACCAGCA
GCAGATCTTCAGAAACATTACAAAGCATCATTCAGTCCCCCAAAACTCCAAAAAGACCGTC
CAGCAGTATTCTGGACCCTCCAGTGTCTTCAGGAACAGCATCCAGCAGTGCCCAGGCACT
AGGAGTTCCTCTAGCAACAATAGCCAAGAGACAGAGACTGAAAAATGTACCCAAAGAACCT
TCTGAGGAAAATGGTCACCAGCAAGGTTCCAAAAAAGTGATGGTGTTAAAAGTAACAGAAC
CATTTTCATATGACGTGACAGAAGAGAAGATGTTCCATGCCACCGTGGCTACGGAAACAGA
ATTCTTTAGAGTAAAGGTGTTTGACATTGTCCTGAAGGAGAAGTTCATACCAAATAAGGTCC
TTACCATCTCAAACTATGTTGGCTGCAATGGGTTCATAAATATATACAGTGCATCTAGTGTG
TCAGAGGTAAATGATGGCGAACCAATGAATATCCCACTTTCACTGAGGAAAAGTGCCAATC
GAACACCTAAAATCAACTATCTTTGCTCGAAGAAAAGAGGAATATTTGTGAATGGAGTGTTT
ACTGTATGTAAGAAAGAAGAAAAGTGGAACTACATTTGCTATGAGATAGGAGATGATACAG
GGATGATGGAAGTGGAGGTCTACGGACGACTGACCAATATTGCTTGTAATCCTGGAGATAA
ACTTAGACTCATATGCTTTAAACTGATCCCAGATGAAGAAAAAGCACAGCTGAGATCTACAA
CACACAGTAACATGCAGGTTATCAAAGCTAAGAATTGA

Gene name: *Pyr-rv1*
Other names: Pydc3
Open reading frame:
ATGGTGAATTACTACAAGCAAATTGTTCTTCTGTCTGGATTAGAATATATGAATGACTATAAC
TTTAGAGCACTTAAGTCCTTACTGAACCATGACCTAAAACTGACTAAGAACATGCAAGATGA
TTATGACAGAATTAAGATTGCTGACTTAATGGAAGAAAAGTTCCCAGAGGATGCTGGATTG
AGCAAACTGATAGAAGTATGTGAAGACATTCCAGAACTTGCAGCTCGTGTTGATATACTTAG
AAAAGAGATGGAGAAAGTTAAAAATAAAACCAAAATAAAATCGGAATCCAGTCCTCTCCCAC
TTACCTCCAGCCTGATGGAAGCTTGGGAAGTAGAACCAGCTATGGTGACAGCTTCATCAGA
GGAAAGCAAAGATACTATCCCTGAATCACCTGATACCATGACAACACAGTTTTTGGAGGAA
AAGCCTAAGTTTCCATTACTTTCAGCAACTAGCACAAGCCAGGCTGAGGGTGAACCACTGA
CTCCCCAGAGGTTTCCAACAACAGCCTCCAGCAGTCTCCAGACACCCCTGGAGCCTACAG
AAATATTGTCCACCATATTGGCAACATCTCAGGGTTCTTCAGCACCATATTCCACCTGTGAC
AAATCTTCTCGGGTGCCTCCAGTGACAGTCTCCAGCAGCCTCCAGACCATCCAGACCTGT
CAGGCAACCTCAACACTTCCTTGTAGTCACCACACTTCTCTGGAGTCTCCAAAAACAGAGC
CCAGCAGTGTCCAGGCCACTCAGATGACTCAAGCAATAAAGGCAAGTGGTCATAATTGTCC
TCAGGTGCCTGCATCAGCAGTATCCAGCAGTTTCATTAAACCTCAGGTGACCCCAGCAATG
TTACTCAGCGGTGTCCAAACTCCTTTGATGCCTCAAGCAACAGTGCCCAGCCGTGCCCAG
ACCTTTCAACTAACCCCAGCAAAAATGACCAGTGGTTGTAATAGTCCTCAGATGTCTGCAG
CAACAGTATACAGCAGTTACAGTAACCCTCAGGTGACTCCAGTGCACAGCAGTGTCCAGAT
CTTGCAAATGAACCTAGCAGCAATGACCATTGGTTGTAATAGTCCTCATGTATCTGCAGCAA

CAGTATCCAGCCCTTACAATAACCCTTGGGTGACTCCAGCAACATTTCCCCGAAATGCCCA
GACCTTGCAACTCTACCCAGCAGCCATGGCCCATGCCTGTAATAGTCCTCAGGTGTCTGCA
GCAACAATATCCAGCAGTTACAATAATACCCCTCAGGTGTCTTCTGTAACAGTGCCCAGGA
GTTTTCCTGCCATGTCTCTGTCTCCAGCAAAGCCACTCAAGGCAAAGATAGGATCTCATCT
GGGTGCTACAGACCAGCTGGTAAAGGATCACCAGTTTGAGGACAATGAAATGCAGAATCC
CCAGTCTGGATTAGGGACAGGCTTATCAGATCAGCCCCGCCTGTCCCGAACACAAAGCCG
CACAGAGAAGAGAAGGAGAAAGCAGGAAGAACAGCCGCTTGCAACCAAAATGAAGAGACT
CTCCCTATCTCCACCACAGACAACAGCAAGGAGAGTCAACCCGATGCCAACCTGGGGACA
ACTAAAGAAGCTGACTCGTGAGGCGGAGGGACTAGTACAGAGAACGGGCAACAAGCTTTC
ATCAGAGACAATGTTTTTGGCCATGCTTGCTTTAATAGCTATGCAGAGTTCTTCTGAGATTT
GTAACTACGACAAGGGAGGACATTCTGGAACTTCATCTGAAAAATAA

The last three coding exons of the murine *Pyr-rv1* ORF are an in-frame fusion of three distinct regions of ERVB2

PYR-RV1 exons 4-6 translated sequence:
AKIGSHLGATDQLVKDHQFEDNEMQNPQSGLGTGLSDQPRLSRTQSRTEKRRRKQEEQPLAT
KMKRLSLSPPQTTARRVNPMPTWGQLKKLTREAEGLVQRTGNKLSSETMFLAMLALIAMQSSS
EICNYDKGGHSGTSSEK*

ERVB2 “spliced” sequence, translated:
ARIGSHLGATDQLVKDHQFEDNEMQNPQSGLWTGLSDQPHLSRTQSCTEKRRKKKEEQPLAT
KMKRLSLSPPQTTARRVNPMPSWGQLKKLTREAEGLVHRTGNKLSSESMFLAMLALSSSEICN
YDKAGHSGTSSEK*

Gene name: *Pyhin-A*
Other names: AI607873
Open reading frame:
ATGGTGAAAGAATATAAAAGAATTGTTCTGCTGAAAGGACTTGAATGTATCAATAAGCATCA
TTTTAGCTTATTTAAGTCATTGCTGGCCAGAGATTTAAGTCTGGAAAGAGACAACCAAGAGA
AATACTCCACAATTCAGATTGCTAACATGATGGAAGAGAAATTCAAACCCGATGCTGGATTG
GGTGAACTGATCGAGTTTTGTGAAAAAGTACCAGCTCTTAGAAAACGTGCTGAAATTCTTAA
AAAAGAGAGATCAGAAGTAACAGGAGAAACATCACTGGAAAAAAATGGTCAAGAAGCAGGT
CCTGCAACACCTACATCAACTACAAGCCACATGCTAACATCTGAAAGAGGAGAGACTTCCG
CAACCCAGGAAGAGACTTCCACAGCTCAGGGGGAGACTTCCACAGCTCAGGCGGGGACT
TCCACAGCTCAGGCGGGGACTTCCACAGCTCAGGCGGGGACTTCTACAGCCCAGACAAA
GAGAAAACAAATCACAAAAAGTGAAGGTGAAAAGAAAAAGAAGCTCACCCAGGAGCAGGC
TCAGCTTTCAGAACCTTTAGGAACCGATAGAAAAAAAAGATGAAGATTGTCTCCAGACTCCTC
TTATGGCACCACCAACACCACCCAGCAGTTCCTCAAACAAGAAACAGAAAAACACAACCAT
CCCAAAACATGGTATCATAAAGACAAGGGGTCCTCAGGAAATTCATCAACTTGTAGAATTTT
CATCATCCAGCAACTTCCCAGCTGTAAGTGAACTCCCGACCTTTGAAGGACTTTCAGCAAT
TCCTTCTAGCAGACTTCAGTCTTCTCAGAAGCCTCTAGAAGCTCACTTGGATTTAAAGATGT
CTCCAAGCTCTTCAAGATCACCTTGCCATAATCTTTCAGTGTCTCTAACGTCAGACTCAAAT
GTTCATCTGAACTCTAATGCACATTCCATCCAGTCCAGTGGTGCCCAGGTTCCTTATGTAC
CATCAGCAACAGGATTCAGTAATGTCAGGGTTCCTCTTATGCCTTCAAAAACAGTGTCCTG
CTCTTCCATTGCTCCTCACATGTCTTTAATAAAGGTACCCAGCAGTATCCAGGACCTTCACC
TTCCTACACAAGAAGCATTCAGTAGCACAGAGGATCCTCACAATCAAGCAATAGCATCCAA
AAATCTCCAAACTACTAAAATGCCTCCAACAACACTGACAAGCAAAGCCCAGCTCCTAAAG
ATGCCTTCTTCAGCAACATCATCCAGCAGTGTTCAGGTTTCTCATACTCTAGCAACTGTGTC
CAGAAGCACCTCTGTCAAACAGATACTTCAAACAACAAGATCCAGTAGCATTCAGATTCTTA

ACTCTGCTACAGTAAAAGCATCCAAGAATGTCCAGGCCTCTCAAGTATCTTTACCAACAGAA
CCCAATTATTTCCAGGCTTCTGTGGCACCTCCATCAGCAACACCCAGCAGTTCTACATCTC
TTCTGGTACCTTTGTCAAAAGCCACAAGCAGAGCCCAGAGTACTCAAATTCATCCAGAAAG
AGAATCCATTTGTGTCCAGGCTCATCGTGCTCCTTCATCAACAGTGTCCAGAAATAAGTGC
ACCACTCAGTTGACACAAGGGGCAGCATCCGGTACTGGGAAGGCCTTTTCCCTCCCTGAA
GTAAAAGCATCCATGAAAGTCCAAGCACCTCAGGTGTCTTCCCCAACAGCATCCATGAGTA
TCCTGAATCCAAATGCAACCCCCCCAACAACATCCAGCAATCTCCTGGCTCCACATGCAAC
TTCATCAACAACATATAGCATCCTCCTGGCTCCATATGCAACTCTATCAACAGCATCTAGCA
ACCTCCTGGCTCCACAAGCAACTGTATCAACAGCATCTAGCAACCTCCTGGCTCCACATGC
AACTCTATCAACAACATCTAGCAACCTCCTGGCTCCACATGCAACTCTATCAACAACATCTA
GCAACCTCCTGGCTCCACATGCAACTTCATCAACAGCATCTAGCAACCTCCTGGCTCCACA
AGCAACTCTATCAACACCATCCAATAGTCTCCTGGCTCCACATGCAACTCTATCAACAACAT
CTAGCAACCTCCTGGCTCCACATGCAACTTCATCAACAGCATCTAGCAACCTCCTGGCTCC
ACATGCAACTTCATCAACAGCATCTAGCAACCTCCTGGCTCCACATGCAACTCTACCAACA
GCATCCAGCAATCTTCTGGCTCCACAGTTGTGTCCAGTAACAGCATCTAGGGCTCTCAGTG
CTATTCCAGTGCCTTCAGCAACAGTACACAGCAGTCCATCCTGGACACCAAGGAGAGGAA
CTGTACCAAAGGAGCCTTCTAGGGAAGAAGGTCACCATCAAGGTCCCAAACAAGTGATGG
TGCTGAAAGTAACAGAACCATTTACATATGACCTGGAAGAAGATAAAAGGATGTTTCATGCT
ACAGTGGCTACTGAAACTGAGTTCTTCAGAGTGAAGGTTTTTGACACAGCTCTAATAAGCA
AGTTCATCCCAAGAAATATCATTGCCATATCAGATTATTTTGGGTGCAATGGATTTTTGGAG
ATATACAGAGCTTCCTGTGTCTCTGATGTGAACGTTAACCCAACAATGGTCATCTCAAATAC
ACTGAGACAAAGAGCTAATGCAACTCCTAAAATTTCTTATCTTTTCTCACAAGCAAAGGGGA
CATTTGTGAGTGGAGAGTACTTAGTAATTAAGAAAACTGAGAGGAATAAAGCCATTTACTAT
GGAGTTAAAGATAATACAGGGAAAATGGAAGTGATGGTTTATGGACGACTCACCAATATCA
CGTGTGAACCAGGCAATAAACTTAGACTTGTCTGCTTTGAATTGAATCTCATTGAAGATGGG
TTGCAGCTGAAGTCTGTAAGGCACAGTTACATGCAGGTCATCAATGCTAGAAGGTGA

