SUPPLEMENTARY METHODS AND EXPERIMENTS

Design of Experimental Controls. For HSATII and GSAT negative controls were designed in two ways and both negative controls were compared to HSATII and GSAT for all experiments. First, full RNA sequences of both satellites were randomly permuted until scrambled sequences were generated that fell within one half of a standard deviation from the mean value of the strength of statistical bias against CpG and UpA dinucleotides for humans and mice respectively. These sequences are denoted as HSATII-sc and GSAT-sc. In other words these sequences had the same length and nucleotide content as HSATII and GSAT but fell within the inner ellipse in Figures 2a (HSATII-sc) and Figure 2b (GSAT-sc). In addition we checked that in both cases the minimum RNA folding energy was not lowered during the scrambling process so that our permutations did not seem to produce more RNA secondary structure thereby creating the possibility of innate immune stimulation via TLR3. The free energy was calculated using the MATLAB RNAfold routine (1,2). We created endogenous negative controls by searching Repbase for the repetitive elements that fell within one standard deviation of the mean bias against CpG and UpA in humans and mice but were also closest in length to HSATII and GSAT. These were UCON38 for HSATII and RMER16A3 for GSAT.

GSAT RNA Expression Level Detection. GSAT RNA expression levels were investigated by a custom Taqman Assay in normal mouse tissue versus mouse tumor tissue samples (Supplementary Figure 3). The tumor mouse models that were investigated were a model of testicular teratoma (p53-/- 129/Sv^{SL}) and a model of liposarcoma ($p53^{LoxP/LoxP}$;$Pten^{LoxP/LoxP}$). In all instances GSAT levels were increased in the tumor samples as compared to normal samples however to varying degrees. There was no significant difference in GSAT levels between tumors arising in females versus those arising in males in the liposarcoma model. Also there was no difference in GSAT levels in p53-/- 129/Sv^{SL} that developed teratomas at a young age (~1 month old) versus at an older age (~3-4 months old) (3,4).

i-ncRNA generation. Sequences encoding for murine GSAT and human HSATII were generated by custom gene synthesis (Genscript) and cloned into a pCDNA3 backbone (EcoRI/EcoRV) that carries a T7 promoter on the + strand and a SP6 promoter on the – strand (Invitrogen). Sequences encoding for GSAT-sc, HSATII-sc, UCON38 and RMER16A3 were generated as minigenes and sub-cloned in a pIDT-blue backbone with a T7 promoter on the + strand and a T3 promoter on the – strand surrounding the sequence of interest (IDT). To produce high quality RNA, plasmids were digested by the restriction enzymes NotI/NdeI (pCDNA3) and ApaLI (pIDT blue) to isolate the fragment containing the sequence of interest by gel purification (Qiagen). Then the sequences of interest containing the T7 promoter were amplified by PCR (Accuprime-PFX Invitrogen) using the following primer pairs:
pIDT blue - Forward: GCGCGTAATACGACTCACTATAGGCGA;
Reverse: CGCAARRAACCCTCACTAAAGGGAACA) and

pCDNA.3 - Forward: GAAATTAATACGACTCAATAGG;
Reverse:TCTAGCATTTAGGTGACACTATAGAATAG).
PCR products were purified by PCR-Cleanup (Qiagen) and controlled by electrophoresis (0 8% Agarose gel). RNAs were generated by in-vitro transcription using the mMESSAGE mMACHINE T7 ultra kit (Ambion) followed by a capping and short polyA reaction. RNAs were then purified using RNA-cleanup (Qiagen) quantified using a nanodrop and checked by electrophoresis after denaturation at 65 C for 10 minutes (1 5% Agarose gel).

Cell stimulation. MoDCs and imBM were both stimulated by i-ncRNA in the same way. The culturing of these cells is described below. Briefly cells were plated in 96 flat well plates at 200,000 cells per well for primary cells (MoDCs) and 100,000 cells per well for lines (IMBM). i-ncRNA were transfected via liposomes formed using DOTAP (Roche Life Science) at a ratio of 1ug DNA per 6 ul DOTAP diluted in HBS following the user-guide recommendations. The cells were stimulated using 2ug/ml of purified i-ncRNA versus 10ug/ml total RNA. To stimulate the TLR4 pathway we used 100ng/ml Ultrapure LPS (Invivogen) for TLR2: 500ng/ml Pam2CSK4 (Invivogen) for TLR3: 2ug/ml HMW PolyIC (Invivogen) TLR7/8: 1ug/ml CLO97 (Invivogen) and 100 ng/ml R848 (Invivogen) TLR9: CpG B-ODN 1826 3uM or STING CDN 5ug/ml (Aduro).

