



Supplementary Materials for

The Ro60 Autoantigen Binds Endogenous Retroelements and Regulates Inflammatory Gene Expression

Tiffany Hung^{1,*}, Gabriel Pratt², Balaji Sundararaman², Michael J. Townsend, Christina Chaivorapol¹, Tushar Bhangale¹, Robert R. Graham¹, Ward Ortmann¹, Lindsey A. Criswell³, Gene Yeo^{2,4,5,*}, Timothy W. Behrens^{1,*}

*correspondence to: behrens.tim@gene.com, hungt2@gene.com, and geneyeo@ucsd.edu

This PDF file includes:

Materials and Methods
Supplementary Text
Figs. S1 to S14

Other Supplementary Materials for this manuscript includes the following:

Additional Data table S1 (separate file)

Ro60 iCLIP peaks shared between GM12878 and K562

Additional Data table S2 (separate file)

Genes differentially expressed between wildtype and Ro60 -/- GM12878 cells

Materials and Methods

Cell lines

GM12878 (Coriell) and K562 (ATCC) cell lines were grown in RPMI-1640 supplemented with 15% FBS and 1X Glutamax (Invitrogen). GM12878 Ro60-null clones were generated using zinc finger nucleases targeting exon 2 of TROVE2 (Sigma; see Fig. S5). Recombinant interferon alpha (Sigma I4276) was applied to culture media at 1000 U/ml for 6 hours unless otherwise indicated.

Crosslinking and Immunoprecipitation (iCLIP)

iCLIP was performed as previously described(12). Briefly, 10 million cells were UV cross-linked at 400mJ/cm² and flash frozen at -80°C. Cells were lysed in buffer containing 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate and protease inhibitors (Roche). Lysate was treated with RNase I at units indicated in Fig.S1D and 4U DNase at 37C for 3min. Ro60 and associated RNA were immunoprecipitated at 4°C overnight using 10µg of indicated antibodies pre-coated in 100µl of Protein G beads. Precipitation using normal IgG antibody and non-cross-linked cell lysate were used as negative controls. Bead-precipitated RNP complexes were washed three times with buffer containing 1M NaCl to remove non-specifically associated protein/RNA. A 3' linker was ligated on-bead, and RNPs were radiolabelled using T4 PNK, resolved on SDS-PAGE and transferred to nitrocellulose membranes. RNA associated with the protein was confirmed by autoradiogram; immunoprecipitation was confirmed by Western blot. RNP complexes from ~60-125 kDa were cut as one band, membrane pieces were treated with Proteinase-K, and RNA was isolated by urea/phenol:CHCl₃ extraction. RNA was reverse transcribed using an adapter complementary to the 3' linker. The RT oligo also contained barcodes for multiplexing and randomers to track PCR duplcates. cDNA was size selected, circularized using CircLigase II, annealed with an oligo containing a BamHI restriction site, then re-linearized by digesting with BamHI. The final libraries were amplified using Solexa P3 and P5 primers and sequenced on an Illumina HiSeq 2500; basecalls were performed with CASAVA v.1.8.

Reads were mapped to human genome build hg19 using STAR (<https://code.google.com/p/rna-star/>) with the “outFilterMultimapNmax 10” option, then PCR duplicates were removed using unique randomers in the barcode sequence. Peak calling was performed using pyicoclip (<http://regulatorygenomics.upf.edu/Software/Pyicoteo/pyicoclip.html>) and the cutoff for further analysis was a minimum peakscore of 5.

MEME motif analysis

MEME was performed on shared peaks between GM12878 and K652 iCLIP using the peak sequence plus 10 flanking bases. The default settings were used: see documentation at <http://meme.nbcr.net/meme/doc/meme.html>. To reverse search the genome for the motif, the MEME motif file was used as input to FIMO (<http://meme.nbcr.net/meme/doc/fimo.html>) to search for all matches in the hg19 genome with a minimum $-\log_{10}p$ val of 6. These sites were then collapsed to Refseq annotated

genes. Alu motif density per 10Kb= (instances of FIMO-defined Ro60 Alu motifs)/(Total gene length)/(10Kb).