Gene name: *Ifi204*
Other names: Ifi16; p204
Open reading frame:
ATGGTGAATGAATACAAGAGAATTGTTCTGCTGAGAGGACTTGAATGTATCAATAAGCATTA
TTTTAGCTTATTTAAGTCATTGCTGGCCAGAGATTTAAATCTGGAAAGAGACAACCAAGAGC
AATACACCACGATTCAGATTGCTAACATGATGGAAGAGAAATTTCCAGCTGATTCTGGATTG
GGCAAACTGATTGAGTTTTGTGAAGAAGTACCAGCTCTTAGAAAACGAGCTGAAATTCTTAA
AAAAGAGAGATCAGAAGTAACAGGAGAAACATCACTGGAAAAAAATGGTCAAGAAGCAGGT
CCTGCAACACCTACATCAACTACAAGCCACATGTTAGCATCTGAAAGAGGCGAGACTTCTG
CAACCCAGGAAGAGACTTCCACAGCTCAGGCGGGGACTTCCACAGCTCAGGCGAGGACTT
CCACAGCTCAGGCGAGGACTTCCACAGCTCAGGCGAGGACTTCCACAGCTCAGGCGAGG
ACTTCCACAGCTCAGGCGGGGACTTCCACAGCCCAGAAAAGAAAAAGTATGAGAGAAGAA
GAGACTGGAGTGAAAAAGAGCAAGGCGGCTAAGGAACCAGATCAGCCTCCCTGTTGTGAA
GAACCCACAGCCATGTGCCAGTCACCAATACTCCACAGCTCATCTTCGGCTTCATCTAACA
TTCCTTCGGCTAAGAACCAAAAATCACAACCCCAGAATCAGAATATTCCCAGAGGTGCTGT
TCTCCACTCAGAGCCCCTGACAGTGATGGTGCTCACTGCAACAGACCCATTTGAATATGAA
TCACCAGAACATGAAGTAAAGAACATGCTTCATGCTACAGTGGCTACAGTGAGCCAGTATT
TCCATGTGAAAGTTTTCAACATCAACTTGAAAGAAAAGTTCACAAAAAAGAATTTTATCATCA
TATCCAATTACTTTGAGAGCAAAGGCATCCTGGAGATCAATGAGACTTCCTCTGTGTTAGA
GGCTGCTCCTGACCAAATGATTGAAGTGCCCAACAGTATTATCAGAAATGCAAATGCCAGT
CCTAAGATCTGTGATATTCAAAAGGGTACTTCTGGAGCAGTGTTCTATGGAGTGTTTACATT
ACACAAGAAAACAGTGAACCGAAAGAACACAATCTATGAAATAAAAGATGGTTCAGGAAGC
ATAGAAGTGGTGGGGAGTGGAAAATGGCACAACATCAACTGCAAGGAAGGAGATAAACTC

CACCTCTTCTGCTTTCACCTGAAAACAATTGACAGGCAACCAAAGTTAGTGTGTGGAGAAC
ACAGTTTCATCAAGATATCAAAGAGAGGAAATGTACCAAAGGAGCCTGCTAAGGAAGAAGA
TCACCATCATGGTCCCAAACAAGTGATGGTGCTGAAAGTAACAGAACCATTTACATATGAC
CTGAAAGAGGATAAAAGAATGTTTCATGCTACCGTGGCTACTGAAACTGAGTTCTTCAGAG
TGAAGGTTTTTGACACAGCTCTAAAGAGCAAGTTCATCCCAAGAAATATCATTGCCATATCA
GATTATTTTGGGTGCAATGGGTTTCTGGAGATATACAGAGCTTCCTGTGTCTCTGATGTGAA
CGTTAATCCAACAATGGTTATCTCAAATACACTGAGACAAAGAGCTAATGCAACTCCTAAAA
TTTCTTATCTTTTCTCACAAGCAAGGGGGACATTTGTGAGTGGAGAGTACTTAGTAAATAAG
AAAACGGAGAGGAATAAATTCATTTACTATGGAATTGGAGATGATACAGGGAAAATGGAAG
TGGTGGTTTATGGAAGACTCACCAATGTCAGGTGTGAACCAGGCAGTAAACTAAGACTTGT
CTGCTTTGAATTGACTTCCACTGAAGATGGGTGGCAGCTGAGGTCTGTAAGGCACAGTTAC
ATGCAGGTCATCAATGCTAGAAAGTGA

Gene name: *Mndal*
Other names:
Open reading frame:
ATGGCTGAATACAAGAAAATTGTTCTTCTAAAAGGATTAGAAAGCATGGAGGACTATCAGTT
TAGGACAGTCAAGTCCTTACTAAGAAAAGAACTAAAACTTACTAAAAAATTGCAAGAAGATT
ATGATAGGATTCAGCTGGCGGACTGGATGGAGGACAAGTTCCCCAAATATGCTGGACTAG
ACAAACTGATAAAAGTGTGTGAGCACATAAAAGATCTTAAAGACCTTGCTAAAAAACTTAAA
ACCGAGAAGGCAAAAGTTCAAAAGAAAAAGCAAGGAAAATGCAAAACGGCTGTGAAAAAAA
AAGGTCAAGATGAACTCAGTTCCTCCGAATCTCTTTTCATCAATAAAGAATCATACAAGAGT
GTACCGTCTTCAAAGAAAAAGGGAAAAGCGATCGCAAAAACTGAAGGTGAAAAAAAAAATA
AGCTCACCCAGGATCAGGATCATCTTCCAGAAACTTCAGGAACCGATATAAAAACAGAGGA
AGATTGTCTCCAGAATTCTCCTAAGCCTCCACCAACATCACCCAGCAGTTCCTCAAACAAG
AAAAAGAGAAAAGAGATCACAAAAACTGAAGGTGGAAAGAAAAAGAAGCTCACCCAGGAG
CAGGCTCAGCTTCCAGAACCTTTAGGAACTGATATAAAAAAGGATGAAGATTGCCTCCAGA
CTCCTCCTAAGCCTCCACCAACACCACCCAGCAGTTCCTTAAACAAGAAAAGAAAAAGTAG
GAGAGAAGAAGAGACTGGAGTGAAAAAGAGCAAGGCGGCTAAGGAACCAGATCAGCCTC
CCTGTTGTGAAGAACCCACAGCCAGGTGCCAGTCACCAATACTCCACAGCTCATCTTCAGC
TTCATCTAACATTCCTTCAGCTACGAACCAAAAACCACAACCCCAGAACCAGAACATTCCCA
GAGGTGCTGTTCTCCACTCAGAGCCCCTGACAGTGATGGTGCTCACTGCAACAGACCCGT
TTGAATATGAATCACCAGAACATGAAGTAAAGAACATGTTTCATGCTACAGTGGCTACAGTG
AGCCAGTATTTCCATGTGAAAGTTTTCAACATCAACTTGAAAGAGAAGTTCACAAAAAAGAA
TTTTATCATCATATCCAATTACTTTGAGAGCAAAGGCATCCTGGAGATCAATGAGACTTCCT
CTGTGTTAAAGGCTGATCCTGACCAAATGATTGAAGTGCCCAACAATATTATCAGAAATGCA
AATGCCAGTCCTAAGATCTGTGATATTCAAAAGGGTACTTCTGGAGCAGTGTTCTATGGAG
TGTTTACATTACACAAGAAAAAAGTGAAAACACAGAACACAAGCTATGAAATAAAAGATGGT
TCAGGAAGTATAGAAGTGGAGGGGAGTGGACAATGGCACAACATCAACTGTAAGGAAGGA
GATAAGCTCCACCTCTTCTGCTTTCACCTGAAAAGAGAAAGAGGACAACCAAAGTTAGTGT
GTGGAGACCACAGTTTCGTCAAGATCAAGGTCACCAAGGCTGGGAAAAAAAAGGAAGCAT
CAACTGTCCTGTCAAGCACAAAAAATGAAGAAGAAAATAATTACCCAAAAGATGGAATTAAG
GTAGAGATGCCAGACTATCACGTCTAA

Gene name: *Mnda*
Other names: Ifi205b
Open reading frame:
ATGGTGAATGAATACAAGAGAATTGTTCTGCTGAGAGGACTTGAATGTATCAATAAGCATTA
TTTTAGCTTATTTAAGTCATTGCTGGCCAGAGATTTAAATCTGGAAAGAGACAACCAAGAGC
AATACACCACGATTCAGATTGCTAACATGATGGAAGAGAAATTTCCAGCTGATTCTGGATTG

GGCAAACTGATTGAGTTTTGTGAAGAAGTACCAGCTCTTAGAAAACGAGCTGAAATTCTTAA
AAAAGAGAGATCAGAAGTAACAGGAGAAACATCACTGGAAAAAAATGGTCAAGAAGCAGGT
CCTGCAACACCTACATCAACTACAAGCCACATGTTAGCATCTGAAAGAGGCGAGACTTCTG
CAACCCAGGAAGAGACTTCCACAGCTCAGGCGGGGACTTCCACAGCTCAGGCGGGGACT
TCCACAGCTCAGGCGGGGACTTCCACAGCCCAGAAAAGAAAAAGTATGAGAGAAGAAGAG
ACTGGAGTGAAAAAGAGCAAGGCGGCTAAGGAACCAGATCAGCCTCCCTGTTGTGAAGAA
CCCACAGCCATGTGCCAGTCACCAATACTCCACAGCTCATCTTCGGCTTCATCTAACATTC
TTTCGGCTAAGAACCAAAAATCACAACCCCAGAACCAGAACATTCCCAGAGGTGCTGTTCT
CCACTCAGAGCCCCTGACAGTGATGGTGCTCACTGCAACAGACCCGTTTGAATATGAATCA
CCAGAACATGAAGTAAAGAACATGTTTCATGCTACAGTGGCTACAGTGAGCCAGTATTTCC
ATGTGAAAGTTTTCAACATCGATTTGAAAGAGAAGTTCACAAAAAATAATTTTATCACCATAT
CCAATTACTTTGAGAGCAAAGGCATCCTGGAGATCAATGAGACTTCCTCTGTGTTAGAGGC
TGCTCCTAAACAAATGATTGAAGTGCCCAACTGTATTACCAGAAATGCAAATGCCAGTCCTA
AGATCTGTGATATTCAAAAGGGTACTTCTGGAACAGTGTTCTATGGAGTGTTTACATTACAC
AAGAAAAAAGTGAAAACACAGAACACAAGCTATGAAATAAAAGATGGTTCAGGAAGCATAG
AAGTTGTGGGGAGTGGACAATGGCACAACATCAACTGTAAGGAAGGAGATAAGCTCCACC
TCTTCTGCTTTCACCTGAAAAGAGAAAGAGGACAACCAAAGTTAGTGTGTGGAGACCACAG
TTTCGTCAAGGTCACCAAGGCTGGGAAAAAAAAAGAAGCATCAACTGTCCAGTGA