Cell culture. *Human moDCs:* Human monocyte derived DCs were differentiated as previously described (5); briefly PBMCs were prepared by centrifugation over Ficoll-Hypaque gradients (BioWhittaker) from healthy donor buffy coats (New York Blood Center). Monocytes were isolated from PBMCs by adherence and then treated with 100 U/ml GM-CSF (Leukine Sanofi Oncology) and 300 U/ml IL-4 (RandD) in RPMI plus 5% human AB serum (Gemini Bio Products). Differentiation media was renewed on day 2 and day 4 of culture. Mature moDCs were harvested for use on days 5 to 7. For all experiments harvested DCs were washed and equilibrated in serum-free X-Vivo 15 media (Lonza).
Murine imBMs: Immortalized macrophages were immortalized by infecting bone marrow progenitors with oncogenic v-myc/vraf expressing J2 retrovirus as previously described (6) and differentiated in macrophage differentiated media containing MCSF. ImBM were maintained in 10% FCS PSN DMEM (Gibco). ImBM lines have been kindly provided by several collaborators and also obtained from the BEI resource: ICE (Casp1/Casp11), MAVs, IFN-R, IRF3-7 (Dr. K.Fitzgerald University of Massachusetts), STING and their rescues (Dr. R. Vance University of California Berkeley), Unc93b1 3d/3d (Dr. G. Barton University of Californina Berkeley), TLR 3, 4, 7, 9, 2-9, 2-4, MYD88, TRIF, TRAM, TRIF-TRAM (BEI resource ATCC/NIAID).

Investigation of Type I Interferon Pathway. To characterize whether this pathway could be modulated in our models, we evaluated production of type I interferon in response to stimulation by our i-ncRNA using human and murine interferon stimulated response element (ISRE) reporter cell lines, and monitored transcriptome regulation of a panel of immune genes related to the interferon pathway. Whereas the effect on the

inflammatory response is significant in terms of TNFalpha, IL-6, or IL-12 production, the effect on the type I interferon pathway was less prominent.

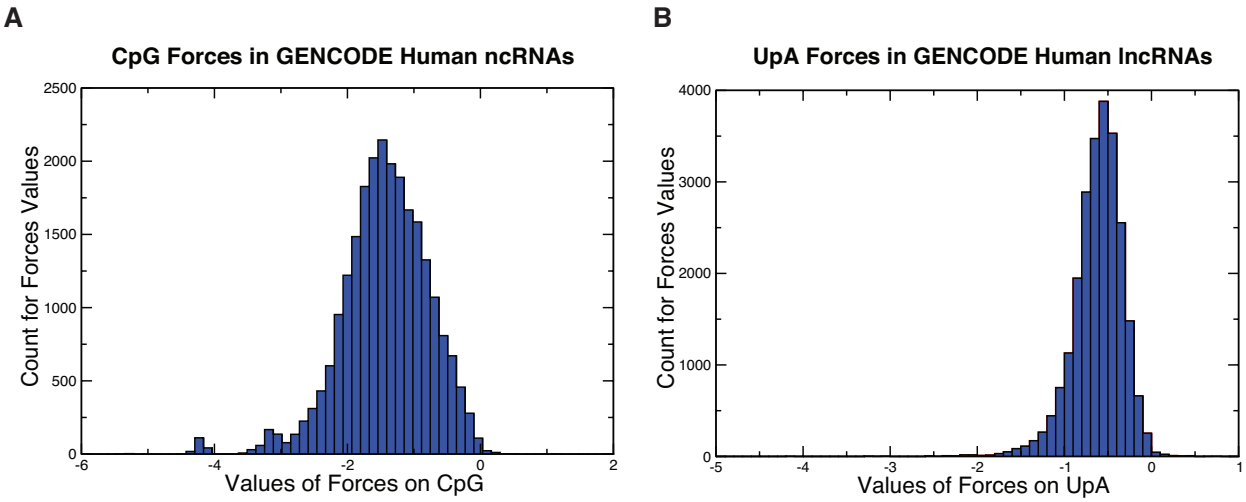
Additional Pathways Investigated. TLR2 or TLR4 were not required, indicating the observed effect was independent of contamination from bacterial products such as lipoproteins and endotoxins (Supplementary Figure 8). TRIF, TRIF/TRAM, and IRF3/IRF7, which participate down stream in the signaling of TLR3, TLR4, and TLR7, were also not obligatory (Supplementary Figure 5). We did not identify a role for candidate molecules for sensing murine GSAT, such sensors related to cGAS-STING signaling or DEAD box RNA helicases such as RIG-I and MDA5 (7-10). Inflammatory responses to GSAT did not depend upon the stimulator of interferon genes (STING), which induces type I interferon production when cells are infected with intracellular pathogens. RIG-I (retinoic acid-inducible gene 1) is a dsRNA helicase enzyme that senses RNA viruses through activation of the mitochondrial antiviral-signaling protein (MAVS) (11-13). MAVS deficient imBMs failed to respond to GSAT stimulation ruling out a contribution of RIG-I in our i-ncRNA signaling (Supplementary Figure 7B). Finally we ruled out a role for inflammasome related pathways using ICE-KO imBM that are essentially a knockout for Caspase 1 and which carry an inactive mutation for Caspase 11.