Antibodies

Ro52 western blot: A302-519A (Bethyl labs)
Ro52 Immunoprecipitation: D-12 (Santa Cruz)
Ro60 Immunoprecipitation: (Santa Cruz Biotech AA-3)
Ro60 western blot: A303-692A (Bethyl labs)
La western blot / Immunoprecipitation: Origene TA500406
IgG: SC-3877 (Santa Cruz Biotech)
Tubulin: ab-15246 (Abcam)

RNA Immunoprecipitation

10 million cells were UV crosslinked (400 mJ/cm²), lysed using cell lysis buffer (same buffers as iCLIP), and incubated with 5 µg of antibodies conjugated to protein-G Dynabeads (Invitrogen). Following overnight incubation, bead-bound complexes were washed 3 times in iCLIP high salt wash buffer, three times in iCLIP wash buffer, and either boiled for SDS-page for western blot or incubated in Proteinase K digestion buffer (100 mM NaCl, 1 mM Tris pH 7, 1 mM EDTA, 0.5% SDS, 5% by volume Proteinase K (Ambion AM2548)) for 45 minutes at 45°C. RNA was extracted using TRIZOL (Ambion) and purified using mirVANA (Ambion AM1560). RNA was DNase treated (Turbo DNA-free, Ambion) for 30 minutes prior to downstream analysis. **For SLE RIP**, 50 µl of patient sera was preincubated with Protein G beads for immunoprecipitation. **For RIP-seq**, RNA was converted to cDNA using Ovation RNA-seq system V2 (Nugen), Covaris fragmented, PCR amplified and barcoded (NEBnext, NEB) and sequenced on the Illumina Hi-seq platform (50 bp, single end). Reads were mapped to the human genome hg19 using bowtie2 and peak calling was performed using Input and IgG as negative controls using Homer (homer.salk.edu). Mapped reads in each condition were normalized to 10⁷ reads prior to comparison.

RT-PCR

RT-PCR was performed using the Taqman RNA-to-Ct One Step kit or SYBR Green RNA to Ct One-step kit (Applied Biosystems) using 50 ng of DNase treated (Turbo DNA-free, Ambion) RNA and the manufacturer's standard protocol. For RIP-PCR experiments, 1/50 of the total isolated RNA was used. Gene expression quantification was normalized to GAPDH using the ΔCt method. All error bars represent standard deviations of three technical replicates. Primers and probes are listed in Supplemental Text. For strand specific RT-PCR, one primer was used for RT reaction (48°C 30 minutes at, 95°C 10 minutes), and the second primer was added prior to PCR.

RNA Electromobility Shift Assay

Custom RNA oligos were P³² end-labeled using KinaseMax (Ambion) and incubated with recombinant Ro60 (Reagent Proteins) in REMSA binding buffer (10 mM HEPES 7.4, 20 mM KCl, 1 mM MgCl₂, 1 mM DTT) for 30 minutes at 30C. Complexes were electrophoresed on a 6% TBE acrylamide gel and imaged on a Typhoon PhosphoImager. Quantification of bands was performed using ImageQuant software.

SLE patient sera

SLE patient sera was obtained from UCSF Medical Center. Immune complexes were isolated by incubation of 50 μ l of sera with 25 μ l of protein G Dynabeads (Invitrogen) in 1 ml of iCLIP lysis buffer for 2 hours at room temperature. Immune complexes were then subjected to TRIZOL extraction or used for immunoprecipitation in cell lysates.

SLE RNA-seq patient selection and clinical analysis

SLE RNA-seq blood samples were collected and processed as described(21). See Fig. S12 for patient demographics. ISM scores were calculated as described(21). Briefly, the mean Ct (as measured by RT-PCR) for seven interferon stimulated genes were combined for the score. Autoantibodies to Ro were measured in serum samples using Quantaplex SLE Profile 8 kits (Inova Diagnostics), which does not distinguish between anti-Ro52 and anti-Ro60.

Alu quantification by RNA-seq

Raw reads were mapped to all Repeatmasker Alu elements +/- 15bp using bwa aln (<http://bio-bwa.sourceforge.net/>). %Alu reads = (# of reads mapping to Alu elements)/total reads). The top 500 expressed Alu RNAs in each patient blood sample (controls and SLE) were collapsed for a total of 1,641 unique Alus for further analysis. Unsupervised hierarchical clustering based on average linkage criterion was used to generate heatmaps of Alu expression. Alu score= average expression within the cluster containing the most highly differentially expressed Alu RNAs in control vs. SLE, n=128.