Gene name: *Ifi203*
Other names:
Open reading frame:
ATGGCTGAATACAAGAATATTGTTCTTCTAAAAGGATTAGAGAACATGGAGGACTATCAGTT
TAGGACAGTCAAGTCCTTACTAAGAAAAGAACTAAAACTTACTAAAAAAATGCAAGAAGATT
ATGATAGGATTCAGCTGGCGGACTGGATTGAGGACAAGTTCCCCAAAGATGCTGGACTAG
ACAAACTGATAAAAGTGTGTGAGCACATAAAAGATCTTAAAGACCTTGCTAAAAAACTTAAA
ACAGAGAAGGCAAAAGTTCAAAAGAAAAAGGAAGGAAAATGCAAAACGGCTGGGAAAAAA
AAAGGGCAAGATGAACTCAGTTCTTCTGAATCTCTTTTCAACAATAAAGAATCAAACAAGAG
TGTACCGTCTTCAAAGAAAAAGAGAAAACAGATCACAAAAACTGAAGGTGGAAAGAAAAAG
AAGCTCACCCAGGAGCAGGCTCAGCTTCCAGAAACTTCAGGAACCAACATAAAAAAAGAG
GAAGATTGCCTCCAGAATCCTCATAAGTCTCCACCAACACCATCCAGCAGTTCCTCAAACA
AGGCACCAAGGAGAGGAACTGTACCAAAGGAGCCTTCTAGAGAAGAAGGTCACCATCAAG
GTCCCAAACAAGTGATGGTGTTGAAAGTAACAGAACCATTTACATATGACTTTGAAGAAACG
AAAAGGATGTTTCATGCTACAGTGGCTACTGAAACTGAGTTCTTCAGAGTGAAGGTTTTTGA
CACAGCTCTAATGAGCAAGTTCATCCCAGGAAAGATCATTGCCATATCACATTATATTGGGT
GCAATGGGTTTCTGGAGATATACAGAGCTTCCTGTGTGTCCGATGTGAACATTAACCCAAC
AATGATTATCTCAAATACACTGAGTGAAAGTGCTATTGCAACTCCAAAAATTTCTTATCTTCT
CTCACAAGCAAAGGGGACATTTGTAAATGGAGAATTTGTAGTATTTAAGAAAAGTGAGAGG
CATGAGTGCATTTGCTATGGAATTGGAGATGATACAGGAAAAATGGCAGTGGTGGTTTATG
GACGACTCACCAATGTCAGGTGTGAACCAGGCAGTAAACTTAGACTTGTCTGCTTTGAATT
GACTTCCACTAAAGATGTGTGCCTCCTGAGGTCTGTAAGGCACAGTTACATGCAGGTCATC
AATGAAGGAAAGCCACTCAACCCAGACTCAGTCAGGAGAAACTCTCTGGAACCATACTTCT
GA

Gene name: *Ifi202b*
Other names: Ifbip-1; Ifi202; Ifi202a; p202
Open reading frame:
ATGTCCAACCGTAACTTAAGGTCATCTACCAACTCAGAATTTTCTGAGGGTCAACATCAGAC
CCCTTCCAGTGATTCATCTGGCCATGGGGAGGATCAACCTCAAGCCTCTCCTGGACCTAAC
AAAAAGTCACACACCCCAAAAAAGAACATTAGCAAAGGTGCTGTTCTTCATGAGAAACCCA

TGACAGTGATGGTACTCACTGCAACAGAACCATTTAATTATAAAGAGGGAAAAGAGAACAT
GTTTCATGCTACAGTGGCTACAGAGAGCCAATATTACCGTGTGAAAGTTTTCAACATGGAC
TTGAAAGAGAAGTTCACAGAAAATAAATTTATTACCATTTCCAAATACTTCAACAGCAGTGG
CATCCTAGAGATCAATGAAACTGCCACTGTGTCAGAGGCTGCTCCTAACCAAATTATTGAA
GTGCCCAAAAATATTATCAGAAGTGCAAAAGAAACTCTTAAGATCTCTAAAATTAAAGAACT
TGATTCTGGAACACTGATTTATGGTGTGTTTGCAGTAGAGAAGAAAAAAGTGAATGATAAAA
GTATAACCTTCAAAATAAAAGATAATGAAGATAATATAAAAGTGGTGTGGGATAAAAAACAG
CACAATATCAACTATGAGAAAGGAGATAAACTCCAACTCTTCTCCTTTCACCTGAGAAAAGG
AAATGGGAAACCAATATTACACTCTGGAAATCACAGTTTCATCAAGGGAGAAAAGCTACTAA
AAGAATCTTTTGAAGGGGATGGTTACCACAAAGGTCCCAAACAAGTGGTGGCATTGAAAGC
AACAAAACTATTTACTTATGATAGTATAAAAAGTAAAAAGATGTTCCATGCCACAGTGGCTA
CTGATACAGAATTCTTCAGAGTGATGGTCTTCGAGGAAAACCTAGAGAAAAAGTTTATCCC
GGGAAACACCATTGCTTTATCAGATTATTTTGGTATGTATGGGTCTCTGGCAATACATGAAT
ATTCCAGCGTGTCTGAGGTGAAGAGCCAAAATAAGGAAGACTCAAGTTCATCAGATGAAAG
ACTCATAGAACATCTTAAAATTTGTGATCTTCACTTGCAAACAAAAGAAAGGCTGGTTGATG
GAGAGTTTAAAGTATACAGGAAAAGTACCGGAAATAATTGTATATGCTATGGAATTTGGGAT
GATACAGGAGCAATGAAAGTGGTGGTATCTGGACAACTGACCAGTGTCAACTGTGAGATTG
GTAATACAATTAGACTTGTCTGCTTTGAATTGACCTCAAATGCAGATGAGTGGTTTCTGAGA
TCTACGAGGTACAGTTACATGGAGGTCATCATGCCTGAAAAATGA

Gene name: *Ifi205*
Other names: D3; D3cDNA; Ifi205a
Open reading frame:
ATGGAGAATGAATATAAGAGACTTGTTCTGCTGGAAGGACTTGAATGTATCAATAAGCATCA
ATTCAATTTATTTAAGTCATTGATGGTCAAAGATTTAAATCTGGAAGAAGACAACCAAGAGA
AATATACCACGTTTCAGATTGCTAACATGATGGTAAAGAAATTTCCAGCTGATGCTGGATTG
GACAGACTGATCAACTTTTGTGAACGTGTACCAACTCTTAAAAAACGTGCTGAAATTCTTAA
AAAAGAGAGATCAGAAGTAACAGAAGAAACATCACTGGAAATAAATAGGCAAGAAGCAAGT
CCTGCAACACCTACATCAACTACAAGCCACATGTTAGCATCTGAAAGAGGCAAGACTTCCA
CAACCCAGGAAGAGACTTCCACAGCCCAGAAAAGGAAAGGTATGAGTGAAGAAAAGACTG
ACGTGAAGAAGATCAAGGCATCTGGGAAAGCAGATCAGCCTCCCTGTTGTGAAGGACCCA
CAGCCACATGCCAGTCACCAATATCCCAGGTCTCATCTTCGGCTTCATCTAACATTCCTTC
GGCTAAGAACCAAAAATCACAACCCCAGAACCAGAACATTCCCAGAGGTGCTGTTCTCCAC
TCAGAGCCCCTGACAGTGATGGTGCTCACTGCAACAGACCCGTTTGAATATGAATCACCAG
AACATGAAGTAAAGAACATGTTTCATGCTACAGTGGCTACAGTGAGCCAGTATTTCCATGT
GAAAGTTTTCAACATCGATTTGAAAGAGAAGTTCACAAAAAATAATTTTATCACCATATCCAA
TTACTTTGAGAGCAAAGGCATCCTGGAGATCAATGAGACTTCCTCTGTGTTAGAGGCTGCT
CCTAAACAAATGATTGAAGTGCCCAACTGTATTACCAGAAATGCAAATGCCAGTCCTAAGAT
CTGTGATATTCAAAAGGGTACTTCTGGAACAGTGTTCTATGGAGTGTTTACATTACACAAGA
AAAAAGTGAAAACACAGAACACAAGCTATGAAATAAAAGATGGTTCAGGAAGCATAGAAGT
GGTGGGGAGTGGACAATGGCACAACATCAACTGTAAGGAAGGAGATAAGCTCCACCTCTT
CTGCTTTCACCTGAAAAGAGAAAGAGGACAACCAAAGTTAGTGTGTGGAGACCACAGTTTC
GTCAAGGTCACCAAGGCTGGGAAAAAAAAAGAAGCATCAACTGTCCAGTGA

Table 3.2 Primers for quantitative RT-PCR analysis of murine ALR expression

Gene name	Sense Q-PCR Primer	Antisense Q-PCR Primer	Product Size (bp)	Efficiency (%)	R^2 value
mAIM2	GTTGAATCTAACCACGAAGTCC	CTACAAGGTCCAGATTTCAACTG	301	104.30	0.9745
mPYBLHIN-C	CGACAGCAATGTGATGAATATCC	TTCCTAGCCTGGATGACCTGC	329	101.38	0.9998
mPYHIN-B	CCAACAAGATCATCCAGCAGG	TTCTCTCCTTTCATGTCATATGC	223	110.40	0.9890
mPYR-A	TAAGTCCTTACTGAACCATGACC	GTGAGATCTTAGACGGTGTACC	318	126.00	0.9980
mPYHIN-1	CAGAGGTAAATGATGGCGAACC	GTGTTGTAGATCTCAGCTGTGC	308	119.30	0.9360
mPYR-RV1	CTCCAGCAAAGCCACTCAAGG	CCTCACGAGTCAGCTTCTTTAG	302	107.60	0.9835
mPYHIN-A	GTACACAGCAGTCCATCCTGG	AGATGACCATTGTTGGGTTAACG	336	104.40	0.9870
mIFI204	TGTGTGGAGAACACAGTTTCATC	GAGATAACCATTGTTGGATTAACG	342	98.20	0.9850
mMNDAL	ACCCAGCAGTTCCTTAAACAAG	TTCAATCATTTGGTCAGGATCAG	498	111.30	0.9935
mMNDA	TCTGAAAGAGGCGAGACTTCTG	TCTGGGAATGTTCTGGTTCTGG	311	98.20	0.9815
mIFI203	CACTGAGTGAAAGTGCTATTGC	GATGACCTGCATGTAACTGTGC	289	99.40	0.9890
mIFI202b	GGAAGACTCAAGTTCATCAGATG	GTAACTGTACCTCGTAGATCTC	291	112.85	0.9835
mIFI205	CAGTCACCAATATCCCAGGTC	CTTCAATCATTTGTTTAGGAGCAG	375	102.15	0.9975

Section 4: Final Discussion

The work presented here furthers our knowledge of ISD pathway sensors and negative regulators in antiviral defense and autoimmunity. In summary, we identified and confirmed 6 novel Trex1 interacting proteins by yeast 2-hybrid screening (Figure 2.3). However, it is still unclear if these proteins are involved in interferon stimulatory DNA (ISD) sensing, retroelement silencing or another unappreciated cellular process in which Trex1 has a role. Interestingly, we found that all Trex1 interacting proteins were nuclear prompting us to re-evaluate the localization of Trex1 (Fig. 2.5). Previous studies identified Trex1 localization to the endoplasmic reticulum (ER), but we discovered that Trex1 is also found on the inner nuclear membrane (INM) in localization studies with emerin (Fig. 2.4) [159]. Although Trex1 is found both on the ER and the INM, it remains to be determined if these pools of Trex1 both participate in endogenous retroelement silencing or serve distinct functions. Further studies are needed to determine the precise roles of the 6 Trex1 interacting proteins in either ISD sensing or retroelement silencing.

We also characterized a highly divergent family of AIM2-like receptor (ALRs). Upon closer examination of the ALR locus in mice, we found 5 novel genes in addition to the 7 previously described genes (Fig. 3.6). Evolutionary analysis of ALRs across mammals suggests that this locus has expanded and contracted in a species-specific manner (Fig. 3.6). AIM2 is the only gene with orthologous sequence and presumably function across species. Functionally, the murine and human ALRs seemed to fall into two categories, activators of STING-dependent type I IFN and ASC/caspase-1-dependent inflammasome activators. The one exception to this is human AIM2 that activated both responses (Fig. 3.12). In agreement with functional data, each

murine ALR localized preferentially to either STING or ASC depending on its ability to activate type I IFN or IL-1β, respectively (Fig. 3.10).