SUPPLEMENTARY REFERENCES

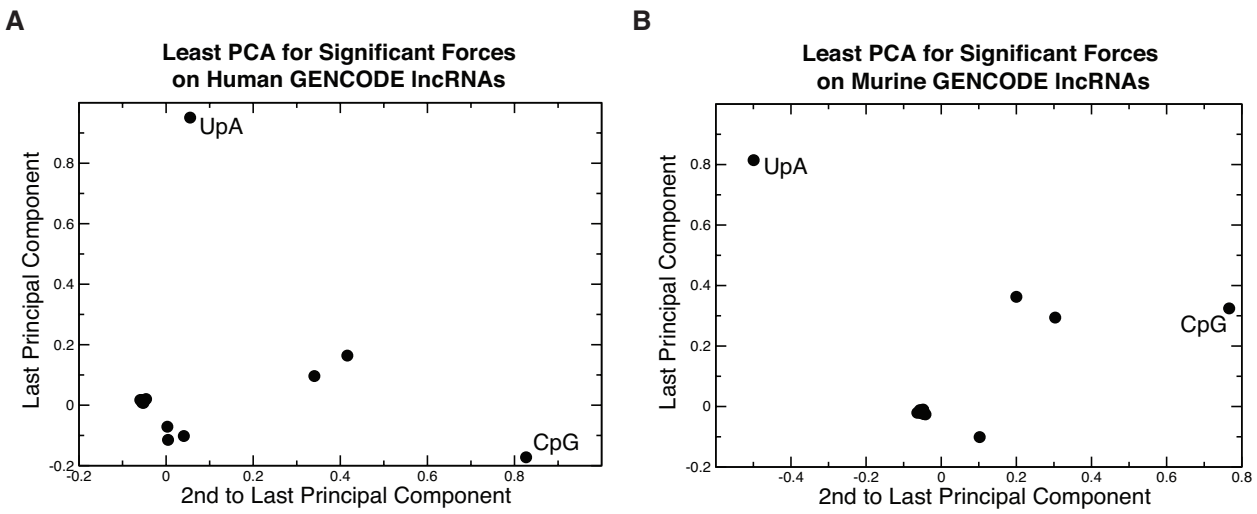
1. Matthews D, Sabina J, Zuker M, Turner D (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J Mol Biol* 288:911–940.
2. Wuchty S, Fontana W, Hofacker I, Schuster P (1999) Complete suboptimal folding of RNA and the stability of secondary structures. *Biopolymers* 49:145-165.
3. Harvey M, McArthur MJ, Montgomery CA, Bradley A, Donehower LA (1993) Genetic background alters the spectrum of tumors that develop in p53-deficient mice. *The FASEB Journal* 7:938-943.
4. Muller AJ, Teresky AK, Levine AJ (2000) A male germ cell tumor-susceptibility determining locus pgct1 identified on murine chromosome 13. *Proc Natl Acad Sci* 97:8421-8426.
5. Frleta D, et al. (2012) HIV-1 infection–induced apoptotic microparticles inhibit human DCs via CD44. *J Clinical Inves* 122:4685.
6. Blasi E, et al. (1985) Selective immortalization of murine macrophages from fresh bone marrow by a raf/myc recombinant murine retrovirus. *Nature* 318:667-670.
7. Atianand MK, Fitzgerald KA (2013) Molecular basis of DNA recognition in the immune system. *J Immunol* 190:1911-1918.
8. Lee BL, et al. (2013) UNC93B1 mediates differential trafficking of endosomal TLRs. *eLife* 2:e00291.
9. Burdette DL, Vance RE (2013) STING and the innate immune response to nucleic acids in the cytosol. *Nature Immunol* 14:19-26.
10. Vanaja SK, Rathinam VA, Fitzgerald KA (2015) Mechanisms of inflammasome

activation: recent advances and novel insights. *Trends Cell Biol*, in press.
11. Zeng M, et al. (2014) MAVS cGAS and endogenous retroviruses in T-independent B cell responses. *Science* 346:1486-1492.
12. Broz P, Monack DM (2013) Newly described pattern recognition receptors team up against intracellular pathogens. *Nature Rev Immunol* 13:551-565.
13. Gajewski TF, Schreiber H, Fu YX (2013) Innate and adaptive immune cells in the tumor microenvironment. *Nature Immunol* 14:1014-1022.