Isolation and treatment of PBMCs

PBMCs from healthy donors were isolated from heparinized peripheral venous blood using LSM-Lymphocyte separation medium (MP Bio) and centrifugation through a Leucosep column (Thomas scientific). PBMCs were lysed with ACK cell lysis buffer (Life Technologies), washed in PBS, and resuspended in complete RPMI media (15% FBS, 1% glutamine/penicillin/streptomycin). PBMCs were either used for transfection experiments or treated with IFN α +/- 6 hours for RNA-seq analysis.

RNA-seq of GM12878 cells, PBMCs, and SLE patient blood

For GM12878 and PBMCs, RNA was extracted using TRIZOL (Ambion) and purified (RNeasy Qiagen). For patients blood, RNA was extracted from baseline whole blood (Paxgene). Sequencing libraries were constructed using 1 μ g of total RNA using the TruSeq kit (Illumina) and sequenced using Illumina HiSeq with the single end protocol at 50 bp length. Reads were mapped to human genome version hg19 using GSNAP (<http://research-pub.gene.com/gmap/>). Differential expression calculations were performed using Deseq2, and differentially expressed genesets represent a minimum of 1.5 fold change (up or downregulated) with a p value <.05.

PBMC Transfection and Cytokine ELISA

PBMCs were isolated from healthy donors and plated in 96 well plates at 0.5*10⁶ cells per well. 25 ng of RNA was transfected using 0.5 μ l of DOTAP (Roche) in HBS buffer according to the manufacturer's protocol. Chloroquine (Invivogen) was added to the

culture media at a pre-determined optimal concentration of 30 μ M. BX-795 (Invivogen) was added to culture media at 20 and 50 nM. After 16-18 hours, cells supernatants were collected, and 5 μ l of PBMC supernatants were used for ELISA following the manufacturer's protocols. The following assays were performed: TNF α (eBioscience 88-7346), IL-6 (eBioscience 88-7066), and Multi-Subtype IFN α (PBL assay science 41105). For experiments using IRS954, custom inhibitory phosphorothioate oligos sequences(18,19) were synthesized (IDT) and co-transfected with Alu RNAs using DOTAP (Roche) at 1 μ M final concentration.

TLR7/9 inhibitory phosphorothioate oligo. *denotes phosphorothioate bonds.

IRS 954

5'-T*G*C*T*C*C*T*G*G*A*G*G*G*G*T*T*G*T-3'

IRS Control

5'-T*C*C*T*G*C*A*G*G*T*T*A*A*G*T-3'

Taqman probes

GAPDH	Hs02758991_g1
TROVE2	Hs01567394_m1
TLR5	Hs01920773_s1
LECT1	Hs00993254_m1
IL1R1	Hs00991002_m1
IL10	Hs00961622_m1
PKIB	Hs00261162_m1
DMRT2	Hs00246364_m1
GPC4	Hs00155059_m1
TNFRSF19	Hs00218634_m1
SMAD1	Hs01077084_m1
BCL11A	Hs01093196_m1
TNFRSF19	Hs00969483_m1

