DNA Tumor Virus Oncogenes Antagonize the cGAS-STING DNA Sensing Pathway

Laura Lau

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2015

Reading Committee:

Daniel B. Stetson, Chair

Jessica A. Hamerman

Ram Savan

Program Authorized to Offer Degree:

Department of Immunology

University of Washington

Abstract

DNA Tumor Viral Oncogenes Antagonizes the cGAS-STING DNA Sensing Pathway

Laura Lau

Chair of Supervisory Committee:

Associate Professor Daniel B. Stetson

Department of Immunology

A key aspect of antiviral immunity is the induction of type I interferons (IFN) to mediate the effective clearance of a viral infection. Cyclic GMP-AMP synthase (cGAS) detects intracellular DNA and signals through the adapter protein STING to initiate a type I IFN-mediated antiviral response to DNA viruses. These viruses, some of which have evolved with their hosts for millions of years, have likely developed means to prevent activation of the cGAS-STING pathway, but such virus-encoded antagonists remain largely unknown. Here, we identify the viral oncogenes of the DNA tumor viruses, including E7 from human papillomavirus (HPV) and E1A from adenovirus, as potent and specific inhibitors of the cGAS-STING pathway. We show that the LXCXE motif of these oncoproteins, which is essential for blockade of the Retinoblastoma tumor suppressor, is also important for cGAS-STING pathway antagonism. We find that E1A and E7 bind to STING, and that silencing of these oncogenes in human tumor cells restores cGAS-STING pathway signaling. Our

findings reveal a host-virus conflict that may have shaped the evolution of viral oncogenes, with implications for the origins of the DNA viruses that cause cancer in humans.

Table of Contents

Acknowledgments

I would like to express my sincere gratitude to all those who have supported me scientifically as well as personally throughout graduate school. I especially would like to thank my mentor, Daniel B. Stetson for his guidance throughout my project but especially his unwavering support in every aspect of this journey. I would like to thank the lab, past and present, for making this a fun and welcoming environment to work in as well as all the helpful scientific discussions. Thank you to the University of Washington Department of Immunology for making our graduate education a priority but at the same time understanding that there are events in life that occur outside of our control and being so accommodating and supportive during such a difficult time.

I would like to thank my committee for their support and scientific guidance: Michael Gale Jr., Jessica A. Hamerman, Ram Savan, and Jason G. Smith. You've all been a great help throughout my graduate work and have really appreciated the time you've taken to support and foster my scientific growth.

Thank you to Gregory M. Barton for hiring a fresh faced new college graduate with limited lab experience and giving me the opportunity to immerse myself in every aspect of the lab. This was an invaluable opportunity toward fostering my scientific interests. Lastly, I would like to thank my classmates for being my Seattle family. You have all made this a fun and memorable ride so thank you for all the laughs, your support, and the memories. I have to especially thank Nikita S. Kolhatkar for always being such a great friend from recruitment to graduation. We made it!

Dedication

Mama & Baba

Thank you for your trust, belief, and love.

Grandma (Ngin Ngin)

Thank you for making me who I am today. Your patience, generosity to others, and love have always guided me in life.

Dennis Otoo-Kofi Djan

Gone too soon but how you embraced, lived, and loved life is a lesson I hold dear.

Chapter 1: Introduction

1.1 Innate Immune Sensing of Nucleic Acids

The innate immune system is the first line of defense against microbial infection, acting to limit the microbial replication and spread before prompting the adaptive immune response. The innate immune system utilizes a number of germline-encoded receptors called pattern-precognition receptor (PRRs) with fixed specificity to recognize conserved features of microbes[1,2]. These receptors primarily recognize a diverse repertoire of microbial features that are foreign to the host; however one exception is the receptors that detect viruses. Innate immune detection of these pathogens typically rely on the recognition of nucleic acids, as viruses possess few unique features given they utilize the cell host machinery to replicate. Because nucleic acids are common to both pathogen and host, it opens the risk of self nucleic acid recognition by the innate immune sensors. This self recognition has been linked to several autoimmune and autoinflammatory diseases due to these sensors imperfect ability to discriminate viral from host nucleic acid[3,4].

Innate immune sensors involved in nucleic acid recognition can broadly be divided into two groups based on their cellular localization. The first group includes members of the Toll-like receptor (TLR) family. TLR3, TLR7, TLR8, TLR9 and TLR13 are localized to endosomes and lysosomes where they monitor the lumen of these compartments for nucleic acids from viruses and bacteria. The second group includes the cytosolic sensor RIG-I-like receptors (RLRs) and cyclic GMP and AMP synthase (cGAS) which recognize cytoplasmic pathogen-derived RNA and DNA in the cytosol respectively[5-7]

TLRs and the Detection of Exogenous Nucleic Acids

The TLR family members are transmembrane receptors with multiple leucine-rich repeats (LRRs) in the extracellular domain, a transmembrane domain, and a conserved cytosolic signaling domain called the Toll/IL-1 receptor homology (TIR) domain. The nucleic acid sensing TLRs are unique in their localization compared to their microbial sensing TLR family members that are all expressed on the surface of cells. The first TLR to be implicated in nucleic acid sensing was TLR3, which binds double stranded RNA (dsRNA) with low sequence specificity, and is largely expressed on macrophages and conventional dendritic cells (cDCs). In contrast, TLR7/8 recognizes single stranded RNA (ssRNA) that is guanosine/uridine rich and imidazoquinolines, including R848. Despite being closely related, TLR7 and TLR8 are expressed in different cell types. In humans, TLR8 is largely expressed on macrophages and cDCs, whereas TLR7 is expressed on plasmacytoid dendritic cells (pDCs). TLR9 recognizes nonmethylated cytosine-guanosine (CpG) motifs in DNA and is expressed on pDCs[8,9]. Lastly, TLR13 recognizes 23S ribosomal RNA and is only expressed in mice. Notably, these receptors are only expressed on specialized cells, and only recognize exogenous ligands that are internalized into endosomal compartments.

Cytosolic Sensors and the Detection of Endogenous Nucleic Acids

Viruses that enter the cytosol are detected by a different set of sensors to initiate a type I interferon (IFN) immune response. The two families of cytosolic sensors are divided by their mode of ligand recognition: one RNA, the other DNA. For the cytosolic RNA sensing pathway, there are the RIG-I-like Receptors (RLRs) composed of three members: RIG-I,

melanoma differentiation associated gene 5 (MDA5), and laboratory of genetic and physiology 2 (LGP2). These RLRs all have a DExD/H-box RNA helicase and a regulatory C-terminal domain (CTD), whereas only RIG-I and MDA5 have an N-terminal caspase recruitment domains (CARD)[1,10]. The helicase domain interacts with dsRNA, while the CARD domains mediate downstream signals. It was initially thought that because LGP2 lacks a CARD domain and cannot induce a downstream signal that it may be a negative regulator of RIG-I and MDA5[3,11]. However, recent findings have shown that it may assist in RIG-I/MDA5 signaling[5,7,12].

Despite both being able to respond to dsRNA poly(I:C), RIG-I and MDA5 have distinct length preferences for their RNA ligands. In the context of synthetic ligands, RIG-I prefers shorter fragments (~300bp) whereas MDA5 recognizes longer fragments (>4 kb). When looking at more physiologically relevant ligands such as RNA derived from viruses, there are ligands only recognized by RIG-I or MDA, or ligands that are recognized by both receptors. RIG-I can recognize viruses from many viral families such as the Paramyxoviridae family (Sendai virus and Newcastle disease virus), Flaviviridae family (Japanese encephalitis virus and HCV), and Rhabdoviridae family (vesicular stomatitis virus). MDA5 recognizes viruses from the Picornaviridae family (Polio and EMCV). Viruses detected by both sensors include Dengue virus and West Nile virus. The most important RNA feature for RIG-I recognition is a triphosphate group at the 5' end, a feature typically unique to viral RNAs and not present in host mRNAs[8,13]. In contrast, the ligands for MDA5 have not been as well defined. Because MDA5 recognizes long fragments of RNA such as long poly(I:C), it was thought that it could be activated by long dsRNA, particularly from viruses in the Piconaviridae family. However, there is still much debate in

the field as recent studies have shown vaccinia virus can produce high molecular weight RNA that consist of both dsRNA and ssRNA forms thought to activate MDA5[10,14].

Once in contact with their ligand, RIG-I and MDA5 signal downstream to the mitochondrial antiviral signaling (MAVS) adaptor protein, which is essential for mediating the signal from RLRs to IRF3 for the induction of type I IFNs. MAVS contains an N-terminal CARD domain that interacts with RLRs, and a C-terminal transmembrane domain that traffics MAVS into the outer mitochondrial membrane. This localization is a prerequisite for MAVS interaction with Interferon Regulatory Factor 3 (IRF3) and further downstream activation of the type I IFN response.

While the primary receptors and signaling components of the cytosolic RNA sensing pathways are well characterized, the receptors and signaling proteins involved in the cytosolic DNA sensing pathway are still being elucidated. One of the first studies suggesting cytosolic DNA sensing exists used a specific DNA ligand to explore TLR9-independent responses. This ligand, a 45mer dsDNA oligonucleotide lacking CpG motifs, could activate IRF3 to initiate the induction of type I IFN independent of TLR signaling. This became known as the Interferon Stimulatory DNA (ISD) pathway[11,15]. The identity of the DNA sensor in this pathway remained unknown for some time. However, cyclic GMP-AMP synthase (cGAS) was recently shown to be the sensor for cytosolic DNA using biochemical fractionation and quantitative mass spectrometry. cGAS belongs to a family of nucleotidyl transferases (NTases). It possesses a structurally conserved NTase core domain, an N-terminal helical extension, and a carboxyl-terminal domain (C-domain). Upon direct DNA binding, cGAS catalyzes the synthesis of cyclic GMP-AMP (cGAMP) from ATP and GTP[7,12]. The phosphodiester linkages between GMP and AMP were not initially

determined, but through subsequent studies it was revealed that cGAS produces a mixed phosphodiester product that has different phosphodiester linkages. One product is 2'-OH of GMP and 5'-phosphate of AMP, and the other a 3'-OH AMP and 5'-phosphate of GMP. The former molecule, known as 2'3'-cGAMP[7,12,16], is the endogenous form produced in mammalian cells, whereas the latter is produced by bacteria.

This secondary messenger, cGAMP, then binds the signaling adapter stimulator of IFN genes (STING), which is an important signaling component of the DNA sensing pathway for the induction of type I IFNs. The structure of STING can be described in three parts. The N-terminal domain (NTD) is comprised of four transmembrane domains that anchor it to the endoplasmic reticulum. The carboxyl-terminal domain (CTD) has been reported to extend out into the cytosol, as the crystal structure is known. At the end of the CTD is the carboxyl-terminal tail (CTT) which was not included in the crystal structure studies[13,17]. In the absence of ligand, STING exists as a symmetrical dimer where two monomers of the CTD form a cleft. Binding studies have been shown that each STING dimer can bind one molecule of c-di-GMP in this cleft. When bound to cGAMP, STING relocalizes to discrete foci in the cytoplasm where it recruits TANK binding kinase 1 (TBK1) to phosphorylate IRF3 for the induction of type IFNs. The important thing to note about the cytosolic sensors involved in the DNA and RNA sensing pathway is they are expressed on all cell types (not just cells of the immune system) and recognize endogenous ligands.

1.2 Type I Interferon (IFN) signaling and production

Type I IFNs are polypeptides that are secreted by infected cells during a viral infection and are an important antiviral mechanism to limit viral growth and spread. Type I IFNs proteins are encoded by multiple IFN-α genes and a single IFN-β gene. The type I IFN response can be divided into two parts. One is a primary response where most host cells have the ability to active IFN-β gene expression. The secondary response is interferon produced by cells in the primary response bind to the type I IFN receptor (IFNAR), which can activate the expression of interferon-stimulated genes (ISGs) and promote further secretion of IFN-β and initiate the production of IFN-α. This feedback loop is responsible for a highly potent antiviral response.

The initiation of the primary type I IFN response occurs in cells expressing either nucleic acid-sensing TLRs, or intracellular nucleic acid sensors. TLR signaling is initiated by ligand binding that facilitates the dimerization of receptors. The TIR domain then can engage downstream adapters. TLR3 engages with TIR domain-containing adapter protein inducing IFN-β (TRIF) and TRIF-related adapter molecule (TRAM) whereas TLR7, 8, and 9 engage myeloid differentiation primary-response protein 88 (MyD88). Despite knowing 23S rRNA can bind TLR13, the signaling of the receptor still remains poorly understood. TLR3 signaling continues with TNF receptor-associated factors 3 (TRAF3) association, which then recruits the kinases, TBK1 and IKKε to activate IRF3 for the transcription of *ifnα* or *β* genes. For TLR7, 8, and 9, association with adapter molecules initiates the recruitment of IL-1R-associated kinases (IRAKs) and TRAFs, particularly IRAK1, 2, 4 and TRAF6, to form a signaling complex that interacts with IKKα. This then activates IRF7, which translocates to the nucleus for the transcription of type I IFNs.