SUPPLEMENTARY FIGURES

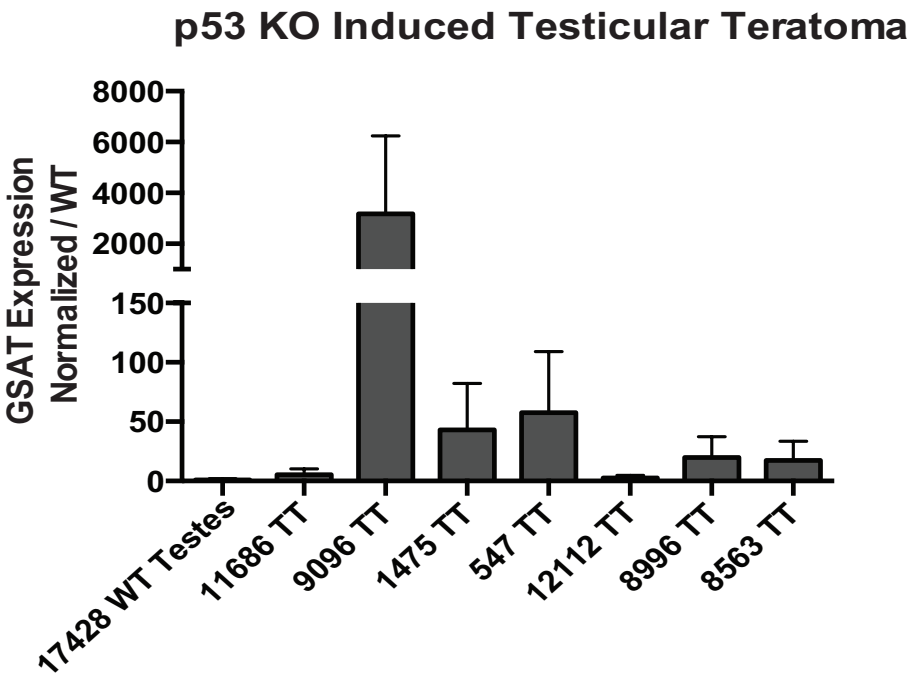


Supplementary Figure 1. CpG and UpA Are Generally Under-represented in ncRNA. Histogram of forces (strength of statistical bias) on (A) CpG and (B) UpA for lncRNA from the GENCODE Human transcript database. These forces are consistent with those observed in mice and those from coding regions.

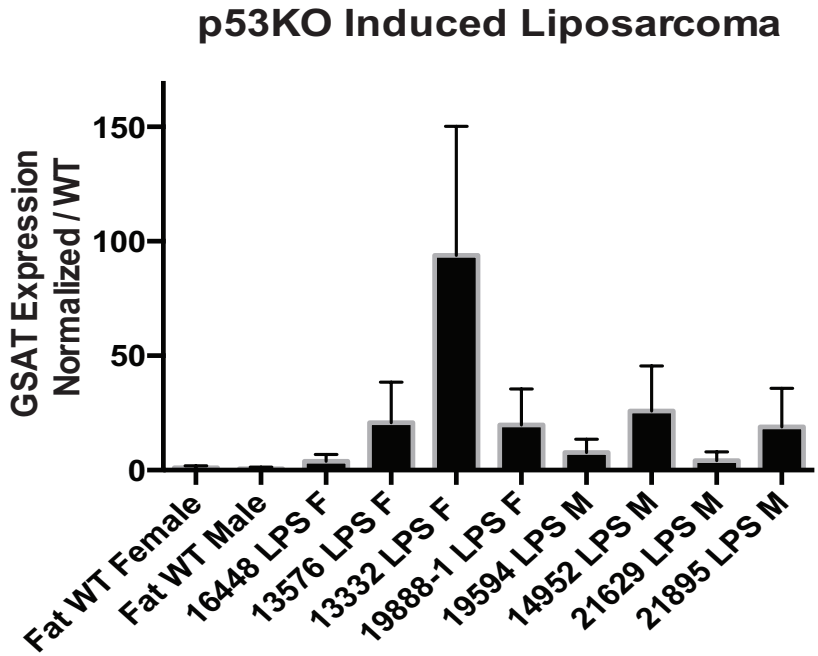


Supplementary Figure 2. Forces on CpG and UpA Dinucleotides Are Independent. Least principal components for all significant forces on motifs for (A) human and (B) mouse GENCODE ncRNA. In both cases CpG and UpA dominantly project onto the two least axes of variation.