It is unusual for innate immune receptors to highly diverge between species. Toll-like receptor-like proteins are present from nematode worms through higher vertebrates [294]. The highly species-specific expansion and contraction of the AIM2-like receptor loci may reflect the evolutionary struggle between DNA viruses and their hosts. Since many DNA viruses, especially the herpesviruses, establish latent, lifetime infections, they are exquisitely adapted for living within a specific host and likely subvert the host's ability to induce type I IFN after DNA sensing [295]. In response to this evolutionary pressure, hosts may have adapted to infection with viruses by diversifying their DNA sensing receptors. Mammalian hosts may have evolved more DNA sensing receptors and/or changed the coding sequences of existing receptors to recognize genome from specific DNA viruses the host encounters. This could explain why mammals lack orthologous ALRs other than AIM2, and also why it has been difficult to definitively identify STING-dependent intracellular DNA sensors.

In addition to species-specific diversity in DNA sensor proteins, our data suggests that the initial DNA sensing of viral genomes takes place in the nucleus. The detection of viral RNA in the cytoplasm of cells led to many studies trying to identify cytosolic DNA sensor proteins. AIM2 was successfully identified using this approach as a cytosolic DNA binding protein that activates the inflammasome to induce the processing and secretion of mature IL-1β [126-129]. We also utilized this approach to discover cytosolic DNA binding proteins that activate the ISD pathway without success (Fig. 3.3 and 3.4). In addition, self and non-self discrimination by DNA sensors is thought to occur by compartmentalization in which detection of viral DNA occurs by sensors in the cytosol [239]. However, our findings imply that DNA sensing can take place in the

nucleus. With the exception of AIM2, all of the ALRs are expressed in the nucleus under steady-state conditions (Fig. 3.10). We also discovered that Trex1 interacts with the nuclear proteins Brd7, Btf3l4, Chaf1a, Pias1, Rfc2 and Sae2. We re-examined the localization of Trex1 and discovered that in addition to the endoplasmic reticulum (ER) Trex1 is localized to the inner nuclear membrane implying that negative regulation of the ISD pathway could take place in the nucleus. Presumably, ISD detection also initially occurs in the nucleus, but activation of signaling would occur in the cytosol since STING and ASC are both cytosolic proteins. Our data supports this model. Only upon overexpression with either STING and/or ASC do the murine ALRs localize to the cytosol to interact with these adaptor proteins (Fig. 3.10).

Based on the lifecycles of DNA viruses, nuclear sensing seems plausible. Upon entry, all DNA viruses, except for poxviruses that replicate in the cytosol using their own replication machinery, need to deliver their genomes to the nucleus for viral gene expression and replication [296]. Viruses have 5 major ways of delivering their genomes to the nucleus. Viruses deliver their genomes to the nucleus during mitosis-induced disruption of the nuclear envelope (murine leukemia retrovirus), by disrupting the nuclear envelope and lamina with viral proteins (parvovirus), cytoplasmic disassembly and nuclear pore entry through nuclear localization signals (NLS) on viral proteins (HIV), capsid binding to nuclear pore complexes and genome insertion through the pore (herpesviruses and adenoviruses) and lastly capsid diffusion through the nuclear pore for small DNA viruses (hepatitis B virus and baculoviruses) [297]. With the exception of HIV infection, viral genomic DNA is not exposed until the final uncoating or insertion into the nucleus. Therefore, the initial detection of DNA viruses could take place in the nucleus, although a cytosolic DNA detection system would also be necessary to recognize viruses that uncoat or replicate in the cytoplasm.

If viral DNA sensing occurs in the nucleus it is unclear how viral DNA would be distinguished from host genome. One potential mechanism could be the differing structures of host and virus DNA. Host DNA is associated with histone proteins that are organized into nucleosome units that serve to condense the genome and aid in regulation of gene expression [298]. Viral DNA genomes are not associated with histone proteins when packaged into virions. However, upon entry into the nucleus the genomes of adenovirus, herpes simplex virus, Epstein-Barr virus and Kaposi's sarcoma-associated herpesvirus all associate with histone proteins as early as 6 hours post-infection as is the case with adenovirus [250]. Nuclear DNA sensors would have a narrow window to recognize this foreign DNA that is not associated with chromatin to activate type I IFN and could explain why host DNA does not normally activate the ISD pathway. Conversely, it is possible that distinct histones could "mark" foreign DNA. We found histone H1 and H2 variants specifically upregulated in the ISD treated cells by microarray (Fig. 3.1). Although we have no evidence to support active "marking" of viral DNA, we can speculate that some ISD sensors might recognize viral DNA associated with a specific antiviral histone variant.

Another potential mechanism for specific recognition of viral DNA could be the localization of ISD sensors within the nucleus. Based on the methods of viral DNA entry into the nucleus discussed previously, the nuclear periphery would be an ideal place for recognition. ISD sensors could be associated with the inner nuclear membrane (INM) in close proximity to Trex1 (Fig. 2.4) or part of the nuclear pore complex to act as sentinels for recognition of foreign DNA entering the nucleus. In addition to nuclear sensing of viral DNA, we also re-examined the localization of Trex1 following discovery of nuclear interacting proteins. Since Trex1 localizes to the INM and the endoplasmic reticulum it is possible that the nuclear pool of Trex1 is actually

responsible for metabolizing retroelement DNA. Nuclear localization of Trex1 is also in line with the implication that DNA sensing occurring in the nucleus and places Trex1 in an ideal location near the sensors themselves to negatively regulate this pathway. Nuclear sensing of viral DNA must be tightly regulated and silenced during normal cellular replication to prevent improper activation of the ISD pathway and autoimmune disease.

References

1. Breitbart, M. and F. Rohwer, *Here a virus, there a virus, everywhere the same virus?* Trends Microbiol, 2005. **13**(6): p. 278-84.
2. Bergh, O., et al., *High abundance of viruses found in aquatic environments.* Nature, 1989. **340**(6233): p. 467-8.
3. Wommack, K.E. and R.R. Colwell, *Virioplankton: viruses in aquatic ecosystems.* Microbiology and Molecular Biology Reviews, 2000. **64**(1): p. 69-114.
4. Heldal, M. and G. Bratbak, *Production and decay of viruses in aquatic environments.* Marine Ecology Progress Series, 1991. **72**(3): p. 205-212.
5. Steward, G.F., et al., *Estimation of virus production in the sea. II. Field results.* Marine Microbial Food Webs 1992. **6**: p. 57-78.
6. Hammad, A.M.M., *Evaluation of alginate-encapsulated Azotobacter chroococcum as a phage-resistant and an effective inoculum.* Journal of Basic Microbiology, 1998. **38**(1): p. 9-16.
7. Hanlon, G.W., et al., *Reduction in exopolysaccharide viscosity as an aid to bacteriophage penetration through Pseudomonas aeruginosa biofilms.* Appl Environ Microbiol, 2001. **67**(6): p. 2746-53.
8. Castillo, F.J. and P.F. Bartell, *Studies on the bacteriophage 2 receptors of Pseudomonas aeruginosa.* J Virol, 1974. **14**(4): p. 904-9.
9. Nordstrom, K. and A. Forsgren, *Effect of protein A on adsorption of bacteriophages to Staphylococcus aureus.* J Virol, 1974. **14**(2): p. 198-202.
10. Bertani, G. and J.J. Weigle, *Host controlled variation in bacterial viruses.* J Bacteriol, 1953. **65**(2): p. 113-21.
11. Tock, M.R. and D.T. Dryden, *The biology of restriction and anti-restriction.* Curr Opin Microbiol, 2005. **8**(4): p. 466-72.
12. Labrie, S.J., J.E. Samson, and S. Moineau, *Bacteriophage resistance mechanisms.* Nat Rev Microbiol, 2010. **8**(5): p. 317-27.
13. Godde, J.S. and A. Bickerton, *The repetitive DNA elements called CRISPRs and their associated genes: evidence of horizontal transfer among prokaryotes.* J Mol Evol, 2006. **62**(6): p. 718-29.
14. Kunin, V., R. Sorek, and P. Hugenholtz, *Evolutionary conservation of sequence and secondary structures in CRISPR repeats.* Genome Biol, 2007. **8**(4): p. R61.
15. Grissa, I., G. Vergnaud, and C. Pourcel, *The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats.* BMC Bioinformatics, 2007. **8**: p. 172.
16. Jansen, R., et al., *Identification of a novel family of sequence repeats among prokaryotes.* OMICS, 2002. **6**(1): p. 23-33.
17. Mojica, F.J., et al., *Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria.* Mol Microbiol, 2000. **36**(1): p. 244-6.
18. van der Ploeg, J.R., *Analysis of CRISPR in Streptococcus mutans suggests frequent occurrence of acquired immunity against infection by M102-like bacteriophages.* Microbiology, 2009. **155**(Pt 6): p. 1966-76.
19. Barrangou, R., et al., *CRISPR provides acquired resistance against viruses in prokaryotes.* Science, 2007. **315**(5819): p. 1709-12.

20. Deveau, H., et al., *Phage response to CRISPR-encoded resistance in Streptococcus thermophilus.* J Bacteriol, 2008. **190**(4): p. 1390-400.
21. Horvath, P., et al., *Diversity, activity, and evolution of CRISPR loci in Streptococcus thermophilus.* J Bacteriol, 2008. **190**(4): p. 1401-12.
22. Marraffini, L.A. and E.J. Sontheimer, *CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea.* Nat Rev Genet, 2010. **11**(3): p. 181-90.
23. Sorek, R., V. Kunin, and P. Hugenholtz, *CRISPR--a widespread system that provides acquired resistance against phages in bacteria and archaea.* Nat Rev Microbiol, 2008. **6**(3): p. 181-6.
24. Karginov, F.V. and G.J. Hannon, *The CRISPR system: small RNA-guided defense in bacteria and archaea.* Mol Cell, 2010. **37**(1): p. 7-19.
25. Hamilton, A.J. and D.C. Baulcombe, *A species of small antisense RNA in posttranscriptional gene silencing in plants.* Science, 1999. **286**(5441): p. 950-2.
26. Silhavy, D. and J. Burgyan, *Effects and side-effects of viral RNA silencing suppressors on short RNAs.* Trends Plant Sci, 2004. **9**(2): p. 76-83.
27. Wang, M.B. and M. Metzlaff, *RNA silencing and antiviral defense in plants.* Curr Opin Plant Biol, 2005. **8**(2): p. 216-22.
28. Zambon, R.A., V.N. Vakharia, and L.P. Wu, *RNAi is an antiviral immune response against a dsRNA virus in Drosophila melanogaster.* Cell Microbiol, 2006. **8**(5): p. 880-9.
29. Wang, X.H., et al., *RNA interference directs innate immunity against viruses in adult Drosophila.* Science, 2006. **312**(5772): p. 452-4.
30. van Rij, R.P., et al., *The RNA silencing endonuclease Argonaute 2 mediates specific antiviral immunity in Drosophila melanogaster.* Genes Dev, 2006. **20**(21): p. 2985-95.
31. Galiana-Arnoux, D., et al., *Essential function in vivo for Dicer-2 in host defense against RNA viruses in drosophila.* Nat Immunol, 2006. **7**(6): p. 590-7.
32. Lu, R., et al., *Animal virus replication and RNAi-mediated antiviral silencing in Caenorhabditis elegans.* Nature, 2005. **436**(7053): p. 1040-3.
33. Wilkins, C., et al., *RNA interference is an antiviral defence mechanism in Caenorhabditis elegans.* Nature, 2005. **436**(7053): p. 1044-7.
34. Schott, D.H., et al., *An antiviral role for the RNA interference machinery in Caenorhabditis elegans.* Proc Natl Acad Sci U S A, 2005. **102**(51): p. 18420-4.
35. Hammond, T.M., et al., *Aspergillus mycoviruses are targets and suppressors of RNA silencing.* Eukaryot Cell, 2008. **7**(2): p. 350-7.
36. Segers, G.C., et al., *Evidence that RNA silencing functions as an antiviral defense mechanism in fungi.* Proc Natl Acad Sci U S A, 2007. **104**(31): p. 12902-6.
37. Obbard, D.J., et al., *The evolution of RNAi as a defence against viruses and transposable elements.* Philos Trans R Soc Lond B Biol Sci, 2009. **364**(1513): p. 99-115.
38. Ding, S.W. and O. Voinnet, *Antiviral immunity directed by small RNAs.* Cell, 2007. **130**(3): p. 413-26.
39. Tolia, N.H. and L. Joshua-Tor, *Slicer and the argonautes.* Nat Chem Biol, 2007. **3**(1): p. 36-43.
40. Baulcombe, D., *RNA silencing.* Curr Biol, 2002. **12**(3): p. R82-4.
41. Carrington, J.C. and V. Ambros, *Role of microRNAs in plant and animal development.* Science, 2003. **301**(5631): p. 336-8.
42. Voinnet, O., *RNA silencing: small RNAs as ubiquitous regulators of gene expression.* Curr Opin Plant Biol, 2002. **5**(5): p. 444-51.