Intracellular nucleic acid sensors can also initiate the primary type I IFN response. RIG-I/MDA5 and cGAS have distinct adapters of the RNA and DNA sensing pathway but converge downstream of their respective adapter on MAVS and STING. In order to continue the signaling activation of the type I IFN response, both MAVS and STING have a C-terminal consensus motif, *p*L*x*IS (*p*, hydrophilic residue, x, any residue, S, phosphorylation site) where phosphorylation occurs when they come in contact with their ligand. The phosphorylation of MAVS is facilitated by both TBK1 and I kappa B Kinase (IKK). In contrast, phosphorylation of STING only requires TBK1[14,15]. Phosphorylation of both MAVS and STING is required to recruit IRF3, as mutations that impair the phosphorylation of either adapter at their consensus *p*L*x*IS motif impair IRF3 recruitment. Phosphorylated MAVS and STING bind to conserved positively charged patches on one IRF3 molecule. In addition to this binding, a second molecule of IRF3 is recruited and phosphorylated. An IRF3 monomer then disassociates from MAVS or STING to bind with the recruited IRF3 to form a dimer. This IRF3 dimer then translocates to the nucleus to initiate the transcription of *ifn* genes such as IFN-β[14].

Following its primary production, IFN-β binds the heterodimeric transmembrane receptor IFNAR, comprised of IFNAR1 and IFNAR2 subunits. This engagement activates Janus kinase (JAK1) and tyrosine kinase (TYK2). These kinases then recruit and phosphorylate the latent cytoplasmic transcription factors, signal transducer and activator of transcription (STAT) proteins. Phosphorylated STAT1 and STAT2 then recruit IRF9, forming a complex known as interferon-stimulated gene factor 3 (ISGF3). This complex translocates to the nucleus to bind IFN-stimulated response elements (ISRE) in promoters of antiviral genes including IRF7[15,18]. Following its production, IRF7 translocates to the

nucleus to drive production of several IFN-α subtypes and a secondary IFN-β response. This IFN-α/β can then engage with IFNAR in an autocrine manner to reinforce the IFN response.

1.3 Evasion Strategies of the Type I IFN Response by Viruses

The Type I IFN response is the most potent host anti-viral response and the earliest defense against viruses. As a result, viruses that have co-evolved with their host have developed measures to evade detection and prevent activation of the type I IFN response. The engagement of ligand with their sensor is one of the first steps in triggering this response. Thus one evasion strategy viruses use is to sequester ligands by encoding proteins that bind nucleic acids to prevent binding to the cytosolic sensor. For example, influenza encodes a dsRNA binding protein called NS1. Deletion of NS1 results in attenuation of influenza as the ability of the virus to inhibit type I IFN production is compromised[16,19].

Additionally, there are viral antagonists that directly antagonize essential components of the type I IFN signaling pathway. The E6 protein of human papilloma virus 16 (HPV16) can directly bind to IRF3 and prevent its primary activation[17,20]. An example of an antagonist that blocks the secondary IFN response is Large T antigen of murine polyoma virus (MPyV), which binds to and inhibits JAK1 preventing engagement with Tyk2[15,21]. Other mechanisms that viruses employ will be highlighted in more detail in subsequent chapters. Taken together, viruses have employed multiple strategies to evade the host type I IFN response,

This highlights the dynamic host-pathogen relationship where both are co-evolving to either retain detection and mount an appropriate response, or evade detection, respectively.

Chapter 2: Viral oncogenes antagonize the cGAS-DNA sensing pathway

2.1 Introduction

Interferons have been well characterized as an important antiviral defense mechanism. The initial discovery of interferons in 1957 by Isaacs and Lindenmann illustrated its importance in antiviral immunity as their study showed interferons were a critical viral restriction factor during influenza infection.[18,22]. Since this initial finding, much work has been done to figure out the signaling components that initiate this antiviral response. Moreover, due to the potent nature of interferons and their effectiveness in controlling and clearing viruses, viruses have evolved mechanisms to evade this response. For example, many viruses encode antagonists that deliberately target the sensors and signaling proteins that detect viral RNA and DNA in infected cells to evade immune detection[19,23]. These antagonists can be broadly grouped into two classes. The first there are those that target common signaling components of the RNA and DNA sensing pathway, such as TBK1 and IRF3. Examples of this include several herpesvirus (HSV) and RSV where both encode proteins to block the activity of STAT to interfere with type I IFN signaling[20 21]. The second class of virus-encoded antagonists are those that target the upstream components of either the RNA or DNA sensing pathways, which would mediated the inhibition of one while leaving the other intact. The best examples of RNA virus-based antagonists are NS3/4A from hepatitis C virus (HCV), which inhibits MAVS by cleaving it from the mitochondrial membrane[22] and NS1 from influenza that sequesters RNA ligands from RIG-I.

We had reasoned that analogous to the proximal inhibitors of the RLR pathway uniquely found in RNA viruses, there must be cGAS-STING-specific antagonists present in DNA viruses. We began our search for cGAS-STING DNA sensing pathway antagonists by looking at a previous observation that most immortalized and tumor cell lines had failed to respond to cGAS-STING transfection, whereas primary cells mounted a vigorous DNA-activated response[23]. This suggested there was a factor present in immortalized cell lines that inhibited the DNA sensing pathway. The manner in which these cell lines are immortalized are predominantly with a virus that encodes a transformative oncogene. These dominant oncogenes can bind to and inactivate the two principal tumor suppressors in cells: p53 and Retinoblastoma (Rb)[24,25]. This process not only disrupts normal cell cycle, but also is largely responsible for oncogenesis during a persistent viral infection. Two of the most widely used immortalized human cell lines are HEK 293 and HeLa cells. HEK 293 cells were transformed by introduction of human adenovirus 5 (hAd5) DNA into fetal kidney cells[26]. HeLa cells are the oldest human cell line in existence, derived in 1951 from a cervical carcinoma that was later shown to be caused by infection with human papillomavirus 18 (HPV18)[27]. Immortalized mouse fibroblasts are generated by introduction of the large T antigen of simian virus 40 (SV40). hAd5, HPV18, and SV40 are known as "DNA tumor viruses"[24], owing to their ability to cause cancer in experimental animals and (in the case of HPV18) in humans[28].

As mentioned, numerous studies have demonstrated that some viral oncogenes interfere with IRF3- or IFN-dependent transcription downstream of both the cGAS-STING and RIG-I-like receptor pathways, mediating generic inhibition of the antiviral response[29-32]. Based on our initial observations in immortalized and tumor cells, we wondered whether

these DNA virus-encoded oncogenes might also specifically disrupt the unique proximal signaling components of the cGAS-STING DNA sensing pathway.

2.2 Results

Transformed cells lack the cGAS-STING pathway

To further explore the phenomenon that immortalized cells lack a type I IFN response to DNA ligand, we wanted to compare the response between primary and immortalize cells line from human and murine when stimulated with DNA or RNA ligands to further interrogate if this lack of type I IFN response was unique to the cGAS-STING pathway. The primary human cell line we use was human foreskin fibroblasts (HFF) while the primary murine cell lines we used were mouse embryonic fibroblasts (MEFs) and bone marrow-derived macrophages (BMDM). For immortalized human cell lines, we used HEK 293 and HeLa cells while using the immortalized murine cell line, Jax MEFs. We confirmed the expression of STING in each of these primary and immortalized cell lines by western blot (data not shown), because STING is known to be absent in certain immortalized cell lines[33,34]. We transfected these cells with calf thymus (CT) DNA, a specific activator of the cGAS-STING pathway[11], or with a triphosphate RNA ligand that activates the RNA sensor RIG-I[35]. Whereas primary HFF robust type I IFN responses to both DNA and RIG-I ligand (Fig. 1A), HEK 293 cells and HeLa cells responded only to RIG-I ligand and not to DNA (Fig. 1A). We also observed the same findings in SiHas that are immortalize with HPV16 E7 (data not shown). Similarly, MEFs and BMDMs responded to both DNA and RNA ligands, but immortalized mouse fibroblasts responded only to RNA (Fig. 1B). These data demonstrate a specific loss of the intracellular DNA sensing pathway in multiple immortalized cell lines and supports our hypothesis that there is a potential antagonist present in immortalized cell lines that specifically targets the cGAS-STING DNA sensing pathway. To further investigate the antagonist(s) of the cGAS-

STING DNA pathway, we took a closer look into the viral oncogenes that facilitated the immortalization of these cell lines.

Viral oncogenes specifically antagonize the cGAS-STING DNA sensing pathway

In order to address the role of viral oncogenes on the cGAS-STING DNA sensing pathway, we generated retroviral expression vectors containing the oncogenes derived from these viruses: HPV18 E6, HPV18 E7, hAd5 E1A, and SV40 Large T antigen. We chose these specific viral oncogenes to test as they are expressed in the immortalized cell lines we tested. As a control, we included a constitutively activated form of the cellular proto-oncogene H-Ras (G12V)[36]. We transduced primary MEFs with each of these retroviruses, confirmed expression of each oncoprotein by western blot (Fig. 1C), and monitored their responses to transfection with DNA or RIG-I ligand by measuring *Ifnb* mRNA transcription, comparing each oncogene-transduced line to MEFs transduced with an empty control retrovirus. We found that transduction of MEFs with H-Ras G12V mildly affected their response to both RNA and DNA ligands (Fig. 1D). Similarly, HPV18 E6 transduction resulted in a mild impairment of both pathways (Fig. 1D). Remarkably, transduction with HPV18 E7 or hAd5 E1A resulted in a profound blockade of DNA-activated signaling, whereas the RIG-I pathway remained fully intact (Fig. 1D). Interestingly, transduction with SV40 Large T antigen led to a reduction in both the cGAS-STING and the RIG-I responses (Fig. 1D), suggesting that Large T antigen may block a common component of the inducible antiviral response. These data reveal that E1A and E7 specifically block the cGAS-STING pathway when introduced into primary cells.

The role of the LXCXE motif (in antagonizing the cGAS-STING DNA sensing pathway)

The common targets for viral oncogenes are tumor suppressors that regulate the cell cycle. In the case of both hAd5 E1A and HPV18 E7, the common target for both is retinoblastoma (Rb). Rb, p107, and p130 belong to a pocket protein family that binds to proteins containing a Leu-X-Cys-X-Glu (LXCXE) motif. In normal Rb cell cycle regulation, Rb is normally bound to E2F and its associated protein, DP to inhibit the transcriptional activity of E2F by recruiting chromatin remodeling complex to inhibit G1 to S phase cell cycle progression[37]. This repression is regulated by cyclin dependent kinases (CDKs). CDKs phosphorylate Rb, releasing its association with E2F and allowing transcription and transition from G1 to S phase. However, during a viral infection, viral oncogenes hijack this process by mimicking host partners of Rb by interacting through the LXCXE motif and often bind with higher affinity. This results in the displacement of Rb from E2F, thus driving cells into unscheduled cell cycle progression.

We wanted to explore the mechanism in which hAd5 E1A and HPV18 E7 antagonize the cGAS-STING DNA sensing pathway. Considering both of these oncoproteins bind and inhibit the function of Rb using the LXCXE motif, we wanted to test the role of the LXCXE motif in this inhibition. To test this, we introduced conservative mutations into this motif (VXSXD) that are known to disrupt Rb binding (Fig. 2A) [38 39]. We then transduced primary MEFs with retroviruses encoding the VXSXD mutants of E1A or E7 and measured the responses of the transduced cells to DNA and RNA ligands. Whereas the wild-type E1A and E7 potently blocked the cGAS-STING response, the VXSXD mutants of these oncoproteins failed to antagonize (Fig. 2B, C). Importantly, neither the wild-type nor the mutant oncoproteins had an inhibitory effect on RIG-I-mediated signaling (Fig. 2B,

C). These data demonstrate that the mechanism of cGAS-STING pathway antagonism by E1A and E7 is similar to the mechanism by which they disrupt the Rb pathway.

2.3 Discussion

The discovery of type I IFN opened immunology to a new field of research that was focused on discovering the signaling components involved in initiating a response to viral nucleic acids. The RNA sensing pathway has been well characterized as the viral ligands, the viral RNA sensors, and signaling adapters have been identified. However, despite the discovery of the adapter STING, the nature of the DNA sensor and its ligands remained largely unknown[40]. Several candidate sensors have been described. The first was DNA-dependent activator of IFN-regulatory factors (DAI), where *in vitro* findings showed that DAI could induce type I IFN upon overexpression, and a reduced response upon siRNA knockdown. In addition, it was shown to interact with TBK1 and IRF3. However, *in vivo* findings with DAI-deficient mice showed normal type I IFN response with DNA, suggesting it was not the DNA sensor. In addition to DAI, human IFI16 and the mouse IFI204 have been implicated in a STING-dependent IFN response, and were attractive candidates as they belong to the AIM2-like receptor family (ALRs). This receptor family contains a pyrin domain that mediates protein-protein interactions, and a HIN domain that can directly bind to DNA[41,42]. However, findings from our lab suggested otherwise, as IFI204 knockdown did not impair the DNA sensing pathway in BMDMs or MEFs. In addition, we made a complete knock out mouse of all 13 murine ALRs and tested the response to DNA *in vitro*, as well as infections with DNA viruses *in vivo,* and found no difference in response compared to WT. Moreover, human IFI16 did not strongly activate STING in contrast to previous findings[43 44]. It wasn't until the recent discovery of cGAS as the cytosolic DNA sensor that would put to rest the sensor identity debate as the initial *in vitro* cGAS findings were validated *in vivo* in the cGAS knockout mouse[45].