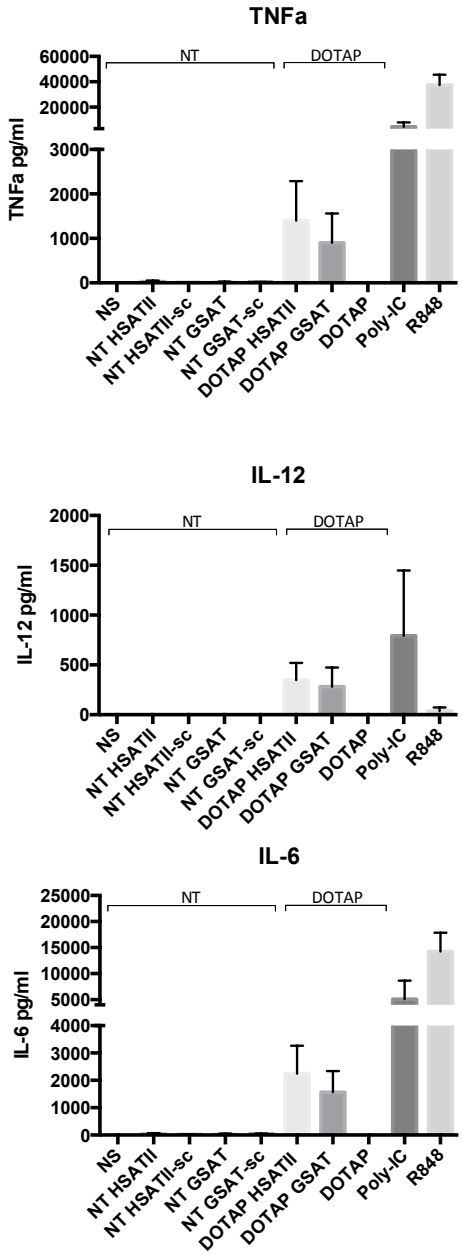
A



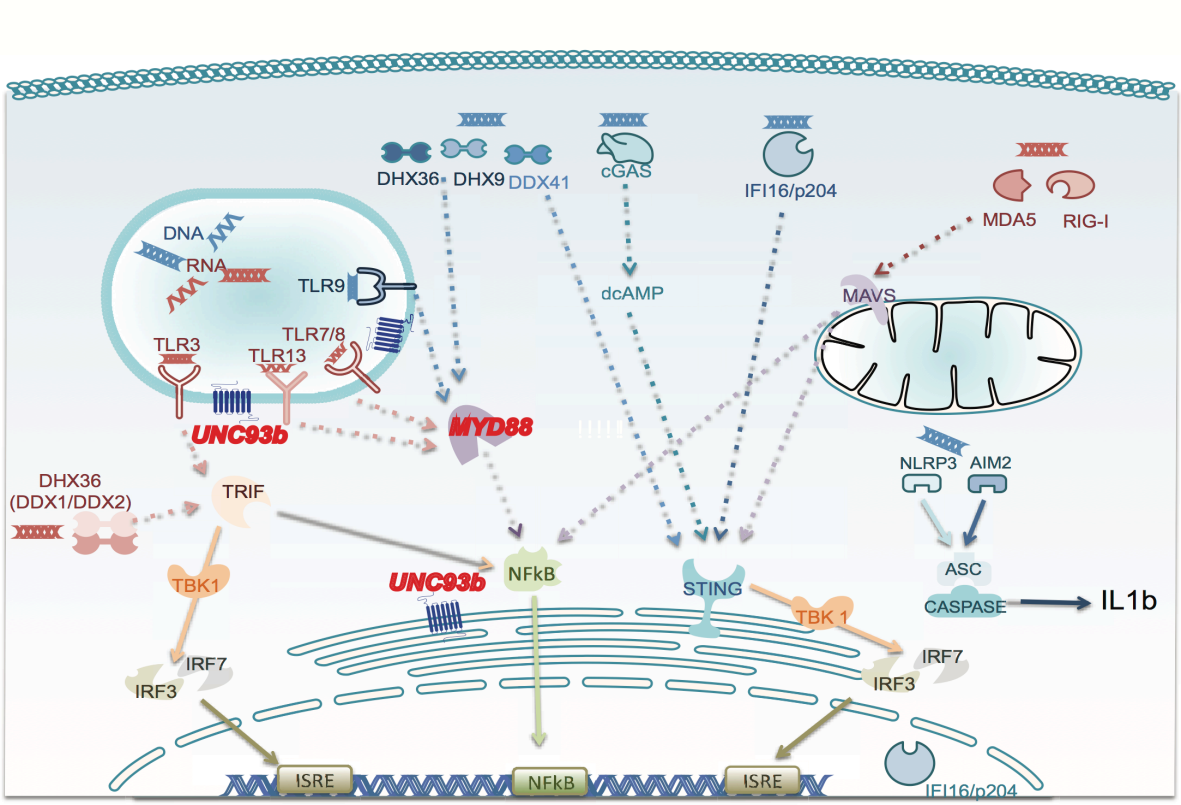
B



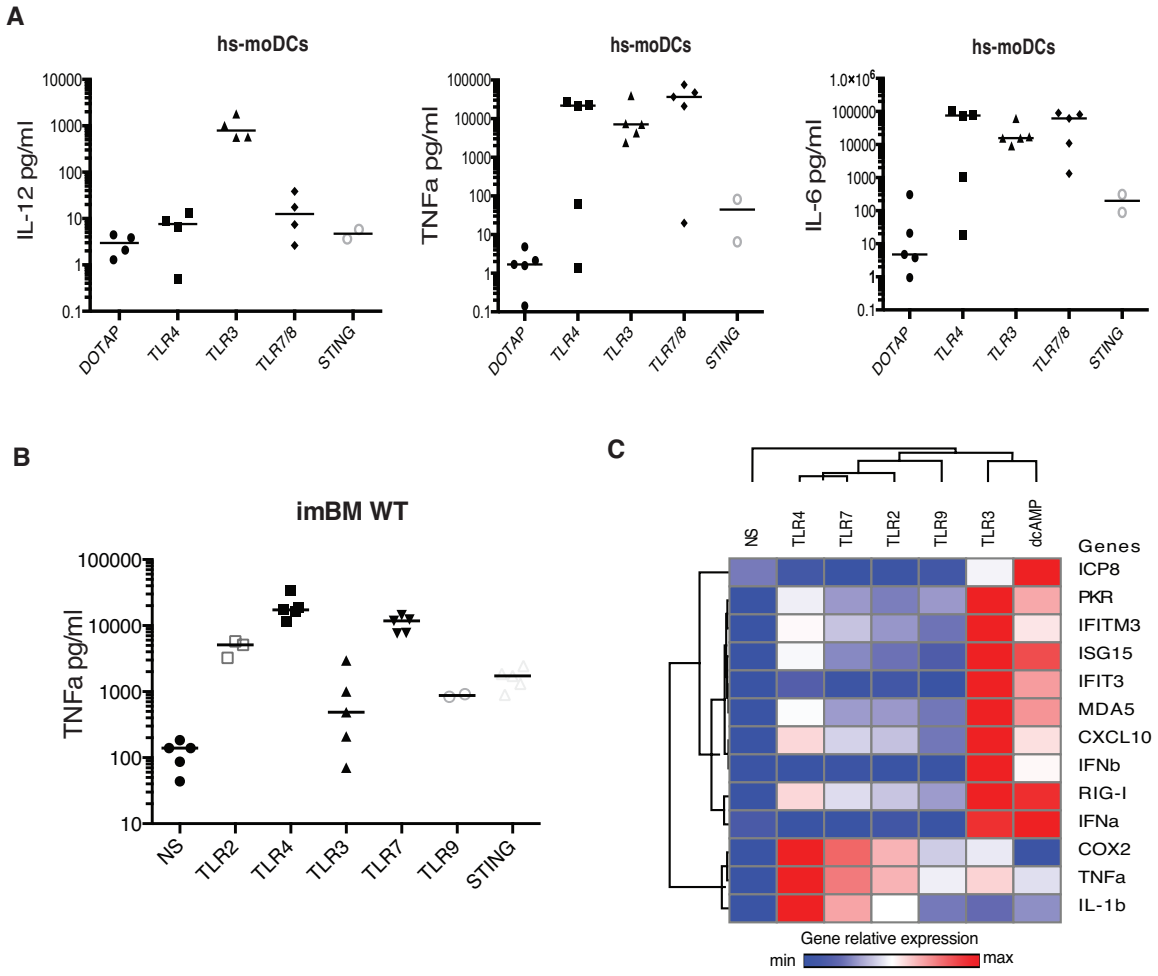
Supplementary Figure 3. GSAT is Expressed in Mouse Testicular Teratoma and Liposarcoma. Study of the relative levels of expression of GSAT RNA by a custom Taqman Assay in normal murine tissue versus murine tumor tissue samples. The tumor mouse models investigated were testicular (A) teratoma and (B) liposarcoma induced tumor in p53KO background. In all instances, GSAT levels were increased in the tumor samples as compared to normal samples, to varying degrees.