43. Plasterk, R.H., *RNA silencing: the genome's immune system.* Science, 2002. **296**(5571): p. 1263-5.
44. Haasnoot, J., E.M. Westerhout, and B. Berkhout, *RNA interference against viruses: strike and counterstrike.* Nat Biotechnol, 2007. **25**(12): p. 1435-43.
45. Cullen, B.R., *Is RNA interference involved in intrinsic antiviral immunity in mammals?* Nat Immunol, 2006. **7**(6): p. 563-7.
46. Akira, S., S. Uematsu, and O. Takeuchi, *Pathogen recognition and innate immunity.* Cell, 2006. **124**(4): p. 783-801.
47. Robinson, M.J., et al., *Myeloid C-type lectins in innate immunity.* Nat Immunol, 2006. **7**(12): p. 1258-65.
48. Brown, G.D., *Dectin-1: a signalling non-TLR pattern-recognition receptor.* Nat Rev Immunol, 2006. **6**(1): p. 33-43.
49. Inohara, et al., *NOD-LRR proteins: role in host-microbial interactions and inflammatory disease.* Annu Rev Biochem, 2005. **74**: p. 355-83.
50. Philpott, D.J. and S.E. Girardin, *Nod-like receptors: sentinels at host membranes.* Curr Opin Immunol, 2010. **22**(4): p. 428-34.
51. Faurschou, M. and N. Borregaard, *Neutrophil granules and secretory vesicles in inflammation.* Microbes Infect, 2003. **5**(14): p. 1317-27.
52. Underhill, D.M. and A. Ozinsky, *Phagocytosis of microbes: complexity in action.* Annu Rev Immunol, 2002. **20**: p. 825-52.
53. Lanier, L.L., *NK cell recognition.* Annu Rev Immunol, 2005. **23**: p. 225-74.
54. Anthony, R.M., et al., *Protective immune mechanisms in helminth infection.* Nat Rev Immunol, 2007. **7**(12): p. 975-87.
55. Guermonprez, P., et al., *Antigen presentation and T cell stimulation by dendritic cells.* Annu Rev Immunol, 2002. **20**: p. 621-67.
56. Berland, R. and H.H. Wortis, *Origins and functions of B-1 cells with notes on the role of CD5.* Annu Rev Immunol, 2002. **20**: p. 253-300.
57. Hansen, T.H., et al., *Patterns of nonclassical MHC antigen presentation.* Nat Immunol, 2007. **8**(6): p. 563-8.
58. Janeway, C.A., Jr., B. Jones, and A. Hayday, *Specificity and function of T cells bearing gamma delta receptors.* Immunol Today, 1988. **9**(3): p. 73-6.
59. Medzhitov, R., *Recognition of microorganisms and activation of the immune response.* Nature, 2007. **449**(7164): p. 819-26.
60. Isaacs, A. and J. Lindenmann, *Virus interference. I. The interferon.* Proc R Soc Lond B Biol Sci, 1957. **147**(927): p. 258-67.
61. Bonjardim, C.A., P.C. Ferreira, and E.G. Kroon, *Interferons: signaling, antiviral and viral evasion.* Immunol Lett, 2009. **122**(1): p. 1-11.
62. Darnell, J.E., Jr., I.M. Kerr, and G.R. Stark, *Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins.* Science, 1994. **264**(5164): p. 1415-21.
63. van Boxel-Dezaire, A.H., M.R. Rani, and G.R. Stark, *Complex modulation of cell type-specific signaling in response to type I interferons.* Immunity, 2006. **25**(3): p. 361-72.
64. Stark, G.R., et al., *How cells respond to interferons.* Annu Rev Biochem, 1998. **67**: p. 227-64.
65. Garcia, M.A., et al., *Impact of protein kinase PKR in cell biology: from antiviral to antiproliferative action.* Microbiol Mol Biol Rev, 2006. **70**(4): p. 1032-60.

66. Kristiansen, H., et al., *The oligoadenylate synthetase family: an ancient protein family with multiple antiviral activities.* J Interferon Cytokine Res, 2011. **31**(1): p. 41-7.
67. Haller, O., P. Staeheli, and G. Kochs, *Interferon-induced Mx proteins in antiviral host defense.* Biochimie, 2007. **89**(6-7): p. 812-8.
68. van den Broek, M.F., et al., *Antiviral defense in mice lacking both alpha/beta and gamma interferon receptors.* J Virol, 1995. **69**(8): p. 4792-6.
69. Kaisho, T. and S. Akira, *Toll-like receptor function and signaling.* J Allergy Clin Immunol, 2006. **117**(5): p. 979-87; quiz 988.
70. Kawai, T. and S. Akira, *TLR signaling.* Cell Death Differ, 2006. **13**(5): p. 816-25.
71. 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.
72. Edelmann, K.H., et al., *Does Toll-like receptor 3 play a biological role in virus infections?* Virology, 2004. **322**(2): p. 231-8.
73. 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.
74. Hoebe, K., et al., *Identification of Lps2 as a key transducer of MyD88-independent TIR signalling.* Nature, 2003. **424**(6950): p. 743-8.
75. Hardarson, H.S., et al., *Toll-like receptor 3 is an essential component of the innate stress response in virus-induced cardiac injury.* Am J Physiol Heart Circ Physiol, 2007. **292**(1): p. H251-8.
76. Wang, T., et al., *Toll-like receptor 3 mediates West Nile virus entry into the brain causing lethal encephalitis.* Nat Med, 2004. **10**(12): p. 1366-73.
77. Gowen, B.B., et al., *TLR3 deletion limits mortality and disease severity due to Phlebovirus infection.* J Immunol, 2006. **177**(9): p. 6301-7.
78. Le Goffic, R., et al., *Detrimental contribution of the Toll-like receptor (TLR)3 to influenza A virus-induced acute pneumonia.* PLoS Pathog, 2006. **2**(6): p. e53.
79. Hornung, V., et al., *RNA recognition via TLR7 and TLR8.* Handb Exp Pharmacol, 2008(183): p. 71-86.
80. 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.
81. Colonna, M., G. Trinchieri, and Y.J. Liu, *Plasmacytoid dendritic cells in immunity.* Nat Immunol, 2004. **5**(12): p. 1219-26.
82. Gorden, K.B., et al., *Synthetic TLR agonists reveal functional differences between human TLR7 and TLR8.* J Immunol, 2005. **174**(3): p. 1259-68.
83. 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.
84. 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.
85. Lee, H.K., et al., *Autophagy-dependent viral recognition by plasmacytoid dendritic cells.* Science, 2007. **315**(5817): p. 1398-401.
86. Wang, J.P., et al., *Cutting Edge: Antibody-mediated TLR7-dependent recognition of viral RNA.* J Immunol, 2007. **178**(6): p. 3363-7.
87. 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.

88. Hemmi, H., et al., *A Toll-like receptor recognizes bacterial DNA.* Nature, 2000. **408**(6813): p. 740-5.
89. 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.
90. 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.
91. Krug, A., et al., *Herpes simplex virus type 1 activates murine natural interferon-producing cells through toll-like receptor 9.* Blood, 2004. **103**(4): p. 1433-7.
92. Fiola, S., et al., *TLR9 contributes to the recognition of EBV by primary monocytes and plasmacytoid dendritic cells.* J Immunol, 2010. **185**(6): p. 3620-31.
93. Lima, G.K., et al., *Toll-like receptor (TLR) 2 and TLR9 expressed in trigeminal ganglia are critical to viral control during herpes simplex virus 1 infection.* Am J Pathol, 2010. **177**(5): p. 2433-45.
94. Wilkins, C. and M. Gale, Jr., *Recognition of viruses by cytoplasmic sensors.* Curr Opin Immunol, 2010. **22**(1): p. 41-7.
95. 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.
96. Kawai, T., et al., *IPS-1, an adaptor triggering RIG-I- and Mda5-mediated type I interferon induction.* Nat Immunol, 2005. **6**(10): p. 981-8.
97. Xu, L.G., et al., *VISA is an adapter protein required for virus-triggered IFN-beta signaling.* Mol Cell, 2005. **19**(6): p. 727-40.
98. 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.
99. Yoneyama, M. and T. Fujita, *RNA recognition and signal transduction by RIG-I-like receptors.* Immunol Rev, 2009. **227**(1): p. 54-65.
100. Hornung, V., et al., *5'-Triphosphate RNA is the ligand for RIG-I.* Science, 2006. **314**(5801): p. 994-7.
101. Pichlmair, A., et al., *RIG-I-mediated antiviral responses to single-stranded RNA bearing 5'-phosphates.* Science, 2006. **314**(5801): p. 997-1001.
102. Schmidt, A., et al., *5'-triphosphate RNA requires base-paired structures to activate antiviral signaling via RIG-I.* Proc Natl Acad Sci U S A, 2009. **106**(29): p. 12067-72.
103. Takahasi, K., et al., *Nonself RNA-sensing mechanism of RIG-I helicase and activation of antiviral immune responses.* Mol Cell, 2008. **29**(4): p. 428-40.
104. 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.
105. 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.
106. Loo, Y.M., et al., *Distinct RIG-I and MDA5 signaling by RNA viruses in innate immunity.* J Virol, 2008. **82**(1): p. 335-45.
107. Roth-Cross, J.K., S.J. Bender, and S.R. Weiss, *Murine coronavirus mouse hepatitis virus is recognized by MDA5 and induces type I interferon in brain macrophages/microglia.* J Virol, 2008. **82**(20): p. 9829-38.

108. Fredericksen, B.L. and M. Gale, Jr., *West Nile virus evades activation of interferon regulatory factor 3 through RIG-I-dependent and -independent pathways without antagonizing host defense signaling.* J Virol, 2006. **80**(6): p. 2913-23.
109. 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.
110. Rothenfusser, S., et al., *The RNA helicase Lgp2 inhibits TLR-independent sensing of viral replication by retinoic acid-inducible gene-I.* J Immunol, 2005. **175**(8): p. 5260-8.
111. Komuro, A. and C.M. Horvath, *RNA- and virus-independent inhibition of antiviral signaling by RNA helicase LGP2.* J Virol, 2006. **80**(24): p. 12332-42.
112. Venkataraman, T., et al., *Loss of DExD/H box RNA helicase LGP2 manifests disparate antiviral responses.* J Immunol, 2007. **178**(10): p. 6444-55.
113. Stetson, D.B. and R. Medzhitov, *Recognition of cytosolic DNA activates an IRF3-dependent innate immune response.* Immunity, 2006. **24**(1): p. 93-103.
114. Ishikawa, H. and G.N. Barber, *STING is an endoplasmic reticulum adaptor that facilitates innate immune signalling.* Nature, 2008. **455**(7213): p. 674-8.
115. Zhong, B., et al., *The adaptor protein MITA links virus-sensing receptors to IRF3 transcription factor activation.* Immunity, 2008. **29**(4): p. 538-50.
116. Sun, W., et al., *ERIS, an endoplasmic reticulum IFN stimulator, activates innate immune signaling through dimerization.* Proc Natl Acad Sci U S A, 2009. **106**(21): p. 8653-8.
117. Jin, L., et al., *MPYS, a novel membrane tetraspanner, is associated with major histocompatibility complex class II and mediates transduction of apoptotic signals.* Mol Cell Biol, 2008. **28**(16): p. 5014-26.
118. 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.
119. Tanaka, Y. and Z.J. Chen, *STING specifies IRF3 phosphorylation by TBK1 in the cytosolic DNA signaling pathway.* Sci Signal, 2012. **5**(214): p. ra20.
120. Sharma, S., et al., *Innate immune recognition of an AT-rich stem-loop DNA motif in the Plasmodium falciparum genome.* Immunity, 2011. **35**(2): p. 194-207.
121. 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.
122. 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.
123. Ishii, K.J., et al., *A Toll-like receptor-independent antiviral response induced by double-stranded B-form DNA.* Nat Immunol, 2006. **7**(1): p. 40-8.
124. Owen, T.J., et al., *Epstein-Barr virus-encoded EBNA1 enhances RNA polymerase III-dependent EBER expression through induction of EBER-associated cellular transcription factors.* Mol Cancer, 2010. **9**: p. 241.
125. 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.
126. Burckstummer, T., et al., *An orthogonal proteomic-genomic screen identifies AIM2 as a cytoplasmic DNA sensor for the inflammasome.* Nat Immunol, 2009. **10**(3): p. 266-72.
127. Fernandes-Alnemri, T., et al., *AIM2 activates the inflammasome and cell death in response to cytoplasmic DNA.* Nature, 2009. **458**(7237): p. 509-13.