The novel finding of hAd5 E1A and HPV18 as specific antagonists of DNA sensing can be used as a tool to further dissect the mechanisms that regulate the cGAS-STING DNA sensing pathway, such as the potential existence of accessory proteins. We initially thought candidates for accessory proteins were the ALRs because several of these receptors contain LXCXE motifs. We hypothesized that these LXCXE containing ALRs served as an accessory protein to STING as previous studies have shown interaction between the two[43]. This interaction with the then unknown DNA sensor formed a signaling complex that was necessary to initiate a type I IFN response. We postulated the viral oncogenes interrupted this interaction between ALRs and STING and this was the mechanism to inhibiting the immune response. Unfortunately, this is not the case as the ALR knockout mice show no affect on the type I IFN response to DNA.

In addition to revealing viral oncogenes as potential a tool to evaluate the innate DNA sensing pathway, we've also revealed a novel function of DNA tumor viral oncogenes. The oncogenes hAd5 E1A and HPV18 E7 were only known in the context of cancer as they targeted Rb through the LXCXE motif to disrupt normal cell cycle. The perceived benefit for viral expression of these oncogenes was their ability to drive quiescent cells into a proliferative state, forming a new viral niche while providing cellular resources for replication. With our findings illustrating viral oncogenes can inhibit cGAS-STING DNA signaling through the LXCXE motif, we revealed a secondary function of these viral oncogenes in the context of inhibiting the innate immune response.

2.4 Figures

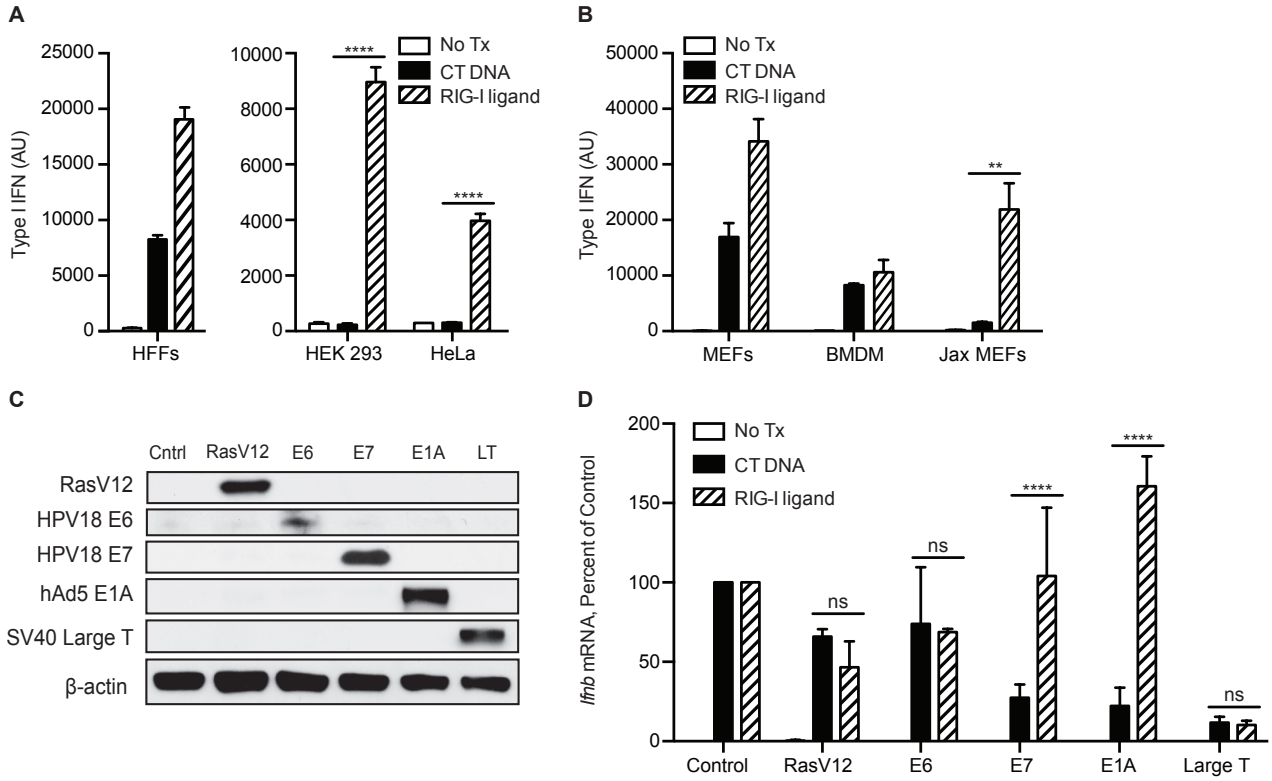

Figure 1: Viral oncogenes antagonize the cGAS-STING DNA sensing pathway.

(A) Primary human foreskin fibroblasts (HFFs), HEK 293 cells, and HeLa cells were treated with the indicated ligands for 8 hours, followed by measurement of type I IFN activity in culture supernatants. **(B)** Primary murine embryonic fibroblasts (MEFs), bone marrow-derived macrophages (BMDM), or Jackson MEFs (Jax MEFs) were treated and measured as in (A). **(C)** Primary MEFs were transduced with retroviruses encoding the indicated oncogenes, and the expression of each oncoprotein was confirmed by western blot of cell extracts. **(D)** Cells transduced with each oncogene were treated with DNA or RNA, and *Ifnb* mRNA responses to were measured by quantitative RT-PCR and compared to cells transduced with control retrovirus. Data are the combined measurements of three independent sets of transductions, with three independent measurements at consecutive cell passages per transduction.

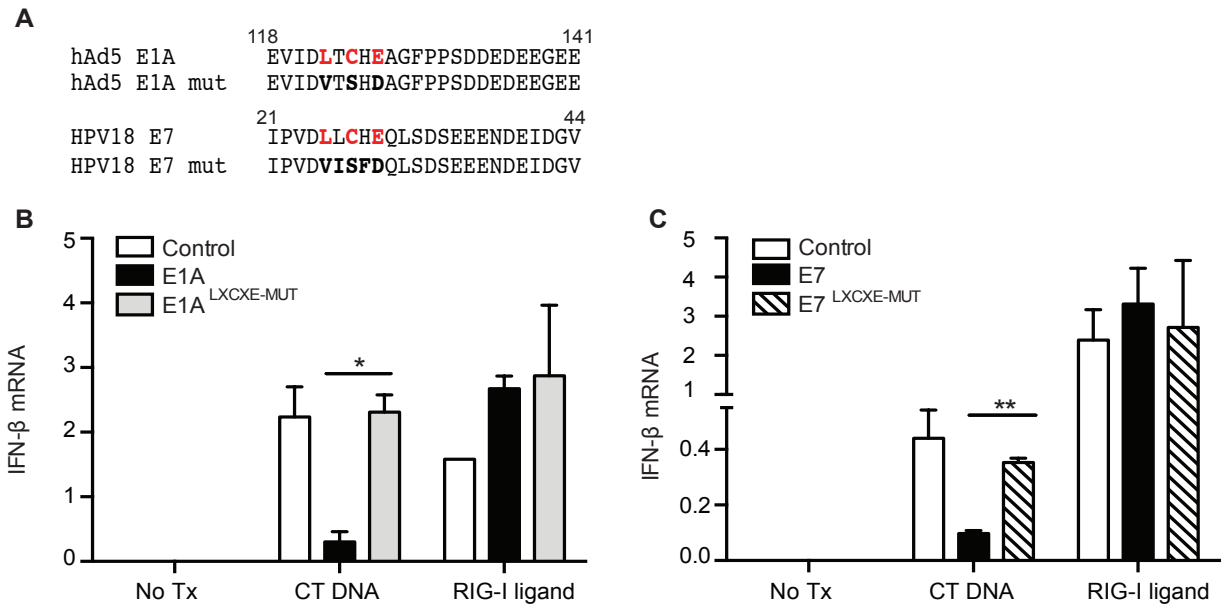

Figure 2: The LXCXE motif of viral oncoproteins is important for cGAS-STING DNA pathway blockade.

(A) Amino acid sequences of the regions of E1A and E7 containing the LXCXE motifs, with the locations of the conservative mutations introduced to disrupt the LXCXE motifs. **(B to C)** Primary MEFs were transduced with retroviruses encoding the wild-type or LXCXE-mutant oncogenes, and *Ifnb* mRNA induction in response to the indicated ligands was measured by quantitative RT-PCR and normalized to *Hprt* mRNA expression within each sample. Results are representative of three independent transductions, with three independent measurements per transduction.

2.5 Materials and Methods

Cell lines, plasmids, and tissue culture

BMDMs and early passage MEFs were generated according to standard techniques. Primary HFFs were generously provided by A.P. Geballe (Fred Hutchinson Cancer Research Center, Seattle, WA). HEK 293 cells were generously provided by S.L. Nishimura (University of California, San Francisco). Jackson MEFs were generously provided by R. Medzhitov (Yale University). HeLa cells were purchased from American Type Culture Collection (ATCC). All cell lines except BMDMs were cultured in DMEM supplemented with 10% fetal calf serum (vol/vol; Gemini Bio-Products), L-glutamine, penicillin/streptomycin, sodium pyruvate, and HEPES (all from Life Technologies). BMDMs were cultured in RPMI media with the above supplements and L929 cell supernatant at 10% (vol/vol).

Primary MEFs were immortalized with retroviral vectors encoding RasV12, SV40 large T antigen cDNA, huAd5 E1A, HPV18 E6, and HPV18 E7 (all plasmids purchased from Addgene). The VISFD mutants of E1A and E7 were created by PCR mutagenesis. Retrovirus was made by transfecting $2.5x10^6$ HEK 293T cells with 10mg retroviral vector and 10mg pCL-Eco using Fugene 6 (Promega) at a ratio of 3ml Fugene per mg of plasmid. 48 hours post transduction, supernatants containing retrovirus were collected, passed through a 0.22mm filter (Millipore), and diluted 1:1 with DMEM before transduction of primary MEFs. Each cell line was monitored and passaged at confluency every 2-3 days. Interferon responses to DNA and RNA were monitored starting nine days post transduction, for at least three consecutive passages per transduction.

Primary BMDMs were transduced with pLKO-puro lentivirus encoding shRNAs as described (*31*), selected for three days in puromycin, and distributed into 12-well tissue culture plates for stimulations.

Cell treatments and analysis

ISD oligonucleotides[23] were annealed and RIG-I ligand was transcribed *in vitro* with T7 Megashortscript kit (Ambion) as previously described[35]. Transduced MEFs, tested at each split starting 9 days post transduction, were plated at 0.15 x 10^6 cells per well in 12 well tissue culture plates for transfection. For transfections, 5mg calf thymus genomic DNA (Sigma-Aldrich), 5mg ISD, and 1mg RIG-I ligand were complexed with Lipofectamine 2000 (Invitrogen) at a ratio of 1 ml lipid per 1mg nucleic acid. DMXAA (Sigma-Aldrich) was added directly to the culture media at 30mg ml^{-1}.

For quantitative RT-PCR analysis of IFN-β mRNA, cells were harvested into RNA-Stat-60 isolation reagent (Tel-Test, Inc.). RNA was treated with DNase (Ambion) and primed with oligo (dT), and reverse-transcribed with Superscript III (Life Technologies). cDNA was used for PCR with EVA Green reagents (Bio-Rad Laboratories) on a Bio-Rad CFX96 Real-Time System. The abundance of IFN-b mRNA was normalized to that of HPRT mRNA in mouse cells and to GAPDH mRNA in human cells, and results were compared with those of untreated cells. PCR primer sequences were previously described[46]. For type I IFN bioassays of human cells, supernatants from treated cells were used to stimulate a HeLa cell line stably expressing an ISRE-luciferase reporter. HeLa-ISRE cells were plated one day prior to treatment at 5×10^4 cells per well in 96-well plates. The following day, growth media was removed from the HeLa-ISRE cells and 75 ml of

supernatant from treated human cells was added and cells incubated for 6 h. Cells were lysed and luciferase activity measured using a Luciferase Assay System (Promega) and Centro LB 960 Microplate Luminometer (Berthold Technologies). For type I IFN bioassays of mouse cells, supernatants from treated cells were used to stimulate a L929 cell line stably expressing an ISRE-luciferase reporter, processed and described as above.

Chapter 3: Characterizing the interaction between viral oncogenes and host

3.1 Introduction

The cellular target of DNA tumor viral oncogenes hAd5 E1A and HPV18 E7 is retinoblastoma (Rb). The Rb gene family includes three members: Rb (p105), p107, and p130. Rb and Rb-related proteins contains two domains: A-box and B-box, which are connected by a short spacer region. This region is the least conserved region of Rb as amino acids in the spacer region varies among the three Rb proteins. They are called "pocket proteins" as each contains a conserved binding pocket region through which they bind host factors where the association is mediated through an LXCXE motif. The LXCXE binding pocket region is located in a groove of the B-box. The interface between the A-box and B- box is an integral part of the pocket structure, as interactions between the side chains of the domains forms a hydrophobic core that extends into the B-box. This suggests the A-box region, while it does not contain the binding groove, may stabilize the structure of the B-box region. In addition to the conserved amino acid residues in the B-box LXCXE binding site, there are also conserved residues in the A-B interface and hydrophobic cores of each domain, all of which are part of the Rb pocket. This pocket groove is a point of contact for most cellular and viral Rb-binding proteins. The crystal structure of the Rb pocket domain was solved bound to a peptide derived from HPV16 E7 protein[47]. This interaction between the LXCXE motif of E7 and Rb drives the transformation by the viral oncogene. Simian virus 40 large T and adenovirus E1A protein also inhibit the tumor suppressor activity of Rb[48].