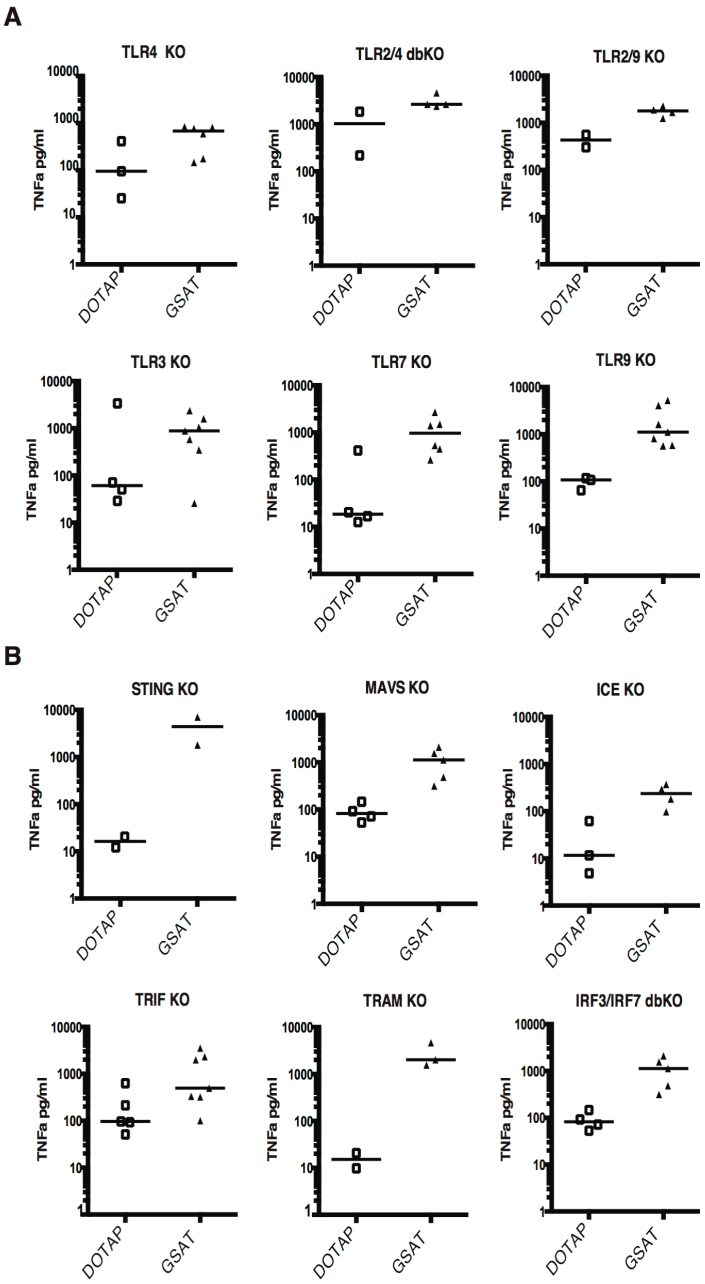
Supplementary Figure 4. NcRNA Require Transfection to Induce Cellular Innate Immune Responses. 2ug /ml of the various ncRNA (HSAT II, HSAT II-sc; GSAT; GSAT-sc) were used to stimulate human DCs in 96 well plates with (DOTAP) or without (NT) the use of DOTAP as a gentle liposomal transfection reagent. In absence of transfection reagent the ncRNA were not sensed by the DCs whereas transfected immunogenic ncRNA HSAT II and GSAT, in addition to Poly-IC and R848, were properly sensed and induced a cellular inflammatory response in (A) TNFalpha, (B) IL-12, and (C) IL-6.