128. Hornung, V., et al., *AIM2 recognizes cytosolic dsDNA and forms a caspase-1-activating inflammasome with ASC.* Nature, 2009. **458**(7237): p. 514-8.
129. Roberts, T.L., et al., *HIN-200 proteins regulate caspase activation in response to foreign cytoplasmic DNA.* Science, 2009. **323**(5917): p. 1057-60.
130. 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.
131. Burdette, D.L., et al., *STING is a direct innate immune sensor of cyclic di-GMP.* Nature, 2011. **478**(7370): p. 515-8.
132. 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.
133. 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.
134. Zhang, Z., et al., *The helicase DDX41 senses intracellular DNA mediated by the adaptor STING in dendritic cells.* Nat Immunol, 2011. **12**(10): p. 959-65.
135. Unterholzner, L., et al., *IFI16 is an innate immune sensor for intracellular DNA.* Nat Immunol, 2010. **11**(11): p. 997-1004.
136. Schattgen, S.A. and K.A. Fitzgerald, *The PYHIN protein family as mediators of host defenses.* Immunol Rev, 2011. **243**(1): p. 109-18.
137. Kerur, N., et al., *IFI16 acts as a nuclear pathogen sensor to induce the inflammasome in response to Kaposi Sarcoma-associated herpesvirus infection.* Cell Host Microbe, 2011. **9**(5): p. 363-75.
138. 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.
139. Lee, Y.H. and M.R. Stallcup, *Interplay of Fli-I and FLAP1 for regulation of beta-catenin dependent transcription.* Nucleic Acids Res, 2006. **34**(18): p. 5052-9.
140. 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.
141. 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.
142. Banchereau, J. and V. Pascual, *Type I interferon in systemic lupus erythematosus and other autoimmune diseases.* Immunity, 2006. **25**(3): p. 383-92.
143. Baccala, R., et al., *TLR-dependent and TLR-independent pathways of type I interferon induction in systemic autoimmunity.* Nat Med, 2007. **13**(5): p. 543-51.
144. Yasuda, K., et al., *Endosomal translocation of vertebrate DNA activates dendritic cells via TLR9-dependent and -independent pathways.* J Immunol, 2005. **174**(10): p. 6129-36.
145. Barton, G.M., J.C. Kagan, and R. Medzhitov, *Intracellular localization of Toll-like receptor 9 prevents recognition of self DNA but facilitates access to viral DNA.* Nat Immunol, 2006. **7**(1): p. 49-56.
146. Gilliet, M., W. Cao, and Y.J. Liu, *Plasmacytoid dendritic cells: sensing nucleic acids in viral infection and autoimmune diseases.* Nat Rev Immunol, 2008. **8**(8): p. 594-606.
147. Napirei, M., et al., *Features of systemic lupus erythematosus in Dnase1-deficient mice.* Nat Genet, 2000. **25**(2): p. 177-81.
148. Ronnblom, L., M.L. Eloranta, and G.V. Alm, *Role of natural interferon-alpha producing cells (plasmacytoid dendritic cells) in autoimmunity.* Autoimmunity, 2003. **36**(8): p. 463-72.

149. 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.
150. Means, T.K., et al., *Human lupus autoantibody-DNA complexes activate DCs through cooperation of CD32 and TLR9.* J Clin Invest, 2005. **115**(2): p. 407-17.
151. Vollmer, J., et al., *Immune stimulation mediated by autoantigen binding sites within small nuclear RNAs involves Toll-like receptors 7 and 8.* J Exp Med, 2005. **202**(11): p. 1575-85.
152. Christensen, S.R., et al., *Toll-like receptor 7 and TLR9 dictate autoantibody specificity and have opposing inflammatory and regulatory roles in a murine model of lupus.* Immunity, 2006. **25**(3): p. 417-28.
153. Yasutomo, K., et al., *Mutation of DNASE1 in people with systemic lupus erythematosus.* Nat Genet, 2001. **28**(4): p. 313-4.
154. Crow, Y.J., et al., *Mutations in the gene encoding the 3'-5' DNA exonuclease TREX1 cause Aicardi-Goutieres syndrome at the AGS1 locus.* Nat Genet, 2006. **38**(8): p. 917-20.
155. Rice, G., et al., *Clinical and molecular phenotype of Aicardi-Goutieres syndrome.* Am J Hum Genet, 2007. **81**(4): p. 713-25.
156. Crow, Y.J., et al., *Cree encephalitis is allelic with Aicardi-Goutieres syndrome: implications for the pathogenesis of disorders of interferon alpha metabolism.* J Med Genet, 2003. **40**(3): p. 183-7.
157. Sanchis, A., et al., *Genetic syndromes mimic congenital infections.* J Pediatr, 2005. **146**(5): p. 701-5.
158. 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.
159. Stetson, D.B., et al., *Trex1 prevents cell-intrinsic initiation of autoimmunity.* Cell, 2008. **134**(4): p. 587-98.
160. Crow, Y.J., et al., *Mutations in genes encoding ribonuclease H2 subunits cause Aicardi-Goutieres syndrome and mimic congenital viral brain infection.* Nat Genet, 2006. **38**(8): p. 910-6.
161. Rice, G.I., et al., *Mutations involved in Aicardi-Goutieres syndrome implicate SAMHD1 as regulator of the innate immune response.* Nat Genet, 2009. **41**(7): p. 829-32.
162. Goldstone, D.C., et al., *HIV-1 restriction factor SAMHD1 is a deoxynucleoside triphosphate triphosphohydrolase.* Nature, 2011. **480**(7377): p. 379-82.
163. Lee-Kirsch, M.A., *Nucleic acid metabolism and systemic autoimmunity revisited.* Arthritis Rheum, 2010. **62**(5): p. 1208-12.
164. Yoneyama, M., et al., *The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses.* Nat Immunol, 2004. **5**(7): p. 730-7.
165. Saito, T. and M. Gale, Jr., *Differential recognition of double-stranded RNA by RIG-I-like receptors in antiviral immunity.* J Exp Med, 2008. **205**(7): p. 1523-7.
166. Leadbetter, E.A., et al., *Chromatin-IgG complexes activate B cells by dual engagement of IgM and Toll-like receptors.* Nature, 2002. **416**(6881): p. 603-7.
167. Pisitkun, P., et al., *Autoreactive B cell responses to RNA-related antigens due to TLR7 gene duplication.* Science, 2006. **312**(5780): p. 1669-72.
168. Subramanian, S., et al., *A Tlr7 translocation accelerates systemic autoimmunity in murine lupus.* Proc Natl Acad Sci U S A, 2006. **103**(26): p. 9970-5.

169. Lande, R., et al., *Plasmacytoid dendritic cells sense self-DNA coupled with antimicrobial peptide.* Nature, 2007. **449**(7162): p. 564-9.
170. 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.
171. Yoshida, H., et al., *Lethal anemia caused by interferon-beta produced in mouse embryos carrying undigested DNA.* Nat Immunol, 2005. **6**(1): p. 49-56.
172. Lindahl, T., J.A. Gally, and G.M. Edelman, *Properties of deoxyribonuclease 3 from mammalian tissues.* J Biol Chem, 1969. **244**(18): p. 5014-9.
173. Hoss, M., et al., *A human DNA editing enzyme homologous to the Escherichia coli DnaQ/MutD protein.* EMBO J, 1999. **18**(13): p. 3868-75.
174. Mazur, D.J. and F.W. Perrino, *Structure and expression of the TREX1 and TREX2 3' --> 5' exonuclease genes.* J Biol Chem, 2001. **276**(18): p. 14718-27.
175. Lee-Kirsch, M.A., et al., *A mutation in TREX1 that impairs susceptibility to granzyme A-mediated cell death underlies familial chilblain lupus.* J Mol Med (Berl), 2007. **85**(5): p. 531-7.
176. Lee-Kirsch, M.A., et al., *Familial chilblain lupus, a monogenic form of cutaneous lupus erythematosus, maps to chromosome 3p.* Am J Hum Genet, 2006. **79**(4): p. 731-7.
177. Rice, G., et al., *Heterozygous mutations in TREX1 cause familial chilblain lupus and dominant Aicardi-Goutieres syndrome.* Am J Hum Genet, 2007. **80**(4): p. 811-5.
178. Lee-Kirsch, M.A., et al., *Mutations in the gene encoding the 3'-5' DNA exonuclease TREX1 are associated with systemic lupus erythematosus.* Nat Genet, 2007. **39**(9): p. 1065-7.
179. Richards, A., et al., *C-terminal truncations in human 3'-5' DNA exonuclease TREX1 cause autosomal dominant retinal vasculopathy with cerebral leukodystrophy.* Nat Genet, 2007. **39**(9): p. 1068-70.
180. Morita, M., et al., *Gene-targeted mice lacking the Trex1 (DNase III) 3'-->5' DNA exonuclease develop inflammatory myocarditis.* Mol Cell Biol, 2004. **24**(15): p. 6719-27.
181. 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.
182. Yan, N., et al., *The SET complex acts as a barrier to autointegration of HIV-1.* PLoS Pathog, 2009. **5**(3): p. e1000327.
183. Yan, N., et al., *The cytosolic exonuclease TREX1 inhibits the innate immune response to human immunodeficiency virus type 1.* Nat Immunol, 2010. **11**(11): p. 1005-13.
184. Lewin, A.R., et al., *Molecular analysis of a human interferon-inducible gene family.* Eur J Biochem, 1991. **199**(2): p. 417-23.
185. Friedman, R.L., et al., *Transcriptional and posttranscriptional regulation of interferon-induced gene expression in human cells.* Cell, 1984. **38**(3): p. 745-55.
186. Bradbury, L.E., et al., *The CD19/CD21 signal transducing complex of human B lymphocytes includes the target of antiproliferative antibody-1 and Leu-13 molecules.* J Immunol, 1992. **149**(9): p. 2841-50.
187. Smith, R.A., et al., *Expression of the mouse fragilis gene products in immune cells and association with receptor signaling complexes.* Genes Immun, 2006. **7**(2): p. 113-21.
188. Lange, U.C., et al., *The fragilis interferon-inducible gene family of transmembrane proteins is associated with germ cell specification in mice.* BMC Dev Biol, 2003. **3**: p. 1.