E1A and E7 are the two viral oncogenes that we found to have a specific effect on the cGAS-STING pathway. These proteins share functionally homologous regions named conserved region (CR) 1, 2, and 3. The LXCXE motif is present in CR2 and largely responsible for the binding and oncogenic activity. Additionally, the CR1 region contributes to the destabilization of pocket proteins[49]. The CR3 region contains a conserved zinc finger motif that also to contributes to transformation as mutations in this region have an effect on the binding affinity to Rb[50].

Our data showing antagonism of the cGAS-STING DNA pathway by hAd5 E1A and HPV18 E7 is dependent on their LXCXE motifs suggests a parallel mechanism of antagonism. Thus we wanted to explore the mechanism of antagonism of the by identifying the cellular targets of hAd5 E1A and HPV 18E7. The most obvious candidates to test were the three members of the Rb gene family to see what effect, if any, they had on the cGAS-STING DNA pathway. If Rb proteins were not important, we would proceed to take a closer look at the proximal components of cGAS-STING signaling and to determine if they are host targets of viral oncogenes.

3.2 Results

Viral oncogenes interact with signaling adapter STING

To test whether the three members of the Rb gene family (Rb, p107, p130) were important for cGAS-STING pathway activation, we took primary mouse macrophages and lentiviral shRNA-mediated the depletion of each of these Rb family members[51]. We confirmed knockdown of each Rb protein by western blot (Fig. 3A) and upon DNA transfection, found that knockdown cell lines responded normally (Fig. 3B). Moreover, triple knockout MEFs lacking Rb, p107, and p130 similarly mounted a potent DNA response, despite their documented growth abnormalities (Fig. 3C)[52]. This demonstrated that the host target(s) of viral oncogenes in the cGAS-STING pathway is distinct from their well-established targets in cell cycle regulation.

We then tested whether E1A and E7 targeted a known proximal signaling component of the cGAS-STING pathway. We hypothesized that E1A and E7 would block the cGAS-STING pathway through direct binding to a cellular protein, similar to how theses oncoproteins disrupt Rb. cGAS was our first target candidate, as it was recently discovered as the DNA sensor and an attractive target for viral oncogenes for inhibiting a type I IFN response by preventing recognition of ligand at the source. However, we did not detect an interaction between viral oncoproteins and cGAS in co-immunoprecipitation assays (data not shown).

Next we decided to look at the adapter STING as a potential target of E1A and E7 for inhibiting the cGAS-STING pathway. We took HEK 293 that we have shown to have higher protein expression of STING and E1A (data not shown) and transfected an epitope-tagged human STING at high to low concentrations of 10ug to 1ug of the expression

vector. Remarkably, we observed a robust interaction between STING and hAd5 E1A (Fig. 4A). Moreover, we noticed that 13S E1A interacted more robustly with STING compared to 12S E1A, especially at lower concentrations of STING expression vector (Fig. 4A). Indeed, a direct comparison of STING binding to FLAG epitope-tagged 13S E1A or 12S E1A revealed a specific interaction with 13S E1A and a much weaker interaction with 12S E1A (Fig. 4B).

We next tested whether specific domains of STING were important for interaction with E1A. We designed the following constructs: N-terminal STING (NTD) which includes the first 148 amino acids comprising the transmembrane domains that anchor it to the endoplasmic reticulum membrane, C-terminal STING (CTD) comprising of amino acids 149-379 which encode a cytosolic signaling domain that mediates STING dimerization and binding to cyclic dinucleotides and IRF3[13,14]. We also considered that STING exists as a dimer and is anchored to the ER, thus we made C-terminal STING (149-379aa) that was N-terminally anchored to the ER membrane by a heterologous transmembrane domain derived from cytochrome p450[53]. We were unable to express the N-terminus of STING to appreciable levels in HEK 293 cells, suggesting instability of this domain in live cells (data not shown). However, using epitope tagged constructs encoding STING amino acids 149-379, or 149-379 anchored to the ER membrane by a heterologous transmembrane domain, we found that the isolated C-terminus of STING is insufficient to mediate interaction with E1A (Fig. 4C), revealing that additional elements of STING are required for robust interaction with E1A. We then overexpressed STING in HeLa cells and detected an interaction with the endogenous HPV18 E7 protein (Fig. 4D). Interestingly, we

reproducibly detected a ~35 kDa protein detected by the E7 antibody that also specifically interacted with STING (Fig. 4D, asterisk), the nature of which is unknown.

We tested whether the LXCXE motif in E1A that is essential for blockade of the cGAS-STING pathway (Fig. 2) is also important for E1A binding to STING. To do this, we generated FLAG epitope-tagged 13S E1A constructs containing either the native LXCXE motif or the mutant VXSXD motif. We transfected these constructs into HEK 293 cells and achieved equivalent expression of the corresponding proteins (Fig. 4E). Interestingly, we found that the VXSXD mutant of E1A was severely compromised for interaction with STING-HA (Fig. 4E). Together, these data identify STING as a novel binding partner of E1A and E7, and suggest a potential mechanism of cGAS-STING pathway blockade by viral oncogenes. Importantly, the contribution of the LXCXE motif of E1A to STING binding precisely mirrors the role of this same motif in Rb binding, revealing a mechanistic parallel between antagonism of the DNA-activated antiviral response and inactivation of a tumor suppressor.

This data reveals that E1A and E7 can both bind to STING and inhibit the cGAS-STING pathway response to DNA. We wanted to explore what effect these viral oncogenes have on cyclic dinucleotide (CDN) activation of the cGAS-STING pathway. As CDNs bind to STING and initiate the cascade of signaling events for the production of type I IFN, we wanted to know if cGAMP production by cGAS was still functional when we stimulated cells expressing E1A or E7 with DNA ligand, and if STING could be activated by cGAMP in these same cells. To answer if viral oncogenes impacted the function of cGAS to produce cGAMP, we took HeLa cells that endogenously express E7 and stimulated them with CT DNA and measured cGAMP production. We also confirmed protein expression of

cGAS in HeLa cells by western blot. We were surprised to see that cGAS expression in HeLa was much higher than in primary HFFs. In contrast, HEK 293 expressed little to no cGAS protein. In accordance with cGAS protein expression, HeLas robustly produced cGAMP when stimulated with DNA compared to HFFs. HEK 293 served as a negative control showing no production of cGAMP when stimulated with DNA. As a control, we also stimulated HeLa cells, HEK 293, and HFFs with RIG-I ligand and observed no detectable production of cGAMP.

To address how cGAMP activation of STING is impacted by in the presence of the viral oncogenes, we again used HeLa cells and HEK 293 as they expressed E7 and E1A respectively. We stimulated these cells with cGAMP and measured type I IFN protein by bioassay. In both cell lines, there was no type I FN response following cGAMP stimulation (Fig. 5). In contrast, we were able to detect type I IFN in response to cGAMP in HFFs. For all three cell types tested, we were able to detect type I IFN with RIG-I stimulation (Fig. 5). Collectively, these results suggest E1A and E7 do not inhibit cGAS activation, as we were able to detect production of cGAMP in the presence of E7. Whether this also holds true for E1A will need to be determined. These results were corroborated in recent findings where HeLa cell line can produce cGAMP in response to DNA[54]. Additionally, these results confirm antagonism of the cGAS-STING pathway by the viral oncogenes is directed at STING. Addition of cGAMP did not bypass the antagonism of either viral oncogene, suggesting the region of contact on STING may be shared by both viral oncogenes and cGAMP (and possibly other CDNs). Further experiments will be needed to determine if viral oncogenes and cGAMP have overlapping contact regions in STING.

Restoring the cGAS-STING pathway in human tumor cell lines

Given we had found the target of E1A and E7 is STING, we wanted to formally test if the active presence of viral oncogenes was responsible for the inhibition of cGAS-STING signaling. We returned to the transformed human cells that are specifically unresponsive to stimulation with DNA. We hypothesized that the constitutive expression of E1A (HEK 293) or E7 (HeLa) renders these cells unresponsive to DNA due to failed to activation of the cGAS-STING pathway. To test this, we used a lenti-CRISPR approach[55] to disrupt expression of these oncogenes. We transduced HEK 293 cells with lentiviruses encoding the Cas9 endonuclease and two distinct guide RNAs targeting the integrated E1A gene, selected for transduced cells for three days, and achieved robust depletion of E1A protein in both targeted lines (Fig. 6A). Remarkably, we found that the E1A-targeted HEK 293 cells were able to mount an IFN response to DNA transfection, but cells transduced with control guide RNAs remained as unresponsive as the parental cells (Fig. 6B). We next used the same lenti-CRISPR approach to target the ~12 copies of E7 in HeLa cells (Fig. 6C). We observed robust recovery of the cGAS-STING pathway in these E7-depleted HeLa cells (Fig. 6D). Moreover, we found that HeLa cells transduced with a retrovirus encoding the bovine papillomavirus E2 protein, which silences the promoter encoding the bicistronic HPV18 E6/E7 mRNA[56], also recovered a potent IFN response to DNA transfection (Fig. 7). Together, our data demonstrate that HEK 293 cells and HeLa cells both retain a functional cGAS-STING pathway, but it is inhibited by the constitutive expression of the viral oncogenes that transforms these cells, illustrating the potent activity these viral oncogenes have on inhibiting an antiviral response.

3.3 Discussion

The evasion of the type I IFN response is a critical feature of viruses to ensure successful replication and infection. Thus viruses have employed numerous strategies to avoid detection or inhibit the response. Previous literature implicated that both adenovirus E1A and HPV18 E7 possessed mechanisms by which they could suppress a type I IFN response. E1A can disrupt the transcription of ISGs through the antagonism of chromatin remodeling complexes, while HPV 18 E7 can bind to IRF9 and prevent its nuclear translocation[32,17]. These evasion strategies are part of the secondary type I IFN response, which is downstream of IFNAR receptor. Each strategy also represents generic mechanisms to avoid the type I IFN response as they block common components of the RNA and cGAS-STING DNA pathway and the secondary interferon response. Our data are one of the first to show specific antagonism of the cGAS-STING DNA pathway by these viral oncogenes, as they specifically target the DNA signaling adapter STING, blunting the primary interferon response. In support that viruses have employed strategies to target the cGAS-STING pathway, a recent report showed that the viral protein viral interferon regulatory factor 1 (vIRF1) from Karposi's sarcoma-associated herpesvirus (KSHV) could bind STING and block its activation. This suggests that the targeting of STING is a general mechanism viruses have developed to inhibit innate immune signaling in a pathway-specific manner[57].

We also revealed the LXCXE motif of viral oncogenes not only contributes to inhibiting signaling of the cGAS-STING pathway, but that this motif is absolutely crucial for binding to STING. The LXCXE motif is highly conserved in the CR2 region of both E1A and E7. Interestingly, our data showed the 13S isoform of E1A binds STING with

higher affinity than the 12S isoform. The difference between the two isoforms is a 46 amino acid consensus zinc finger motif (CXXC) in CR3 of E1A present in the 13S but not 12S isoform. This suggests that the CR3 region plays a contributing role to the LXCXE containing CR2 region in binding to STING. As the addition of zinc and omission of EDTA in our protein lysis buffer enhances the interaction between E1A and STING, this strongly suggests the zinc finger in the CR3 region contributes to the binding of STING (data not shown). Considering that viral oncogenes bind Rb through the LXCXE CR2 region, studies have shown that CR2 region alone may not be sufficient to displace E2F from Rb, and that a contribution from the CR3 region is needed to efficiently disrupt E2F binding to Rb. The mechanism in which viral oncogenes bind and inhibit Rb may parallel as the viral oncogenes require CR2 and CR3 to effectively bind STING. The requirement of the LXCXE motif in inhibiting the signaling of the cGAS-STING DNA pathway makes it difficult to address if immortalization and transformation function of these viral oncogenes can be separated from their inhibition of cGAS-STING DNA pathway. Because the mechanism of the two is centered on the LXCXE motif, it is unlikely either function can be seen independent of the other.