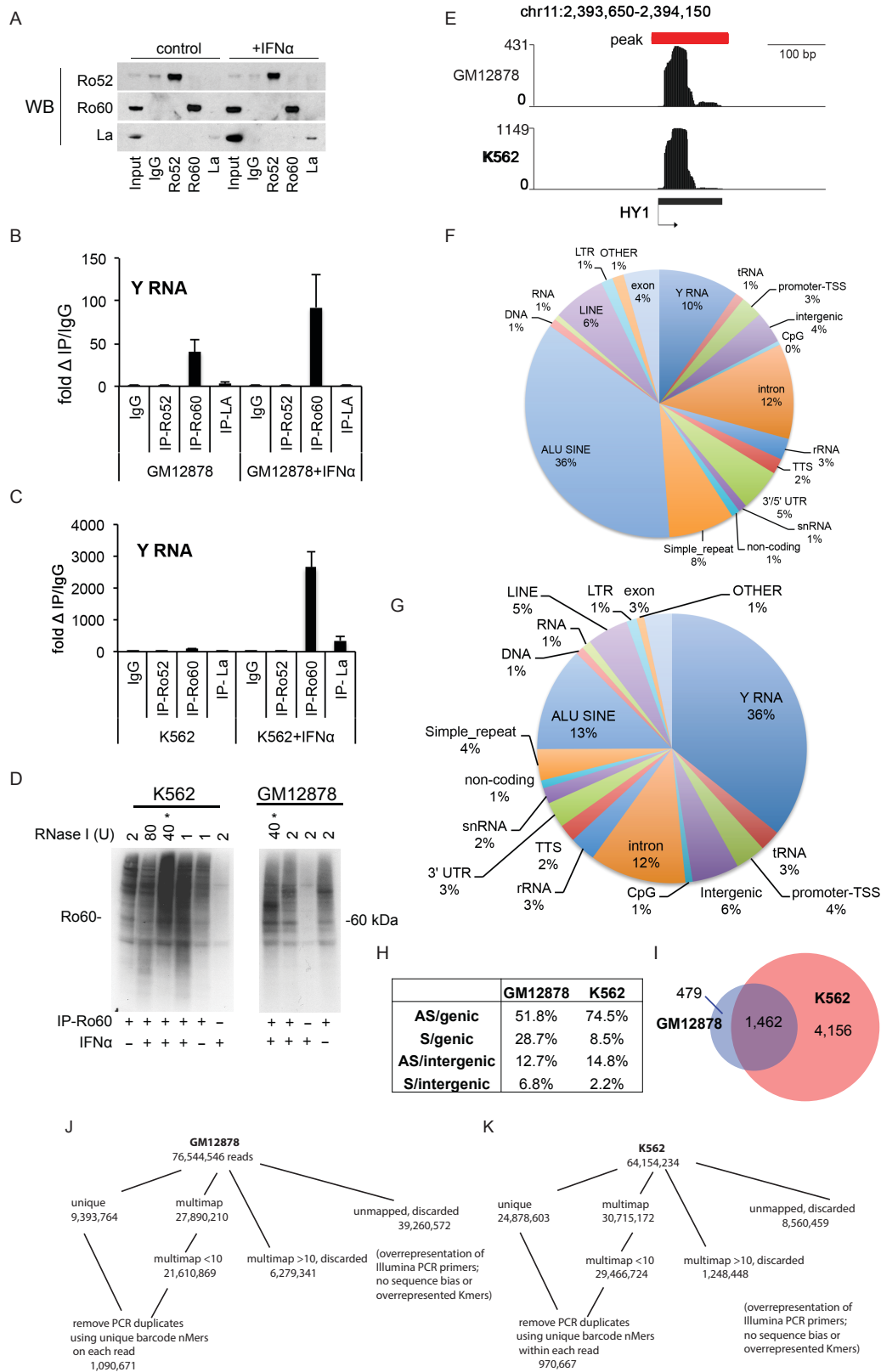


Fig. S1. Ro60 iCLIP

(A) Western blot of Ro52, Ro60 and La immunoprecipitates from resting or IFN α (6 hr) stimulated GM12878 cells. IgG is a control mouse monoclonal antibody. Input lane is 10% of the cell lysate. (B) RNA was isolated from Ro60 immunoprecipitates in (A), followed by RT-PCR for HY4 RNA. Data are presented as fold enrichment over the IgG control. Error bars represent s.d. (C) Same as (B), but with K562 cell line. (D) Autoradiography of Ro60 iCLIP RNAs in K562 and GM12878 cell lines. RNase I treatment as indicated above lanes. RNAs from ~60-125 kDa were extracted from the 40U RNase I treated lanes (*) for sequencing. (E) Genome browser snapshot of iCLIP reads from a representative HY1 locus. Axis indicates read counts. (F,G) Annotation of Ro60 iCLIP tags to repeatmasker elements: (F) K562 (n=8273) (G) GM12878 (n=2440) (H) Breakdown of all Alu tags as Antisense (AS) or Sense (S) in relation to annotated Alu, as well as genic or intergenic. The majority of peaks were Antisense/Genic in both cell lines. (I) Overlap between genes with annotated iCLIP peaks in K562 vs. GM12878. $p < .001$, hypergeometric distribution. (J,K) Ro60 iCLIP mapping summary in (J) GM12878 or (K) K562 cell lines.