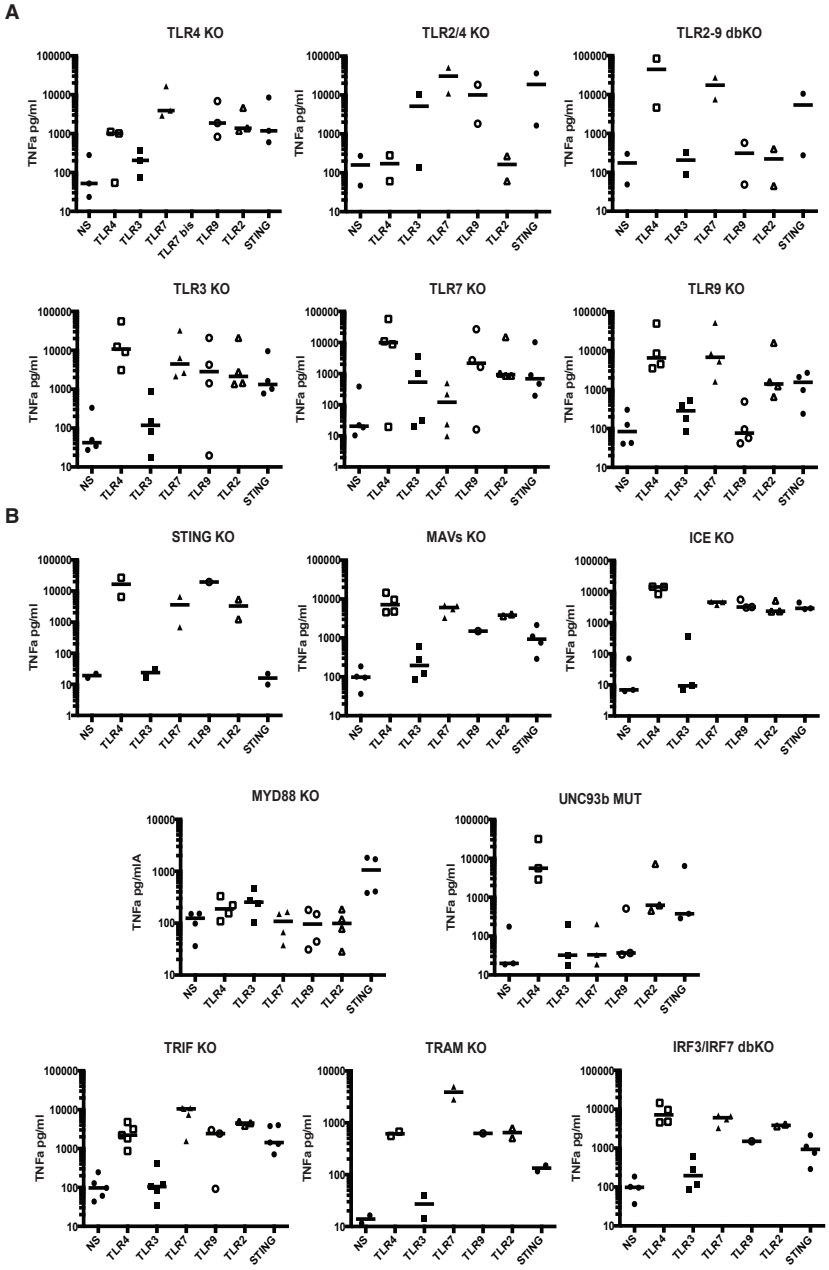
Supplementary Figure 5. Innate Immune Sensing of Nucleic Acids. Summary of the innate immune pathways involved in the sensing of nucleic acids which were investigated in this work. MYD88 and UNC93b, highlighted in red, were directly implicated in i-ncRNA sensing.



Supplementary Figure 6. Human moDCs and Mouse imBM Cells Respond to Common PAMPs and DAMPs. Quantification of inflammatory cytokine production in human moDCs (A) and in murine imBM (B) upon stimulation with common PAMPs or DAMPs known to activate PRR innate immune pathways, which are listed in the Materials and Methods. Each point represents the mean value of the experimental replicates for each individual condition; the bar represents the median. (C) The inflammatory response related to type I IFN pathway induction in imBM upon stimulation of the PRR related innate immune pathways has been analyzed by qRT-PCR. The heat-map represents the log of the relative expression of each gene based on relative quantification analysis using the ddCT bi-dimensional normalization method (house keeping genes and non-stimulated cells).



Supplementary Figure 7. Genetic Screen of Innate Immune Pathways Related to i-ncRNA Function in Murine imBM. (A) imBM cells of different knockout genotypes related to TLR PRRs (TLR2-4 dbKO, TLR3 KO, TLR4 KO, TLR7 KO, TLR9 KO). (B) imBM cells of different knockout genotypes related to STING, inflammasome, and MAV dependent helicases pathways (STING KO, MAV KO, ICE KO); and common innate immune signaling (TRIF KO, TRAM KO, IRF3/IRF7 dbKO). Cells have been stimulated by liposomal transfection of the murine i-ncRNA (GSAT). The TNFa production in the supernatant has been quantified and each point represents the mean value of the experimental replicates for each individual condition; the bar represents the median.