Despite showing interaction of E1A and E7 to STING through co-immunoprecipitation assays, this does not rule out that this might not be a direct interaction between the viral oncogene and STING. There could be an associate or accessory protein facilitating this interaction. This would not be surprising as Rb repression and association with E2F is through the recruitment of multiple proteins in the formation of a complex. This includes the histone deacteylases (HDACs) HDAC1 and HDAC2, which have an LXCXE-like sequence that allows them to interact with the Rb pocket. Studies have also shown that a

bridging protein RBP1 with an LXCXE motif mediates the interaction between HDACs and Rb[58], suggesting there are multiple proteins involved in the regulation of the cell cycle. One way to address if the viral oncogenes interact with STING is through immunofluorescence. We can stain for proteins endogenously like E1A protein in HEK 293 or E7 protein in HeLas with STING and assess their localization. In addition, we can epitope-tag each the viral oncogenes and STING to overexpress them in cell lines if antibodies limited us. It would also be worth checking if localization of viral oncogenes and STING changes upon ligand engagement. Studies have shown E1A largely resides in the nucleus and has nuclear exporting sequence and is also present in the cytosol, while E7 is thought to reside in the cytosol[59]. STING is N terminally ER resident while the C-terminus is in the cytosol. How and where these viral oncogenes interact with STING, and how this is affected by the presence of ligand, needs to be further explored.

cGAS-STING pathway inhibition by viral oncogenes occurs at the DNA adapter STING and not at the DNA sensor cGAS, as cGAS function is not affected by the expression of E7 since HeLa cells that express cGAS are able to produce cGAMP upon DNA stimulation. It is still unclear how viral oncogenes interfere with the activating activity of cGAMP on STING as the expression of E1A and E7 prevented the activation of type I IFNs. Further experiments will need to be done to figure out the exact mechanism, but we can hypothesize cGAMP and viral oncogenes bind STING in the same region. Site-directed mutagenesis of amino acids in STING may help reveal what region(s) are involved. If mutagenesis results only reveals inhibition of binding from cGAMP and not viral oncogene or vice versa, this might support the caveat that viral oncogenes indirectly target STING and there is an accessory protein involved in this antagonism.

Our data also revealed the presence of the viral oncogenes was sufficient for the inhibition of the cGAS-STING pathway, as CRISPR targeting of E1A and E7 restored the type I IFN response in these cell lines. We also performed the experiment where we expressed in HeLa cells the protein regulator BPV-E2, which binds to the promoter of E6 and E7 and inhibits their expression, and then looked at the type I IFN response to DNA. Interestingly, cells expressing BPV-E2 showed a response to DNA. This dispels the idea immortalized cell lines have a blunted type I IFN response because of the absence of STING. Rather, the expression of viral oncogenes is responsible for this blunted type I IFN response.

3.4 Figures

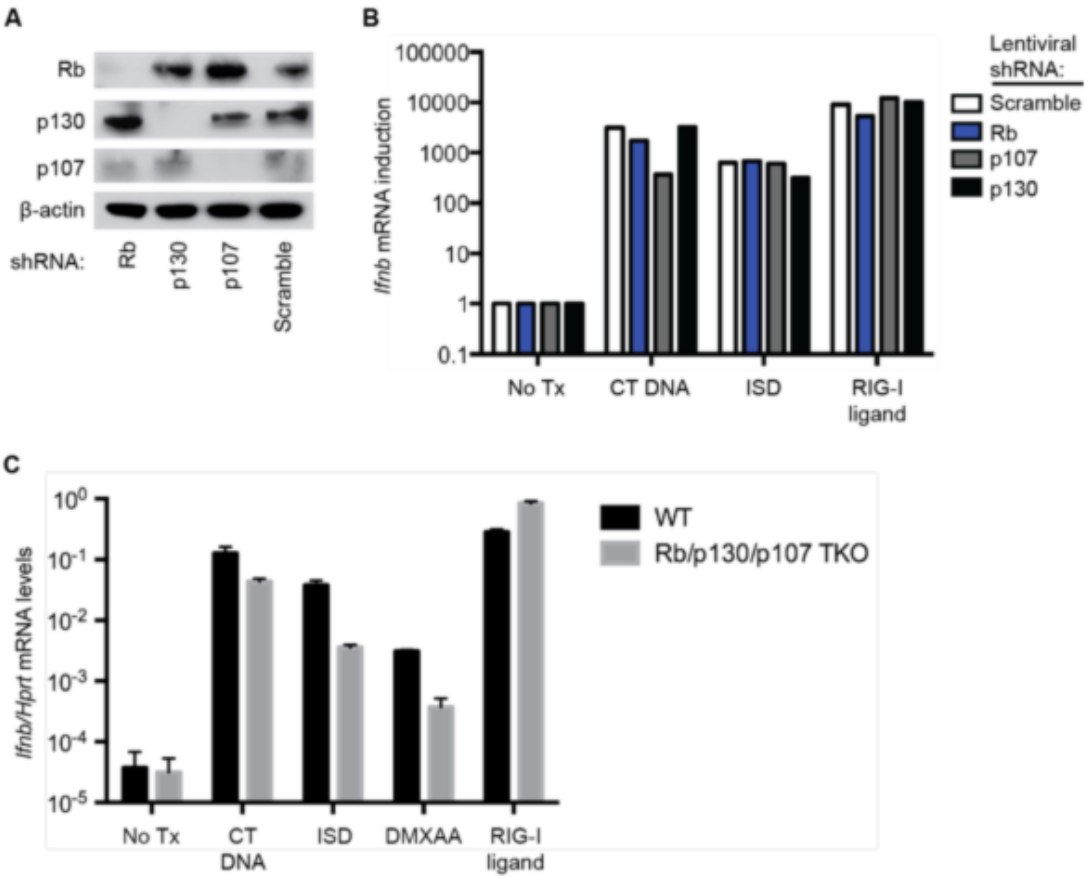

Figure 3: The Retinoblastoma family is dispensable for cGAS-STING pathway activation

(A) Bone marrow-derived macrophages (BMDMs) were transduced with lentiviral shRNAs targeting the indicated mRNAs or a scramble control. Extracts were prepared for western blot analysis of the indicated proteins.

(B) BMDMs transduced as in (A), were treated for four hours with the indicated ligand. *Ifnb* mRNA induction was quantitated by qPCR, normalized to *Hprt* mRNA, and compared to untreated cells to calculate induction.

(C) MEFs from WT or RB/p107/p103 TKO mice were treated with the indicated ligands and measured for *Ifnb* mRNA induction at four hours.

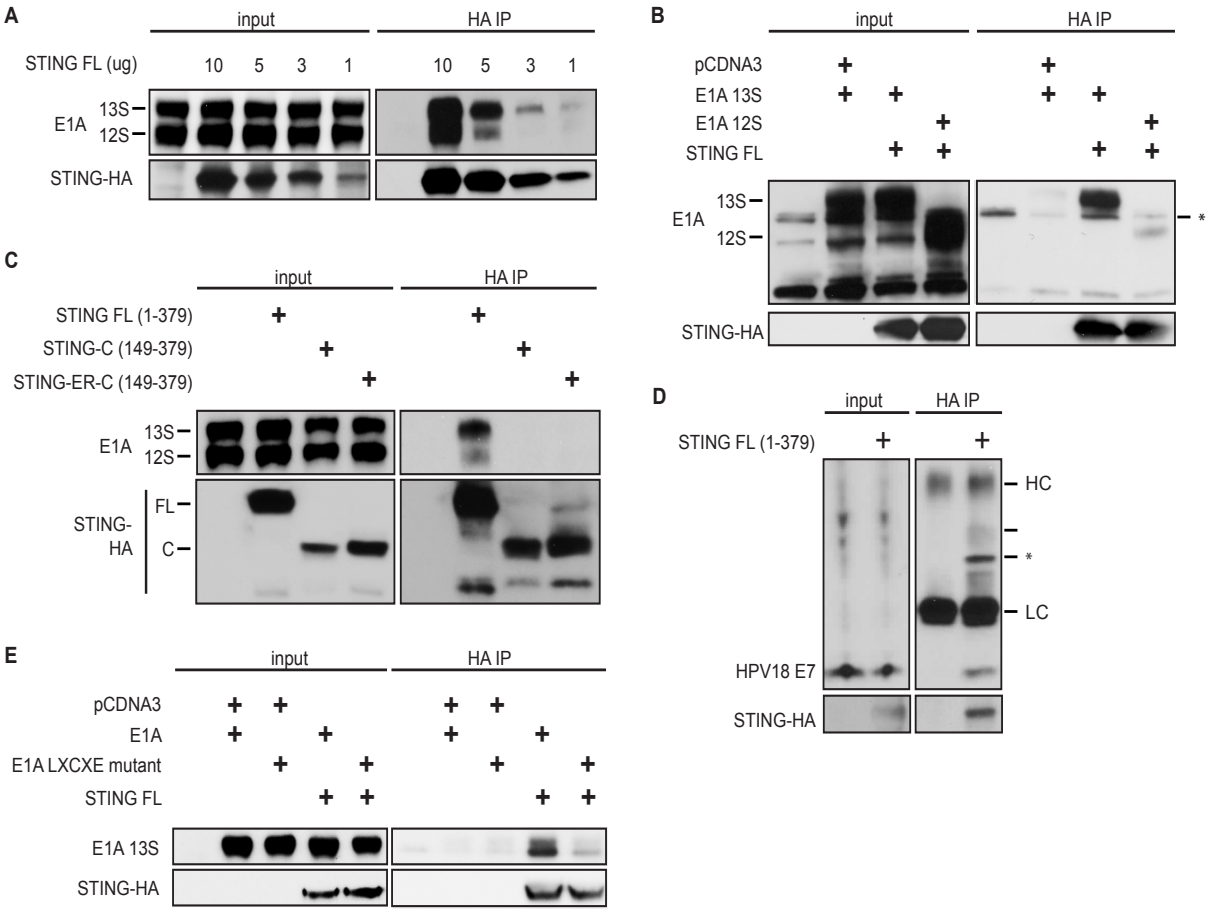

Figure 4: Viral oncoproteins bind to STING using the LXCXE motif

(A) HEK293 cells were transfected with the indicated concentrations of human STING-HA expression vector, followed by immunoprecipitation of STING-HA and western blotting for endogenous E1A protein or STING-HA. Inputs are shown in the left panels, and immunoprecipitations are in the right panels. **(B)** HEK 293 cells were transfected with 5 mg human STING-HA expression vector, together with either 13S E1A-FLAG or 12S E1A-FLAG, followed by immunoprecipitation of STING-HA and western blotting for E1A-FLAG or STING-HA. **(C)** HEK293 cells were transfected with the indicated STING-HA expression vectors, and processed for immunoprecipitations and western blot as in (A). **(D)** HeLa cells were transfected with full-length (FL) STING-HA, followed by immunoprecipitation of STING-HA and western blotting for endogenous E7 protein or STING. The asterisk indicates an unknown E7 antibody-immunoreactive protein migrating at approximately 35 kilo Daltons. HC, immunoglobulin heavy chain. **(E)** HEK 293 cells were transfected with plasmids encoding wild-type or VXSXD-mutant E1A-FLAG, together with STING-HA expression vector, followed by immunoprecipitation of STING-HA. All results are representative of two (B) or three (A,C, D-E) independent experiments.

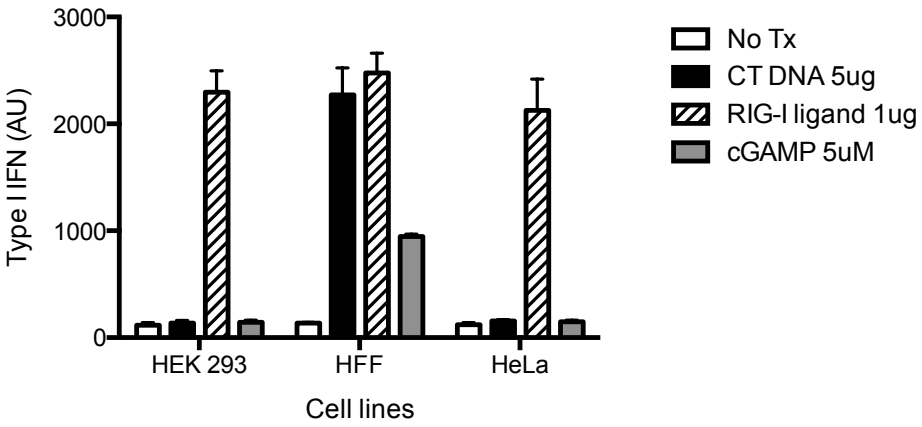

Figure 5: Viral oncogenes inhibit cGAMP activation of type I IFNs
HEK 293, HFF, and HeLa cell lines were treated with the indicated ligands for 8 hours, followed by measurement of type I IFN activity in culture supernatants.