189. Lange, U.C., et al., *Normal germ line establishment in mice carrying a deletion of the Ifitm/Fragilis gene family cluster.* Mol Cell Biol, 2008. **28**(15): p. 4688-96.
190. Ropolo, A., et al., *Cloning of IP15, a pancreatitis-induced gene whose expression inhibits cell growth.* Biochem Biophys Res Commun, 2004. **319**(3): p. 1001-9.
191. Evans, S.S., et al., *IFN-alpha induces homotypic adhesion and Leu-13 expression in human B lymphoid cells.* J Immunol, 1993. **150**(3): p. 736-47.
192. Moffatt, P., et al., *Bril: a novel bone-specific modulator of mineralization.* J Bone Miner Res, 2008. **23**(9): p. 1497-508.
193. Brem, R., et al., *Inhibition of proliferation by 1-8U in interferon-alpha-responsive and non-responsive cell lines.* Cell Mol Life Sci, 2003. **60**(6): p. 1235-48.
194. Hoek, M. and B. Stillman, *Chromatin assembly factor 1 is essential and couples chromatin assembly to DNA replication in vivo.* Proc Natl Acad Sci U S A, 2003. **100**(21): p. 12183-8.
195. Nabatiyan, A. and T. Krude, *Silencing of chromatin assembly factor 1 in human cells leads to cell death and loss of chromatin assembly during DNA synthesis.* Mol Cell Biol, 2004. **24**(7): p. 2853-62.
196. Smith, S. and B. Stillman, *Purification and characterization of CAF-I, a human cell factor required for chromatin assembly during DNA replication in vitro.* Cell, 1989. **58**(1): p. 15-25.
197. Smith, S. and B. Stillman, *Stepwise assembly of chromatin during DNA replication in vitro.* EMBO J, 1991. **10**(4): p. 971-80.
198. Mossi, R. and U. Hubscher, *Clamping down on clamps and clamp loaders--the eukaryotic replication factor C.* Eur J Biochem, 1998. **254**(2): p. 209-16.
199. Hay, R.T., *SUMO: a history of modification.* Mol Cell, 2005. **18**(1): p. 1-12.
200. Desterro, J.M., et al., *Identification of the enzyme required for activation of the small ubiquitin-like protein SUMO-1.* J Biol Chem, 1999. **274**(15): p. 10618-24.
201. Kahyo, T., T. Nishida, and H. Yasuda, *Involvement of PIAS1 in the sumoylation of tumor suppressor p53.* Mol Cell, 2001. **8**(3): p. 713-8.
202. Liu, B., et al., *Inhibition of Stat1-mediated gene activation by PIAS1.* Proc Natl Acad Sci U S A, 1998. **95**(18): p. 10626-31.
203. Liu, B., et al., *PIAS1 selectively inhibits interferon-inducible genes and is important in innate immunity.* Nat Immunol, 2004. **5**(9): p. 891-8.
204. Megidish, T., J.H. Xu, and C.W. Xu, *Activation of p53 by protein inhibitor of activated Stat1 (PIAS1).* J Biol Chem, 2002. **277**(10): p. 8255-9.
205. Ungureanu, D., et al., *PIAS proteins promote SUMO-1 conjugation to STAT1.* Blood, 2003. **102**(9): p. 3311-3.
206. Drost, J., et al., *BRD7 is a candidate tumour suppressor gene required for p53 function.* Nat Cell Biol, 2010. **12**(4): p. 380-9.
207. Zhou, J., et al., *BRD7, a novel bromodomain gene, inhibits G1-S progression by transcriptionally regulating some important molecules involved in ras/MEK/ERK and Rb/E2F pathways.* J Cell Physiol, 2004. **200**(1): p. 89-98.
208. Brass, A.L., et al., *The IFITM proteins mediate cellular resistance to influenza A H1N1 virus, West Nile virus, and dengue virus.* Cell, 2009. **139**(7): p. 1243-54.
209. Feeley, E.M., et al., *IFITM3 inhibits influenza A virus infection by preventing cytosolic entry.* PLoS Pathog, 2011. **7**(10): p. e1002337.

210. Weidner, J.M., et al., *Interferon-induced cell membrane proteins, IFITM3 and tetherin, inhibit vesicular stomatitis virus infection via distinct mechanisms.* J Virol, 2010. **84**(24): p. 12646-57.
211. Lu, J., et al., *The IFITM proteins inhibit HIV-1 infection.* J Virol, 2011. **85**(5): p. 2126-37.
212. Papouli, E., et al., *Crosstalk between SUMO and ubiquitin on PCNA is mediated by recruitment of the helicase Srs2p.* Mol Cell, 2005. **19**(1): p. 123-33.
213. Pfander, B., et al., *SUMO-modified PCNA recruits Srs2 to prevent recombination during S phase.* Nature, 2005. **436**(7049): p. 428-33.
214. Shibahara, K. and B. Stillman, *Replication-dependent marking of DNA by PCNA facilitates CAF-1-coupled inheritance of chromatin.* Cell, 1999. **96**(4): p. 575-85.
215. Oku, T., et al., *Functional sites of human PCNA which interact with p21 (Cip1/Waf1), DNA polymerase delta and replication factor C.* Genes Cells, 1998. **3**(6): p. 357-69.
216. Mossi, R., et al., *Replication factor C interacts with the C-terminal side of proliferating cell nuclear antigen.* J Biol Chem, 1997. **272**(3): p. 1769-76.
217. Tadokoro, T. and S. Kanaya, *Ribonuclease H: molecular diversities, substrate binding domains, and catalytic mechanism of the prokaryotic enzymes.* FEBS J, 2009. **276**(6): p. 1482-93.
218. Chon, H., et al., *Contributions of the two accessory subunits, RNASEH2B and RNASEH2C, to the activity and properties of the human RNase H2 complex.* Nucleic Acids Res, 2009. **37**(1): p. 96-110.
219. Kalantry, S., et al., *Gene rearrangements in the molecular pathogenesis of acute promyelocytic leukemia.* J Cell Physiol, 1997. **173**(2): p. 288-96.
220. Dyck, J.A., et al., *A novel macromolecular structure is a target of the promyelocyte-retinoic acid receptor oncoprotein.* Cell, 1994. **76**(2): p. 333-43.
221. Koken, M.H., et al., *The t(15;17) translocation alters a nuclear body in a retinoic acid-reversible fashion.* EMBO J, 1994. **13**(5): p. 1073-83.
222. Weis, K., et al., *Retinoic acid regulates aberrant nuclear localization of PML-RAR alpha in acute promyelocytic leukemia cells.* Cell, 1994. **76**(2): p. 345-56.
223. Hodges, M., et al., *Structure, organization, and dynamics of promyelocytic leukemia protein nuclear bodies.* Am J Hum Genet, 1998. **63**(2): p. 297-304.
224. Zhong, S., et al., *Role of SUMO-1-modified PML in nuclear body formation.* Blood, 2000. **95**(9): p. 2748-52.
225. Muller, S., M.J. Matunis, and A. Dejean, *Conjugation with the ubiquitin-related modifier SUMO-1 regulates the partitioning of PML within the nucleus.* EMBO J, 1998. **17**(1): p. 61-70.
226. Sternsdorf, T., K. Jensen, and H. Will, *Evidence for covalent modification of the nuclear dot-associated proteins PML and Sp100 by PIC1/SUMO-1.* J Cell Biol, 1997. **139**(7): p. 1621-34.
227. Adamson, A.L. and S. Kenney, *Epstein-barr virus immediate-early protein BZLF1 is SUMO-1 modified and disrupts promyelocytic leukemia bodies.* J Virol, 2001. **75**(5): p. 2388-99.
228. Boutell, C., A. Orr, and R.D. Everett, *PML residue lysine 160 is required for the degradation of PML induced by herpes simplex virus type 1 regulatory protein ICP0.* J Virol, 2003. **77**(16): p. 8686-94.

229. Everett, R.D. and J. Murray, *ND10 components relocate to sites associated with herpes simplex virus type 1 nucleoprotein complexes during virus infection.* J Virol, 2005. **79**(8): p. 5078-89.
230. Kang, H., et al., *Inhibition of SUMO-independent PML oligomerization by the human cytomegalovirus IE1 protein.* J Gen Virol, 2006. **87**(Pt 8): p. 2181-90.
231. Kim, Y.E., et al., *Human cytomegalovirus infection causes degradation of Sp100 proteins that suppress viral gene expression.* J Virol, 2011. **85**(22): p. 11928-37.
232. Wang, L., et al., *Disruption of PML nuclear bodies is mediated by ORF61 SUMO-interacting motifs and required for varicella-zoster virus pathogenesis in skin.* PLoS Pathog, 2011. **7**(8): p. e1002157.
233. Kang, D.C., et al., *mda-5: An interferon-inducible putative RNA helicase with double-stranded RNA-dependent ATPase activity and melanoma growth-suppressive properties.* Proc Natl Acad Sci U S A, 2002. **99**(2): p. 637-42.
234. Fitzgerald, K.A., et al., *IKKepsilon and TBK1 are essential components of the IRF3 signaling pathway.* Nat Immunol, 2003. **4**(5): p. 491-6.
235. Sharma, S., et al., *Triggering the interferon antiviral response through an IKK-related pathway.* Science, 2003. **300**(5622): p. 1148-51.
236. Hemmi, H., et al., *The roles of two IkappaB kinase-related kinases in lipopolysaccharide and double stranded RNA signaling and viral infection.* J Exp Med, 2004. **199**(12): p. 1641-50.
237. 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.
238. Perry, A.K., et al., *Differential requirement for TANK-binding kinase-1 in type I interferon responses to toll-like receptor activation and viral infection.* J Exp Med, 2004. **199**(12): p. 1651-8.
239. Hornung, V. and E. Latz, *Intracellular DNA recognition.* Nat Rev Immunol, 2010. **10**(2): p. 123-30.
240. Sauer, J.D., et al., *The N-ethyl-N-nitrosourea-induced Goldenticket mouse mutant reveals an essential function of Sting in the in vivo interferon response to Listeria monocytogenes and cyclic dinucleotides.* Infect Immun, 2011. **79**(2): p. 688-94.
241. 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.
242. Fernandes-Alnemri, T., et al., *The AIM2 inflammasome is critical for innate immunity to Francisella tularensis.* Nat Immunol, 2010. **11**(5): p. 385-93.
243. Albrecht, M., D. Choubey, and T. Lengauer, *The HIN domain of IFI-200 proteins consists of two OB folds.* Biochem Biophys Res Commun, 2005. **327**(3): p. 679-87.
244. Saito, T., et al., *Innate immunity induced by composition-dependent RIG-I recognition of hepatitis C virus RNA.* Nature, 2008. **454**(7203): p. 523-7.
245. Kent, W.J., et al., *The human genome browser at UCSC.* Genome Res, 2002. **12**(6): p. 996-1006.
246. Young, J.M., et al., *Different evolutionary processes shaped the mouse and human olfactory receptor gene families.* Hum Mol Genet, 2002. **11**(5): p. 535-46.
247. Stajich, J.E., et al., *The Bioperl toolkit: Perl modules for the life sciences.* Genome Res, 2002. **12**(10): p. 1611-8.

248. Punta, M., et al., *The Pfam protein families database.* Nucleic Acids Res, 2012. **40**(Database issue): p. D290-301.
249. Huson, D.H., et al., *Dendroscope: An interactive viewer for large phylogenetic trees.* BMC Bioinformatics, 2007. **8**: p. 460.
250. Lieberman, P.M., *Chromatin regulation of virus infection.* Trends Microbiol, 2006. **14**(3): p. 132-40.
251. Benech, P., et al., *Structure of two forms of the interferon-induced (2'-5') oligo A synthetase of human cells based on cDNAs and gene sequences.* EMBO J, 1985. **4**(9): p. 2249-56.
252. Saunders, M.E., et al., *Human 2-5A synthetase: characterization of a novel cDNA and corresponding gene structure.* EMBO J, 1985. **4**(7): p. 1761-8.
253. Wathelet, M., et al., *Molecular cloning, full-length sequence and preliminary characterization of a 56-kDa protein induced by human interferons.* Eur J Biochem, 1986. **155**(1): p. 11-7.
254. Wathelet, M., et al., *Full-length sequence and expression of the 42 kDa 2-5A synthetase induced by human interferon.* FEBS Lett, 1986. **196**(1): p. 113-20.
255. Chebath, J., et al., *Four different forms of interferon-induced 2',5'-oligo(A) synthetase identified by immunoblotting in human cells.* J Biol Chem, 1987. **262**(8): p. 3852-7.
256. Hovanessian, A.G., et al., *Identification of 69-kd and 100-kd forms of 2-5A synthetase in interferon-treated human cells by specific monoclonal antibodies.* EMBO J, 1987. **6**(5): p. 1273-80.
257. Hovanessian, A.G., *Interferon-induced and double-stranded RNA-activated enzymes: a specific protein kinase and 2',5'-oligoadenylate synthetases.* J Interferon Res, 1991. **11**(4): p. 199-205.
258. Zhou, A., B.A. Hassel, and R.H. Silverman, *Expression cloning of 2-5A-dependent RNAase: a uniquely regulated mediator of interferon action.* Cell, 1993. **72**(5): p. 753-65.
259. Williams, B.R., et al., *Natural occurrence of 2-5A in interferon-treated EMC virus-infected L cells.* Nature, 1979. **282**(5739): p. 582-6.
260. Laurence, L., et al., *Comparison of the effects of rabies virus infection and of combined interferon and poly(I).poly(C) treatment on the levels of 2',5'-adenyladenosine oligonucleotides in different organs of mice.* Virology, 1985. **143**(1): p. 290-9.
261. Chebath, J., et al., *Constitutive expression of (2'-5') oligo A synthetase confers resistance to picornavirus infection.* Nature, 1987. **330**(6148): p. 587-8.
262. Coccia, E.M., et al., *A full-length murine 2-5A synthetase cDNA transfected in NIH-3T3 cells impairs EMCV but not VSV replication.* Virology, 1990. **179**(1): p. 228-33.
263. Schroder, H.C., et al., *(2'-5')Oligoadenylate and intracellular immunity against retrovirus infection.* Int J Biochem, 1992. **24**(1): p. 55-63.
264. Schroder, H.C., et al., *Protection of HeLa-T4+ cells against human immunodeficiency virus (HIV) infection after stable transfection with HIV LTR-2',5'-oligoadenylate synthetase hybrid gene.* FASEB J, 1990. **4**(13): p. 3124-30.
265. Chebath, J., et al., *Interferon-induced 56,000 Mr protein and its mRNA in human cells: molecular cloning and partial sequence of the cDNA.* Nucleic Acids Res, 1983. **11**(5): p. 1213-26.
266. Kusari, J. and G.C. Sen, *Transcriptional analyses of interferon-inducible mRNAs.* Mol Cell Biol, 1987. **7**(1): p. 528-31.