Supplementary Figure 8. Stimulation of KO and Mutant imBM with Common PAMPs and DAMPs. Quantification of inflammatory cytokine production in PRR KO imBM (A) and innate immune signaling related KO and mutant (B) upon stimulation with common PAMPs or DAMPs known to activate PRR innate immune pathways. Each point represents the mean value of the experimental replicates for each individual condition; the bar represents the median.

SUPPLEMENTARY TABLES

	Human	Mouse
CG	-1.419	-1.375
UA	-0.604	-0.548
ACG	-1.7586	-1.6216
CAG	0.5534	0.5612
CCG	-1.5095	-1.3287
CGA	-1.8995	-1.7082
CGC	-1.7304	-1.5525
CGG	-1.511	-1.2629
CGU	-1.7833	-1.6463
CUG	0.669	0.6748
GCG	-1.748	-1.5592
GUA	-0.8632	-0.7451
UAC	-0.7368	-0.6298
UAG	-0.733	-0.592
UCG	-1.9391	-1.7049

Supplementary Table 1. Average Forces on Motifs are Similar Between Humans and Mice. Average force on a given motif in the Human and Mouse GENCODE dataset, for lncRNAs with length greater than 500 nucleotides. The forces are listed for the significant motifs in humans. The force is a measure of the strength of statistical bias to enhance or suppress a motif versus what is expected from that sequences nucleotide content.

ncRNA	Class	Level of Conservation	CpG Force
MER123	DNA_transposon	Amniota	1.1039
HSATII	SAT	Primates	1.036
UCON21	Transposable_Element	Amniota	0.9465
MER6B	Mariner/Tc1	Homo_spaiens	0.923
Eulor1	Transposable_Element	Amniota	0.8481
Eulor5B	Transposable_Element	Tetrapoda	0.8474
Eulor2C	Transposable_Element	Amniota	0.7676
Eulor6A	Transposable_Element	Tetrapoda	0.7466
MER131	SINE	Amniota	0.6223
Eulor4	Transposable_Element	Tetrapoda	0.6067
Eulor10	Transposable_Element	Amniota	0.6064
MER6C	Mariner/Tc1	Eutheria	0.5667
Eulor12	Transposable_Element	Amniota	0.5295
MER5C1	hAT	Eutheria	0.4582
MER47B	Mariner/Tc1	Eutheria	0.4518
UCON39	DNA_transposon	Mammalia	0.4443
UCON16	Transposable_Element	Amniota	0.4436
Tigger3d	Mariner/Tc1	Primates	0.4374
TIGGER5A	Mariner/Tc1	Eutheria	0.4212
MER75	DNA_transposon	Homo_spaiens	0.4134
Tigger4a	Mariner/Tc1	Primates	0.3815
npiggy2_Mm	piggyBac	Microcebus_murinus	0.3725
MER58B	hAT	Eutheria	0.3657
Eulor6C	Transposable_Element	Tetrapoda	0.3571
Eulor11	Transposable_Element	Amniota	0.3561
UCON15	Transposable_Element	Amniota	0.356
Tigger2b_Pri	Mariner/Tc1	Primates	0.3548
MER44B	Mariner/Tc1	Homo_spaiens	0.3536
SUBTEL_sat	Satellite	Primates	0.3527
Eulor9A	Transposable_Element	Amniota	0.3465
MER44C	Mariner/Tc1	Homo_spaiens	0.3439
Eulor8	Transposable_Element	Amniota	0.3416
MER44D	Mariner/Tc1	Eutheria	0.3211
npiggy1_Mm	piggyback	Microcebus_murinus	0.3131
UCON26	Transposable_Element	Amniota	0.2985
MER127	Mariner/Tc1	Amniota	0.2984
MER97d	hAT	Eutheria	0.2939
Eulor6D	Transposable_Element	Tetrapoda	0.2866
Eulor2B	Transposable_Element	Amniota	0.2852
MER119	hAT	Homo_spaiens	0.2794

MER134	Transposable_Element	Amniota	0.2786
Eulor9C	Transposable_Element	Amniota	0.2751
MER8	Mariner/Tc1	Homo_spaiens	0.2669
Ricksha_a	MuDR	Eutheria	0.2607
MER129	SINE	Amniota	0.2444
MacERV6_LTR3	ERV3	Cercopithecidae	0.2404
MER57B2	ERV1	Homo_spaiens	0.2403
HSMAR1	Mariner/Tc1	Homo_spaiens	0.2397
Eulor12_CM	Transposable_Element	Amniota	0.2269
MERX	Mariner/Tc1	Eutheria	0.2207
Tigger12A	Mariner/Tc1	Mammalia	0.217
MER58A	hAT	Eutheria	0.2006

Supplementary Table 2. Many Repetitive Elements Have High CpG Forces.
Listed above are the repetitive elements from Repbase with a significantly high CpG force. These elements are typically not found to be expressed in normal tissue, yet some may be expressed in cancer cells and cell lines.