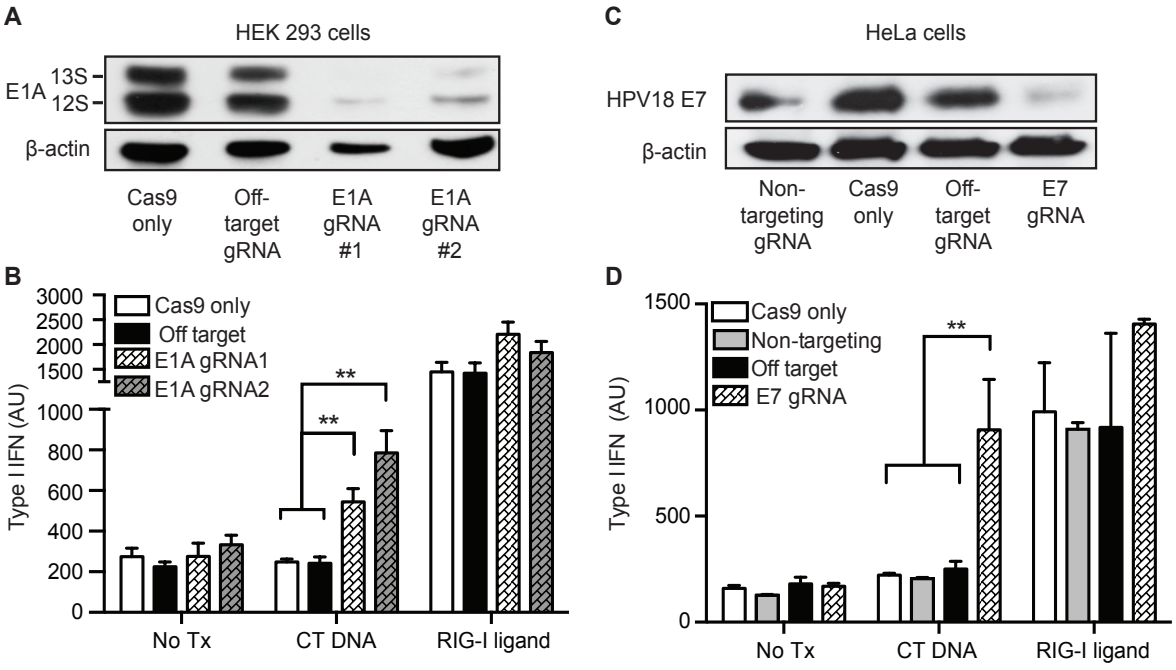

Figure 6: Restoration of the cGAS-STING pathway in human tumor cells.
(A to B) HEK 293 cells were transduced with Lenti-CRISPR constructs containing the indicated guide RNAs (gRNA), selected for three days in puromycin, and then harvested for evaluation of E1A protein expression by western blot **(A)**, or for plating prior to stimulation with the indicated ligands **(B)**. Type I IFN activity was measured in culture supernatants eight hours post stimulation. **(C to D)** HeLa cells were transduced with Lenti-CRISPR constructs containing the indicated gRNAs, selected for three days in puromycin, and then harvested for evaluation of E7 protein expression by western blot **(C)**, or for plating prior to stimulation with the indicated ligands **(D)**. Type I IFN activity was measured in culture supernatants eight hours post stimulation. Data are representative of three independent experiments.

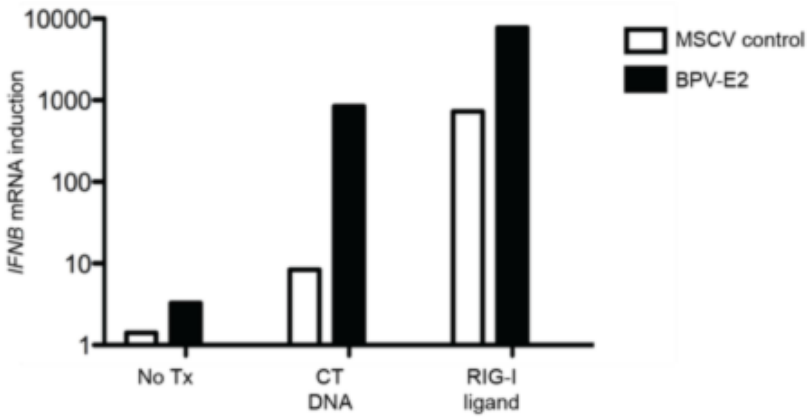

Figure 7: Silencing of E6/E7 with BPV E2 restores the cGAS-STING pathway signaling in HeLa cells

Hela cells were transduced with control retrovirus (MSCV) or retrovirus encoding the BPV E2 gene. Three days post transduction, cells were split and then stimulated with the indicated ligands. *IFNB* mRNA was measured by quantitative RT-PCR, normalized to *GAPDH* mRNA, and then compared to untreated MSCV-transduced cells to calculate the fold induction. BPV E2 transduction resulted in a 3-fold elevation of *IFNB* mRNA in untreated cells, a 100-fold increase in the cGAS-STING pathway, and a 10-fold increase in the RIG-I response.

3.5 Materials and Methods

Plasmids

The human STING open reading frame was cloned from HEK 293 cells, a C-terminal haemagglutinin (HA) epitope tag was introduced, and the resulting ORF was cloned into pCDNA3. STING mutants were generated by PCR-based mutagenesis. Adenovirus-5 E1A 13S with C-terminal FLAG epitope tag was PCR amplified from the E1A retroviral vector encoding huAd5 E1A (see above), and then cloned into pcDNA3 using the In-Fusion HD cloning kit (Clontech). The Adenovirus 5 E1A 12S open reading frame was PCR-amplified from HEK 293 cells using primers to incorporate a C-terminal FLAG tag.

Immunoprecipitations and western blots

HeLa cells were plated at 2.5 x 10^6 cells in a 10cm tissue culture plate and transfected the next day using Fugene 6 with 10 mg pcDNA3 or full length huSTING-HA expression vector. Cells were harvested 24 hours post transfection and resuspended in HeLa lysis buffer (20mM Tris HCl pH 7.4, 150mM NaCl, 0.5% NP-40, 0.5mM DTT, 2mg/ml PMSF, 0.1mM Zinc acetate) supplemented with Complete protease inhibitor cocktail (Roche). After 10 minute incubation on ice, lysates were cleared of insoluble material by centrifugation. For immunoprecipitations, lysates were incubated with anti-HA matrix (Roche), washed extensively in lysis buffer, separated using a 4-12% Bis-Tris SDS-PAGE gel (Life Technologies), and transferred to Immobilon-P PVDF membrane (Millipore). Membranes were probed with HPV18 E7 antibody (F-7; Santa Cruz Biotechnology) or anti-HA (Cell Signaling) followed by goat anti-mouse-HRP (Jackson ImmunoResearch), and then developed by ECL chemiluminescence (Pierce).

HEK 293 cells were plated at 2.5 x 10^6 cells in a 10cm tissue culture plate and transfected the next day with 5μg pcDNA3, 5μg full length human STING-HA (aa 1-379), 10μg STING C-terminus-HA (aa 149-379), or 10μg ER-STING C-terminus-HA. Cells were harvested the next day and resuspended in HEK 293 lysis buffer (20mM HEPES pH 7.4, 150mM NaCl, 10% glycerol, 1% Triton-X 100, 1mM DTT, 0.1mM Zinc acetate) supplemented with Complete protease inhibitor cocktail (Roche). Lysates were incubated on ice for 15 minutes and insoluble material was then cleared by centrifugation. Immunoprecipitations and SDS-PAGE separation were done as described above. Membranes were probed with Adenovirus-5 E1A antibody (M58; Santa Cruz Biotechnology) and developed as described above.

Protein expression of RasV12, SV40 large T antigen cDNA, huAd5 E1A, and HPV18 E7 in transduced primary MEFs were checked by plating 0.5 x 10^6 cells in a 6 well tissue culture plate while HPV18 E6 was plated at 1.0 x 10^6 and harvested the next day with either NP-40 or Trition-X 100 lysis buffer as described above. Lysates were incubated on ice for 15 minutes and insoluble material was then cleared by centrifugation. Lysates from RasV12, SV40 large T antigen cDNA, huAd5 E1A were separated using a 4-12% Bis-Tris SDS-PAGE gel (Life Technologies). Lysates from HPV18 E6 and HPV18 E7 were separated using a 12% Bis-Tris SDS-PAGE gel (Life Technologies). All lysates were transferred onto Immobilon-P PVDF membrane (Millipore). Membranes were probed RasV12 (D2H12; Cell Signaling Technology), SV40 large T (PAb416; Abcam), Adenovirus-5 E1A (see above), and HPV18 E6 (G-7; Santa Cruz Biotechnology), HPV18 E7 (see above) and β-actin (AC-74; Sigma) followed by goat anti-mouse- or goat anti-rabbit-HRP and then developed as described above.

Detection of cGAMP by liquid chromatography-tandem mass spectrometry

HEK 293T, HFFs, and HeLa cells were plated at $0.15x10^5$ cells/well in triplicate in a 12 well TC plate (BD Falcon) and stimulated as described in Section 2.5. Cells were then washed 2x with 1ml of col 0.9% NaCl. Ice cold 0.5ml of 80% MeOH was added to each sample and collected. Extracts were then spiked with isotope-labeled cGAMP (cGAMP*) at 5nM, sonicated, and cleared once by centrifugation at 16000xg, lyophilized, and resuspended in water. Detection of cGAMP in cell extracts was performed by liquid chromatography-tandem mass spectrometry as described (Huynh/Woodward, PNAS 2015 + Gray EE JI 2015 cGAS paper).

CRISPR/Cas9 targeting

For CRISPR-Cas9 targeting, we developed a lentivirus vector similar to the one recently reported by Zhang and colleagues (*28*) in which a U6 promoter-driven guide RNA and a MND promoter-driven Cas9-T2A-puromycin resistance cassette are constitutively expressed from a single, self-inactivating lentivirus upon integration into the host cell genome. HEK 293 cells or HeLa cells were transduced and selected in puromycin for three days, and CRISPR targeting of the E1A or E7 loci was evaluated by restriction fragment length polymorphism (RFLP) using restriction sites that overlapped the CRISPR targeting sites, as well as western blot for endogenous E1A (HEK 293) or E7 (HeLa) proteins. HeLa cells were additionally treated with the caspase inhibitor Q-VD-OPH (SM Biochemicals) for 24 hours prior to harvest. This treatment slightly improved cell recovery but did not affect the IFN response, because E7-targeted HeLa cells without Q-VD-OPH treatment similarly recovered cGAS-STING pathway signaling. Three days post selection, cells were

split into 12-well tissue culture plates at 1.5 x 10^5 cells per well, rested overnight, and stimulated for eight hours before harvest of supernatants for IFN bioassay. The sequences of the guide RNA target sites are as follows, with the protospacer adjacent motif (PAM) sequence underlined: non-targeting H1 gRNA: 5'-(G)ACGGAGGCTAAGCGTCGCAA, where the (G) denotes a nucleotide added to enable robust transcription off the U6 promoter; off-target AIM2 gRNA: GGAGATGTTTCATGCTACAGTGG; E1A gRNA #1: 5'- (G)TAGATTATGTGGAGCACCCCGGG-3'; E1A gRNA #2: 5'-(G)AAGACCTGCAACCGTGCCCGGGG-3'; E7 gRNA: 5'-GACGTTGTGGTTCAGCTCGTCGG.

Chapter 4: Exploring the Functional Impact of STING Single Nucleotide Polymorphisms in Cervical Cancer

4.1 Introduction

A central question of immunology is how the immune system of the host can recognize harmful antigens derived from pathogens while remaining unresponsive to harmless antigens that are often derived from self. The relationship between host and pathogens are intimately connected as the host often employs receptors to recognize unique patterns derived from pathogens (TLRs are a primary example), while pathogens simultaneously evolved mechanisms to avoid a host response. A consequence of this host-pathogen conflict is positive selection in the form of genetic variants in proteins that are intimately involved in host-pathogen interactions. These genetic variants are commonly single nucleotide polymorphisms (SNPs) within a protein. Genetic variants in important components of host innate immune signaling pathways can affect the response and susceptibility of the host to infection. This is particularly true with TLRs as these receptors directly recognize pathogen associated molecular patterns (PAMPs) on microbes. Specific examples of SNPs affecting the immune response of TLRs have been shown for TLRs 4 and 2: D299G in TLR4 is associated with narrowing airways in response to inhaled LPS, while R677W in TLR2 blunts the activation of NF-κB to *Mycobacterium tuberculosis* [60] [61]

Data from the 1000 Genome Project revealed that there are 4 major non-synonymous single nucleotide polymorphism (SNP) variants of STING within the human population [62]. One variant has three substitutions (R71H, G230A, and R293Q) and is usually referenced as HAQ. This particular SNP is the second most frequent in the population, and has been

previously reported as a loss-of-function human STING variant where it exhibits greater than 90% loss in the ability to induce IFNβ production[63]. There are two SNP variants that have a single amino acid substitution: R232H and R293Q. The last variant has 2 substitutions: G230A and R293Q. As mentioned earlier, STING binds cyclic dinucleotides (CDNs) to initiate the downstream activation of a type I IFN response. Due to importance of this direct interaction in host immune responses, it is likely that this interaction may drive the existence of STING SNP variants in the population. The 4 major SNPs can all recognize the metazoan 2'3' cGAMP, while their response to bacterial CDNs (cyclic di-GMP, cyclic di-AMP, and 3'3' cGAMP) varies[62]. This illustrates that all SNP variants have the capability of responding to CDNs, but that the signaling output may vary depending on the specific type of CDNs,

Given the role of STING in host defense and induction of type I IFN responses, it is no surprise that certain amino acids of STING across species are under positive selection. There are 6 positively selected amino acids of STING in bats, with 3 of these residues located in the transmembrane domain that are important for protein dimerization[64]. It has also been found that residue 230 in primate and bat phylogenies is undergoing positive selection. This site is also polymorphic in the human population as G230A, and the residue is located in the loop region of STING, which is predicted to form a lid that covers the CDN binding pocket[64]. The position of this residue and those neighboring are within the ligand binding domain of STING and is the interface where host and CDNs make contact thus giving away to selective pressures that impact host response to DNA from pathogens to trigger the cGAS-STING DNA pathway..