267. Levy, D., et al., *Interferon-stimulated transcription: isolation of an inducible gene and identification of its regulatory region.* Proc Natl Acad Sci U S A, 1986. **83**(23): p. 8929-33.
268. Bluyssen, H.A., et al., *Structure, chromosome localization, and regulation of expression of the interferon-regulated mouse Ifi54/Ifi56 gene family.* Genomics, 1994. **24**(1): p. 137-48.
269. Lee, C.G., et al., *Molecular cloning and characterization of a murine LPS-inducible cDNA.* J Immunol, 1994. **152**(12): p. 5758-67.
270. Smith, J.B. and H.R. Herschman, *The glucocorticoid attenuated response genes GARG-16, GARG-39, and GARG-49/IRG2 encode inducible proteins containing multiple tetratricopeptide repeat domains.* Arch Biochem Biophys, 1996. **330**(2): p. 290-300.
271. Niikura, T., R. Hirata, and S.C. Weil, *A novel interferon-inducible gene expressed during myeloid differentiation.* Blood Cells Mol Dis, 1997. **23**(3): p. 337-49.
272. Yu, M., et al., *Cloning of a gene (RIG-G) associated with retinoic acid-induced differentiation of acute promyelocytic leukemia cells and representing a new member of a family of interferon-stimulated genes.* Proc Natl Acad Sci U S A, 1997. **94**(14): p. 7406-11.
273. Zhu, H., J.P. Cong, and T. Shenk, *Use of differential display analysis to assess the effect of human cytomegalovirus infection on the accumulation of cellular RNAs: induction of interferon-responsive RNAs.* Proc Natl Acad Sci U S A, 1997. **94**(25): p. 13985-90.
274. de Veer, M.J., et al., *IFI60/ISG60/IFIT4, a new member of the human IFI54/IFIT2 family of interferon-stimulated genes.* Genomics, 1998. **54**(2): p. 267-77.
275. 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.
276. 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.
277. 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.
278. Terenzi, F., S. Pal, and G.C. Sen, *Induction and mode of action of the viral stress-inducible murine proteins, P56 and P54.* Virology, 2005. **340**(1): p. 116-24.
279. Terenzi, F., P. Saikia, and G.C. Sen, *Interferon-inducible protein, P56, inhibits HPV DNA replication by binding to the viral protein E1.* EMBO J, 2008. **27**(24): p. 3311-21.
280. Wang, C., et al., *Alpha interferon induces distinct translational control programs to suppress hepatitis C virus RNA replication.* J Virol, 2003. **77**(7): p. 3898-912.
281. Li, Y., et al., *ISG56 is a negative-feedback regulator of virus-triggered signaling and cellular antiviral response.* Proc Natl Acad Sci U S A, 2009. **106**(19): p. 7945-50.
282. McWhirter, S.M., et al., *A host type I interferon response is induced by cytosolic sensing of the bacterial second messenger cyclic-di-GMP.* J Exp Med, 2009. **206**(9): p. 1899-911.
283. Roberts, Z.J., et al., *The chemotherapeutic agent DMXAA potently and specifically activates the TBK1-IRF-3 signaling axis.* J Exp Med, 2007. **204**(7): p. 1559-69.
284. Tate, V.E. and L. Philipson, *Parental adenovirus DNA accumulates in nucleosome-like structures in infected cells.* Nucleic Acids Res, 1979. **6**(8): p. 2769-85.
285. Mirza, M.A. and J. Weber, *Structure of adenovirus chromatin.* Biochimica et biophysica acta, 1982. **696**(1): p. 76-86.

286. Kent, J.R., et al., *During lytic infection herpes simplex virus type 1 is associated with histones bearing modifications that correlate with active transcription.* J Virol, 2004. **78**(18): p. 10178-86.
287. Pichlmair, A., et al., *IFIT1 is an antiviral protein that recognizes 5'-triphosphate RNA.* Nat Immunol, 2011. **12**(7): p. 624-30.
288. Rozzo, S.J., et al., *Evidence for an interferon-inducible gene, Ifi202, in the susceptibility to systemic lupus.* Immunity, 2001. **15**(3): p. 435-43.
289. Elinav, E., et al., *Regulation of the antimicrobial response by NLR proteins.* Immunity, 2011. **34**(5): p. 665-79.
290. Harris, R.S. and M.T. Liddament, *Retroviral restriction by APOBEC proteins.* Nat Rev Immunol, 2004. **4**(11): p. 868-77.
291. Kim, B.H., et al., *A family of IFN-gamma-inducible 65-kD GTPases protects against bacterial infection.* Science, 2011. **332**(6030): p. 717-21.
292. Torgerson, D.G., et al., *Meta-analysis of genome-wide association studies of asthma in ethnically diverse North American populations.* Nat Genet, 2011. **43**(9): p. 887-92.
293. Meredith, R.W., et al., *Impacts of the Cretaceous Terrestrial Revolution and KPg extinction on mammal diversification.* Science, 2011. **334**(6055): p. 521-4.
294. Leulier, F. and B. Lemaitre, *Toll-like receptors--taking an evolutionary approach.* Nat Rev Genet, 2008. **9**(3): p. 165-78.
295. Davis-Poynter, N.J. and H.E. Farrell, *Masters of deception: a review of herpesvirus immune evasion strategies.* Immunology and cell biology, 1996. **74**(6): p. 513-22.
296. Schramm, B. and J.K. Locker, *Cytoplasmic organization of POXvirus DNA replication.* Traffic, 2005. **6**(10): p. 839-46.
297. Cohen, S., S. Au, and N. Pante, *How viruses access the nucleus.* Biochimica et biophysica acta, 2011. **1813**(9): p. 1634-45.
298. Khorasanizadeh, S., *The nucleosome: from genomic organization to genomic regulation.* Cell, 2004. **116**(2): p. 259-72.

Curriculum Vitae

REBECCA L. BRUNETTE

EDUCATION

Ph.D., **Immunology**, December 2012
University of Washington, Seattle, WA

Dissertation: Nucleic acid sensing in anti-viral responses and autoimmunity
Advisor: Daniel B. Stetson

Certificate: Molecular Medicine

B.S., Microbiology, June 2006
University of Washington, Seattle, WA

RESEARCH INTERESTS

- Inflammation associated with cancer and autoimmune disease
- Intra- and extracellular ligands that signal innate immune-mediated inflammation

HONORS/ AFFILIATIONS

- American Association of Immunologists, September 2008 - present
- Howard Hughes Medical Institute Molecular Medicine Training Program Scholar, September 2009 – present
- Sandra H. Clark Immunology Travel Education Fund Recipient, April 2011
- Cancer Research Institute Pre-Doctoral Fellowship, September 2008 – August 2010

RESEARCH EXPERIENCE

Ph.D. Research, Department of Immunology, University of Washington,
Seattle, WA, September 2007 - present

- Designed, conducted and analyzed experiments characterizing the role of Trex1 nuclease as a negative regulator of type I IFN responses
- Identified a family of nucleic acid sensors upstream of STING-dependent type I IFN signaling

Pre-Doctoral Emerging Infectious Disease Laboratory Training Fellow, Association of Public Health Laboratories/Centers for Disease Control and Prevention, California Department of Health Services,
Richmond, CA, September 2006 - August 2007

- Worked on donor variation HIV neutralizing antibody project
- Isolation and culture of primary lymphocytes from human blood

- Flow cytometry of surface markers, ELISA of cytokine profiles and HIV neutralizing antibody assays with human lymphocytes
- Culture, stock preparation, titering and plaque assays with HIV

Undergraduate Research Assistant, Department of Pathology, University of Washington, Seattle, WA, May 2004 – August 2006

- Participated in project to examine the role of the caspase family of proteins in P2X7-mediated macrophage fusion in an atherosclerosis model
- Thioglycollate elicitation, harvest and culture of primary macrophages
- Chemotaxis and immunocytochemistry assays to monitor cell fusion
- Western blotting and quantitative real-time PCR to ascertain caspase protein and mRNA levels

TEACHING EXPERIENCE

Teaching Assistant, Department of Immunology, University of Washington, Seattle, WA, September 2008- December 2008

- Taught quiz section part of an undergraduate Immunology lecture course
- Prepared presentations and lectured in weekly quiz sections
- Advised and tutored students during office hours and review sessions
- Wrote, proctored and graded quizzes and exams

Laboratory Teaching Assistant, Department of Microbiology, University of Washington, Seattle, WA, January 2006 – June 2006

- Peer taught laboratory sections for undergraduate introductory and intermediate microbiology courses
- Lectured and demonstrated proper laboratory techniques
- Wrote, proctored and graded quizzes and exams

PUBLICATIONS

Brunette, R.L., J.M. Young, D.G. Whitley, I.E. Brodsky, H.S. Malik, and D.B. Stetson. Extensive evolutionary and functional diversity among mammalian AIM2-like receptors. 2012. J. Exp. Med.; 209:1969-1983.

Rice, G.I., J. Bond, A. Asipu, **R.L. Brunette**, I.W. Manfield, I.M. Carr, J.C. Fuller, R.M. Jackson, T. Lamb, T.A. Briggs, M. Ali, H. Gornall, L.R. Couthard, A. Aeby, S.P. Attard-Montalto, E. Bertini, C. Bodemer, K. Brockmann, L.A. Brueton, P.C. Corry, I. Desguerre, E. Fazzi, A.G. Cazorla, B. Gener, B.C. Hamel, A. Heiberg, M. Hunter, M.S. van der Knaap, R. Kumar, L. Lagae, P.G. Landrieu, C.M. Lourence, D. Marom, M.F. McDermott, W. van der Merwe, S. Orcesi, J.S. Prendiville, M. Rasmussen, S.A. Shalev, D.M. Soler, M. Shinawi, R. Spiegel, T.Y. Tan, A. Vanderver, E.L. Wakeling, E. Wassmer, E. Whittaker, P. Lebon, D.B. Stetson, D.T. Bonthron, and Y.J. Crow. Mutations involved in Aicardi-Goutières syndrome implicate SAMHD1 as regulator of the innate immune response. 2009. Nat. Gen.; 7:829-32.

SELECT ABSTRACTS

Brunette, R.L., J.M. Young, D.G. Whitley, S.M. Malik and D.B. Stetson. Evolutionary and functional diversity among mammalian AIM2-like receptors. 2012. Amgen/University of Washington Joint Research Symposium, Seattle, WA.

Brunette, R.L., D.G. Whitley and D.B. Stetson. 2012. The role of the AIM2-like receptor family in DNA sensing. Midwinter Conference of Immunologists, Asilomar, CA.

Brunette, R.L., D.G. Whitley, J.M. Duggan and D.B. Stetson. 2011. The role for HIN-200 proteins in DNA sensing. Toll 2011: Decoding Innate Immunity, Riva del Garda, Italy.

Brunette, R.L., T.Q. Nhan, W.C. Liles, and S.M. Schwartz. 2006. The Role of Activated Caspases in P2X7 Dependent Macrophage Fusion. South Lake Union Group Research Symposium, Seattle, WA.