As we discovered the novel association of viral oncogenes to STING, we hypothesized that this class of antagonism may drive the need for host STING variants that can avoid viral antagonism. We decided to focus on human papilloma virus (HPV) because our findings had shown an association with E7 and human STING. We wanted to know if a particular STING SNP correlated with a risk in developing cervical cancer.

Infection with HPV is necessary but not sufficient for the development of cervical cancer, and it is estimated about 300 million women worldwide are infected with the virus. There are over 40 distinct types that can infect the genital tract, and about 90% of these infections are asymptomatic and effectively cleared within two years by an immune response that involves detection by innate immune receptors, subsequent cytokine responses, and infiltration of adaptive immune cells[65-69]. However, for the small percentage of infections that are not cleared, persistent infection with high-risk oncogenic HPVs can result in cervical cancer. High-risk HPVs include HPV16 and HPV18, as they are the found in about 70% of cervical cancer cases[68,70]. HPV infection occurs in the proliferating basal layer of epidermal or mucosal epidermal cells, where the viral DNA is released and transported to the nucleus as episomes. At this stage of infection, viral gene expression of early viral genes like E6 and E7 are largely suppressed. This allows for the proliferation of the infected cells and lateral expansion where the episomal copy number ranges from 50 – 100 genomes per cell[68,71]. The basal cells will eventually cease to divide as they migrate to the suprabasal layers and begin to differentiate. This is where late viral genes are activated and leads into the increased expression of E6 and E7. As a result, the viral genome replicates from hundred to thousands of copies as these cells proliferate. This is due to the ability of both viral oncogenes to negatively regulate the major tumor suppressors p53 and

Rb. E6 protein can bind to p53 and inhibits its regulation of cell growth arrest and initiation of apoptosis in infected cells, while E7 targets the proteasome dependent degradation of Rb[72].

As mentioned the majority of infected individuals clear the virus with a proper immune response. This infection is typically seen as a mild epithelial dysplasia and is characterized as a low-grade cervical intraepithelial neoplasia (CIN1). In contrast, there is a small population (10-15% of cases) that develops a persistent HPV infection due to a compromised immune response. This stage of the infection most likely shows increased expression of E6 and E7 and is characterized by moderate or severe epithelial dysplasia or a high-grade intraepithelial lesion (CIN2/3). From these cases, there is a small fraction that progress into invasive cancer where HPV DNA becomes predominantly integrated and is no longer in episomal form. The integration of HPV DNA is associated with chromosomal instability that is largely due to the expression of E6 and E7 in their targeting of p53 and Rb activity. Since the clearance of HPV infection is dependent on an effective immune response, we wondered if the different genetic variants of STING differed in their ability to coordinate an effective immune response to HPV. In collaboration with Margaret Madeleine and Denise Galloway of the Fred Hutchinson Cancer Research Center, we set out to assess if there was a correlation between STING SNPs and an increased risk of cervical cancer.

4.2 Results

Candidate SNPs: An initial approach and exploring the association of STING SNP R232H and cervical cancer

In order to investigate if genetic variants in STING were associated with increased risk of cervical cancer, we initially proposed a pilot study where we genotyped three of the four major coding SNPs of STING (R71H, R232H, R293Q as mentioned earlier) using 100 cervical cancer cases and 100 control subjects of predominantly European descent. These samples were previously used in a prior cervical cancer case-control study conducted by Dr. Madeleine[73]. The cervical cancer cases are from women who are 18-74 years old who were diagnosed with invasive squamous cervical cancer (SCC) while in the Seattle area. Controls were surveyed and were female residents of the same counties within the Seattle area and were age matched with an intact uterus for the study. We designed specific primers for non-synonymous coding SNPs rs11554776 at position 71 (R71H), rs1131769 at 232 (R232H), and rs7380824 at 293 (R293Q) in STING and proceeded to genotype each using a TaqMan assay. We found no association with either R71H or R293Q (data not shown) as our analysis calculated an odds ratio (OR) of 1, where by definition there is no correlation between exposure and outcome. Interestingly when we looked at R232H, we saw an OR of 2.1 for R/H heterozygotes, OR 2.4 for H/H homozygotes; log additive OR 1.8, 95% confidence interval (CI) 1.1-2.9 where an OR higher than 1 would suggest a correlation between exposure and disease (Table 1). In this pilot, R232H was present in about 17% of our controls and was recovered with higher frequency in our cancer patients at 30.6%. Thus from our initial pilot study, SNP R232H appeared to be associated with a risk of developing cervical cancer. Therefore we decided to expand our sample size to gain more statistical power.

In collaboration with Dr. Madeleine and Dr. Galloway, we were able to obtain an additional 1013 cervical cancer cases and 1068 controls from their previous study. In this larger screen, we decided to further interrogate R232H , and used SNP R293Q as a negative control. Using TaqMan assay and the same genotyping primers from our pilot study, we genotyped all cases and controls for R232H and R293Q. Despite the strong findings in our pilot study, we found in our more expansive screen that SNP R232H does not correlate with a risk for cervical cancer as the OR in this study was 0.98 with CI 0.82-1.16. Our findings for SNP R293Q remained the same as it had an OR 1.02 with CI 0.85-1.23 (Table 2).

4.3 Discussion

Numerous studies have shown SNPs in key immune signaling proteins influence the outcome of an infection[74,75]. The immune response to HPV is important for the clearance and control of the infection. A compromised immune response could lead to a chronic infection and potential progression to cervical cancer. Our data revealed an interaction between STING and the oncogenic protein E7 from HPV, which results in an impaired type I IFN response. Assuming that different STING SNPs might affect immune responses to HPV, we wanted to test if there was a correlation between specific alleles of STING and a risk of cervical cancer. Our results were initially promising with an association established between SNP R232H in STING and cervical cancer, but as we increased our sample size, that correlation did not hold statistically.

Despite not finding an association with SNP R232H and cervical cancer, there still is the question as to what is the functional relevance of different genetic variants of STING in the population and what selective pressures drive their existence. SNP R232H lies within the loop region of where CDNs bind. Included in this region are amino acids 230 and 231, which have both been implicated to have differential responses to CDNs when residues are mutated within that region. This suggests within this interface where host come in contact with CDNs in response to pathogens, it faces unique evolutionary pressure to maintain recognition of pathogens[62,76]. Therefore it could still be feasible that STING SNPs such as R232H do in fact have an affect on the immune response. However, with our STING SNP study, we were unable to reveal any immune consequence between the variants of STING we interrogated with cervical cancer. One explanation is that we only investigated 3 SNPs. Additionally, the population we sampled may not have been optimal, as it was

homogeneous in terms of ethnic background, and from a limited area in Seattle. To improve our approach in future studies, we should first to look into a more expansive sampling of cases with cervical cancer and controls. For instance, the International Cervical Cancer Consortium (ICCC) GWAS study includes 2,800 cervical cancer cases from 8 studies in North America, Europe, and Australia and 6,700 control from the Welcome Trust Case-Control Consortium. In addition to a more expansive sampling, there could be a larger screen on STING SNP and/or particular residues of STING that have functional consequences when deviated from the reference allele. This would be a more thorough interrogation of STING alleles beyond the 4 major non-synonymous SNPs in the population, and may shed light on why genetic variants in STING exist, as well as why do particular point mutations have such a functional impact.

4.4 Table

Table 1:	STING SNP R232H association with cervical cancer: 100 cancer cases and 100 control			
	Cases %	Controls %	OR	95%CI
SNP232 (rs1131769)				
R232R	83	69.4	1	reference
R232H	13	22.4	2.1	0.97-4.40
H232H	4	8.2	2.4	0.70-8.46
R/H, H/H	17	30.6	1.8	1.05-2.93

Table 2:	STING SNP R232H and R293Q genotyping results in association with cervical cancer: 1013 cancer cases and 1068 controls	
	OR	95%CI
SNP232 (rs1131769)	0.98	0.82-1.16
SNP293 (rs7380824)	1.02	0.85-1.23

4.5 Materials and Methods

STING SNP genotyping

Peripheral blood samples of cases and controls were collected, processed, and DNA was extracted as outlined in a previous study[73]. To type STING alleles, there were 2 sets of locus specific primers used. One was designed to amplify (rs11554776) R71H region while the second set could amplify both (rs1131769) R232H and (rs7380824) R293Q. Genomic DNA was amplified using BioRad DNAEngine for 45 cycles at an annealing temperature of 57°F. The PCR reactions were then diluted 1:10,000 and then qPCR reaction was performed using TaqMan probes that were specifically designed to distinguish between each allele within the SNP of interest. TaqMan probes were purchased from Applied Biosystems. Probe for (rs11554776) R71H:AGTGAGTCACCTGGAGTGGATGTGG [C/T]GCAGCTCCTCAGCCAGGCTGCAGAC, (rs1131769) R232H:GTAAACCCGA TCCTTGATGCCAGCA[C/T]GGTCACCGGTCTGCTGGGGCAGTTT, and (rs7380824) R293Q: ATCTGCCAGGATGTCCTCAAGTGTC[C/T]GGCAGAAGAGTTTGGCCTG CTCAAG. The qPCR reactions were run on ViiA 7 Real-Time PCR System for 40 cycles at an annealing temperature of 60°F.

Single-locus analysis

STING SNP genotype analysis was carried out as described in previous cervical cancer study[73]

Chapter 5: Concluding Remarks

We discovered a previously unappreciated function of DNA tumor virus oncogenes: they are potent and specific antagonists of the DNA-activated antiviral response. Until now, the role of DNA tumor virus oncogenes had been largely known in the context of cancer as remarkable advances over the last four decades have revealed how these oncogenes transform cells through inhibition of p53 and Rb[24,25,51]. Here, we've revealed an additional function of these viral oncogenes as antagonists of the DNA sensing pathway. Our findings also reveal STING as a novel binding partner of these viral oncogenes. As this association is facilitated through the same LXCXE motif that inhibits the function of tumor suppressors, this suggests their ability to inhibit an immune response to infection occurs through a parallel mechanism.

Two recent reports have shown proteins from Kaposi's sarcoma-associated herpesvirus (KHSV) also specifically block the cGAS-STING DNA pathway[57,77]. One report suggest that the tegument protein ORF52 can sequester DNA from cGAS and binds to cGAS to inhibit antiviral response, while the other reported the viral interferon regulatory factor 1(vIRF1) can target STING and prevent its interaction with TBK1. These independent studies support our findings that viruses have evolved particular antagonists to the cGAS-STING DNA sensing pathway. However, further work will need to be done to shed light on how viral oncogenes E1A and E7 inhibit cGAS-STING DNA sensing pathway. In particular, how the viral oncogenes affect downstream signaling, such as if they perturb phosphorylation and/or recruitment of proteins in the STING-TBK1-IRF3 axis[14], needs to be investigated. Regardless of the mechanism, a consequence of this antagonism of the host gene could be positive selection of STING variants resistant to antagonism as a means

retain a proper host immune response. Preliminary evolution analysis of primate STING indicates there is evidence of positive selection, which would highlight the importance of STING variants in countering the evasion mechanisms of viruses.

Additionally, our findings may shed new light on the long-standing question of why DNA tumor viruses evolved these oncogenes in the first place. There are two hypotheses that rationalize the origins of DNA tumor virus oncogenes, neither of which are mutually exclusive. The "limited resources hypothesis" postulates that oncogenes evolved to stimulate cell cycle progression in order to facilitate a cellular environment conducive to viral replication: abundant dNTPs for viral DNA synthesis, dissolution of the nuclear envelope to allow for viral egress, and new cellular niches to support transmission. In contrast, the "anti-antivirus hypothesis" [25] [24] holds that the principal function of viral oncogenes is to antagonize innate immune signaling. Thus the viral oncogenes ability to drive unregulated cell cycle division through the LXCXE motif and the progression to cancer was simply and unintended evolutionary consequence. Interestingly, most DNA viruses that infect humans encode proteins with LXCXE motif, but only a minority of these viruses are known to cause cancer[46], suggesting that this motif in viruses is most relevant and originally evolved for the purpose of evading a response from the host. For example, low-risk types like HPV6 can also bind Rb and possess a LXCXE motif, but cannot activate E2F for cell cycle progression into S phase. Thus one could hypothesize that the initial purpose of the LXCXE motif of viral oncogenes is to target the DNA sensing pathway to avoid an innate immune response. Our data support the latter of these two hypotheses and reveal a host-virus conflict that may have shaped the origins of the DNA viruses that cause cancer in humans.

References

1. Yoneyama, M. *et al*. The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nat Immunol* **5,** 730–737 (2004).

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

3. Yoneyama, M. *et al*. Shared and Unique Functions of the DExD/H-Box Helicases RIG-I, MDA5, and LGP2 in Antiviral Innate Immunity. *The Journal of Immunology* **175,** 2851–2858 (2005).

4. Ablasser, A., Hertrich, C., Waßermann, R. & Hornung, V. Nucleic acid driven sterile inflammation. *Clinical Immunology* **147,** 207–215 (2013).

5. Satoh, T. *et al*. LGP2 is a positive regulator of RIG-I- and MDA5-mediated antiviral responses. *Proc Natl Acad Sci USA* **107,** 1512–1517 (2010).

6. Kawai, T. & Akira, S. Innate immune recognition of viral infection. *Nat Immunol* **7,** 131–137 (2006).

7. Sun, L., Wu, J., Du, F., Chen, X. & Chen, Z. J. Cyclic GMP-AMP Synthase Is a Cytosolic DNA Sensor That Activates the Type I Interferon Pathway. **339,** 786–791 (2013).

8. Hornung, V. *et al*. 5'-Triphosphate RNA is the ligand for RIG-I. *Science* **314,** 994–997 (2006).

9. Barbalat, R., Ewald, S. E., Mouchess, M. L. & Barton, G. M. Nucleic Acid Recognition by the Innate Immune System. **29,** 185–214 (2011).

10. Pichlmair, A. *et al*. Activation of MDA5 Requires Higher-Order RNA Structures Generated during Virus Infection. *J. Virol*. **83,** 10761–10769 (2009).

11. Stetson, D. B. & Medzhitov, R. Recognition of Cytosolic DNA Activates an IRF3-Dependent Innate Immune Response. *Immunity* **24,** 93–103 (2006).

12. Wu, J. *et al*. Cyclic GMP-AMP Is an Endogenous Second Messenger in Innate Immune Signaling by Cytosolic DNA. **339,** 826–830 (2013).

13. Burdette, D. L. & Vance, R. E. STING and the innate immune response to nucleic acids in the cytosol. *Nat Immunol* **14,** 19–26 (2012).

14. Liu, S. *et al*. Phosphorylation of innate immune adaptor proteins MAVS, STING, and TRIF induces IRF3 activation. *Science* **347,** aaa2630–aaa2630 (2015).

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

16. Pachler, K. & Vlasak, R. Influenza C virus NS1 protein counteracts RIG-I-mediated IFN signalling. *Virology Journal* **8,** 48 (2011).

17. Kanodia, S., Fahey, L. M. & Kast, W. M. Mechanisms used by human papillomaviruses to escape the host immune response. **7,** 79–89 (2007).

18. Isaacs, A. & Lindenmann, J. *Virus interference. I. The interferon. By A. Isaacs and J. Lindenmann, 1957*. **7,** 429–438 (1987).

19. Goubau, D., Deddouche, S. & Sousa, C. R. E. Cytosolic Sensing of Viruses. *Immunity* **38,** 855–869 (2013).

20. Paladino, P. & Mossman, K. L. Mechanisms Employed by Herpes Simplex Virus 1 to Inhibit the Interferon Response. *Journal of Interferon & Cytokine Research* **29,** 599–608 (2009).

21. Zhao, D.-C., Yan, T., Li, L., You, S. & Zhang, C. Respiratory syncytial virus inhibits interferon-α-inducible signaling in macrophage-like U937 cells. *Journal of Infection* **54,** 393–398 (2007).

22. Loo, Y.-M. *et al*. Viral and therapeutic control of IFN-beta promoter stimulator 1 during hepatitis C virus infection. *Proc. Natl. Acad. Sci. U.S.A.* **103,** 6001–6006 (2006).

23. Stetson, D. B., Ko, J. S., Heidmann, T. & Medzhitov, R. Trex1 Prevents Cell-Intrinsic Initiation of Autoimmunity. *Cell* **134,** 587–598 (2008).

24. Moore, P. S. & Chang, Y. Why do viruses cause cancer? Highlights of the first century of human tumour virology. 1–12 (2010). doi:10.1038/nrc2961

25. Levine, A. J. The common mechanisms of transformation by the small DNA tumor viruses: The inactivation of tumor suppressor gene products: p53. **384,** 285–293 (2009).

26. Graham, F. L., Smiley, J., Russell, W. C. & Nairn, R. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J. Gen. Virol*. **36,** 59–74 (1977).

27. Boshart, M. *et al*. A new type of papillomavirus DNA, its presence in genital cancer biopsies and in cell lines derived from cervical cancer. *EMBO J*. **3,** 1151–1157 (2005).

28. de Villiers, E. M. *et al*. Human papillomavirus infections in women with and without abnormal cervical cytology. **2,** 703–706 (1987).

29. Bhattacharya, S. *et al*. Cooperation of Stat2 and p300/CBP in signalling induced by interferon-alpha. **383,** 344–347 (1996).
30. Juang, Y. T. *et al*. Primary activation of interferon A and interferon B gene transcription by interferon regulatory factor 3. **95,** 9837–9842 (1998).

31. Ronco, L. V., Karpova, A. Y., Vidal, M. & Howley, P. M. Human papillomavirus 16 E6 oncoprotein binds to interferon regulatory factor-3 and inhibits its transcriptional activity. **12,** 2061–2072 (1998).

32. Fonseca, G. J. *et al*. Adenovirus Evasion of Interferon-Mediated Innate Immunity by Direct Antagonism of a Cellular Histone Posttranslational Modification. **11,** 597–606 (2012).

33. Zhong, B. *et al*. The Adaptor Protein MITA Links Virus-Sensing Receptors to IRF3 Transcription Factor Activation. *Immunity* **29,** 538–550 (2008).

34. Sun, W. *et al*. ERIS, an endoplasmic reticulum IFN stimulator, activates innate immune signaling through dimerization. *Proc Natl Acad Sci USA* **106,** 8653–8658 (2009).

35. Saito, T., Owen, D. M., Jiang, F., Marcotrigiano, J. & Gale, M., Jr. Innate immunity induced by composition-dependent RIG-I recognition of hepatitis C virus RNA. *Nature* **454,** 523–527 (2008).

36. Lowy, D. R. & Willumsen, B. M. Function and regulation of ras. *Annual review of biochemistry* (1993).

37. Classon, M. & Harlow, E. The retinoblastoma tumour suppressor in development and cancer. *Nat Rev Cancer* **2,** 910–917 (2002).

38. McLaughlin-Drubin, M. E. & Münger, K. The human papillomavirus E7 oncoprotein. *Virology* **384,** 335–344 (2009).

39. Corbeil, H. B. & Branton, P. E. Functional importance of complex formation between the retinoblastoma tumor suppressor family and adenovirus E1A proteins as determined by mutational analysis of E1A conserved region 2. *J. Virol*. **68,** 6697–6709 (2015).

40. Ishikawa, H. & Barber, G. N. STING is an endoplasmic reticulum adaptor that facilitates innate immune signalling. *Nature* **455,** 674–678 (2008).

41. Schattgen, S. A. & Fitzgerald, K. A. The PYHIN protein family as mediators of host defenses. *Immunol. Rev*. **243,** 109–118 (2011).

42. Albrecht, M., Choubey, D. & Lengauer, T. The HIN domain of IFI-200 proteins consists of two OB folds. *Biochemical and Biophysical Research Communications* **327,** 679–687 (2005).

43. Brunette, R. L. *et al*. Extensive evolutionary and functional diversity among mammalian AIM2-like receptors. *Journal of Experimental Medicine* **209,** 1969–1983 (2012).
44. Unterholzner, L. *et al*. UnterholznerIFI16_NatIm2010. *Nature Publishing Group* **11,** 997–1004 (2010).

45. Li, X.-D. *et al*. Pivotal roles of cGAS-cGAMP signaling in antiviral defense and immune adjuvant effects. *Science* **341,** 1390–1394 (2013).

46. de Souza, R. F., Iyer, L. M. & Aravind, L. Biochimica et Biophysica Acta. *BBA - Gene Regulatory Mechanisms* **1799,** 302–318 (2010).

47. Lee, J.-O., Russo, A. A. & Pavletich, N. P. Structure of the retinoblastoma tumour-suppressor pocket domain bound to a peptide from HPV E7. *Nature* **391,** 859–865 (1998).

48. Whyte, P. *et al*. Association between an oncogene and an anti-oncogene: the adenovirus E 1 A proteins bind to the retinoblastoma gene product. **334,** 124–129 (1988).

49. Munger, K. *et al*. Mechanisms of Human Papillomavirus-Induced Oncogenesis. *J. Virol*. **78,** 11451–11460 (2004).

50. Todorovic, B. *et al*. Conserved Region 3 of Human Papillomavirus 16 E7 Contributes to Deregulation of the Retinoblastoma Tumor Suppressor. *J. Virol*. **86,** 13313–13323 (2012).

51. Weinberg, R. A. The retinoblastoma protein and cell cycle control. **81,** 323–330 (1995).

52. Sage, J. Targeted disruption of the three Rb-related genes leads to loss of G1 control and immortalization. *Genes & Development* **14,** 3037–3050 (2000).

53. Ahn, K., Szczesna-Skorupa, E. & Kemper, B. The amino-terminal 29 amino acids of cytochrome P450 2C1 are sufficient for retention in the endoplasmic reticulum. *Journal of Biological Chemistry* (1993).

54. Gentili, M. *et al*. Transmission of innate immune signaling by packaging of cGAMP in viral particles. **349,** 1232–1236 (2015).

55. Wang, T., Wei, J. J., Sabatini, D. M. & Lander, E. S. Genetic Screens in Human Cells Using the CRISPR-Cas9 System. *Science* **343,** 80–84 (2014).

56. Hwang, E. S. *et al*. Inhibition of cervical carcinoma cell line proliferation by the introduction of a bovine papillomavirus regulatory gene. *J. Virol*. **67,** 3720–3729 (1993).

57. Ma, Z. *et al*. Modulation of the cGAS-STING DNA sensing pathway by gammaherpesviruses. *Proc Natl Acad Sci USA* **112,** E4306–E4315 (2015).

58. Lai, A. *et al*. RBP1 recruits both histone deacetylase-dependent and -independent repression activities to retinoblastoma family proteins. *Mol. Cell. Biol.* **19,** 6632–6641 (1999).

59. White, E., Spector, D. & Welch, W. Differential distribution of the adenovirus E1A proteins and colocalization of E1A with the 70-kilodalton cellular heat shock protein in infected cells. *J. Virol.* **62,** 4153–4166 (2013).

60. Cook, D. N., Pisetsky, D. S. & Schwartz, D. A. Toll-like receptors in the pathogenesis of human disease. **5,** 975–979 (2004).

61. Bochud, P. Y., Hawn, T. R. & Aderem, A. Cutting Edge: A Toll-Like Receptor 2 Polymorphism That Is Associated with Lepromatous Leprosy Is Unable to Mediate Mycobacterial Signaling. **170,** 3451–3454 (2003).

62. Yi, G. *et al*. Single Nucleotide Polymorphisms of Human STING Can Affect Innate Immune Response to Cyclic Dinucleotides. **8,** e77846 (2013).

63. Jin, L. *et al*. Identification and characterization of a loss-of-function human MPYS variant. **12,** 263–269 (2011).

64. Mozzi, A. *et al*. OASes and STING: Adaptive Evolution in Concert. **7,** 1016–1032 (2015).

65. de Villiers, E.-M., Fauquet, C., Broker, T. R., Bernard, H.-U. & Hausen, zur, H. Classification of papillomaviruses. **324,** 17–27 (2004).

66. Stanley, M. Immune responses to human papillomavirus. **24,** S16–S22 (2006).

67. Frazer, I. H. Interaction of human papillomaviruses with the host immune system: A well evolved relationship. **384,** 410–414 (2009).

68. Shulzhenko, N., Lyng, H., Sanson, G. F. & Morgun, A. Ménage à trois: an evolutionary interplay between human papillomavirus, a tumor, and a woman. **22,** 345–353 (2014).

69. Ho, G., Bierman, R. & Beardsley, L. Natural history of cervicovaginal papillomavirus infection in young women. (1998). at <http://www.nejm.org/doi/full/10.1056/NEJM199802123380703>

70. Hausen, zur, H. Oncogenic DNA viruses. *Oncogene* **20,** 7820–7823 (2001).

71. Hausen, zur, H. Papillomaviruses and cancer: from basic studies to clinical application. *Nat Rev Cancer* **2,** 342–350 (2002).

72. Wang, J., Sampath, A., Raychaudhuri, P. & Bagchi, S. Both Rb and E7 are regulated by the ubiquitin proteasome pathway in HPV-containing cervical tumor cells. *Oncogene*

20, 4740–4749 (2001).

73. Madeleine, M. M. *et al*. Comprehensive Analysis of HLA-A, HLA-B, HLA-C, HLA-DRB1, and HLA-DQB1 Loci and Squamous Cell Cervical Cancer Risk. *Cancer Research* **68,** 3532–3539 (2008).
74. Arbour, N. C. *et al*. TLR4 mutations are associated with endotoxin hyporesponsiveness in humans. *Nature genetics* **25,** 187–191 (2000).

75. Shigemoto, T. *et al*. Identification of loss of function mutations in human genes encoding RIG-I and MDA5: implications for resistance to type I diabetes. *Journal of Biological Chemistry* **284,** 13348–13354 (2009).

76. Diner, E. J. *et al*. The Innate Immune DNA Sensor cGAS Produces a Noncanonical Cyclic Dinucleotide that Activates Human STING. *Cell Reports* **3,** 1355–1361 (2013).

77. Wu, J.-J. *et al*. Inhibition of cGAS DNA Sensing by a Herpesvirus Virion Protein. *Cell Host and Microbe* **18,** 333–344 (2015).