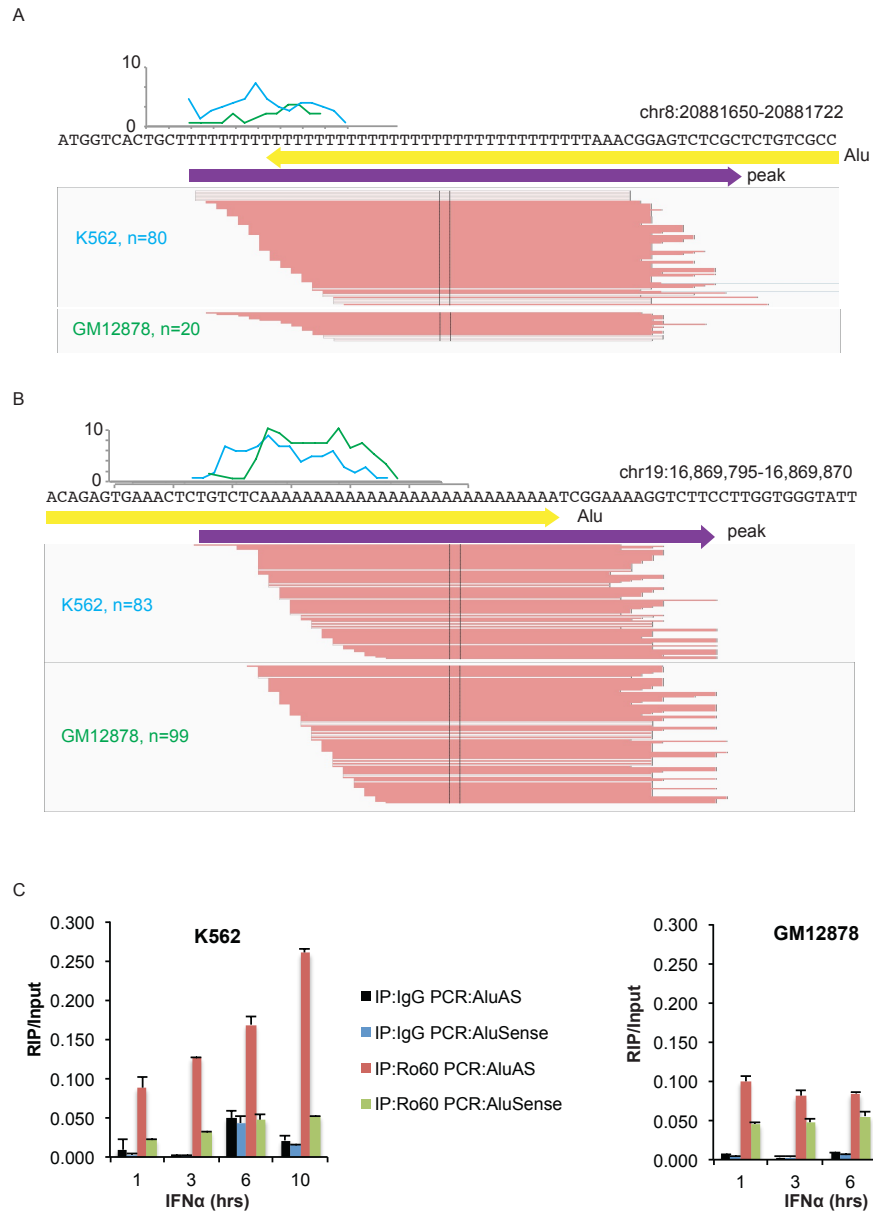


Fig. S2. Enrichment of Antisense Alu in Ro60 bound Alu RNAs. (A, B) Alu Ro60 iCLIP crosslinking sites on selected Alu peaks. (A) Alu RNA antisense to Alu element. (B) Alu RNA sense to Alu element. Purple arrow represents iCLIP peak, yellow arrow represents annotated Alu element, graph represents quantification of crosslinking sites at indicated position (first nucleotide in the genome upstream of a mapping cDNA defined as ‘cross-link nucleotide’ (12)). Blue, K562 crosslinking sites; Green, GM12878 crosslinking sites. Red bars represent plus-strand iCLIP reads (n= read count). (C) Strand specific RT-PCR in Ro60 immunoprecipitates of K562 and GM12878 cells at indicated hours post IFN α treatment. Alu4 primers were used for RT-PCR. AS=Antisense, see(14). Error bars represent standard deviation of technical replicates; representative experiment (of 3).

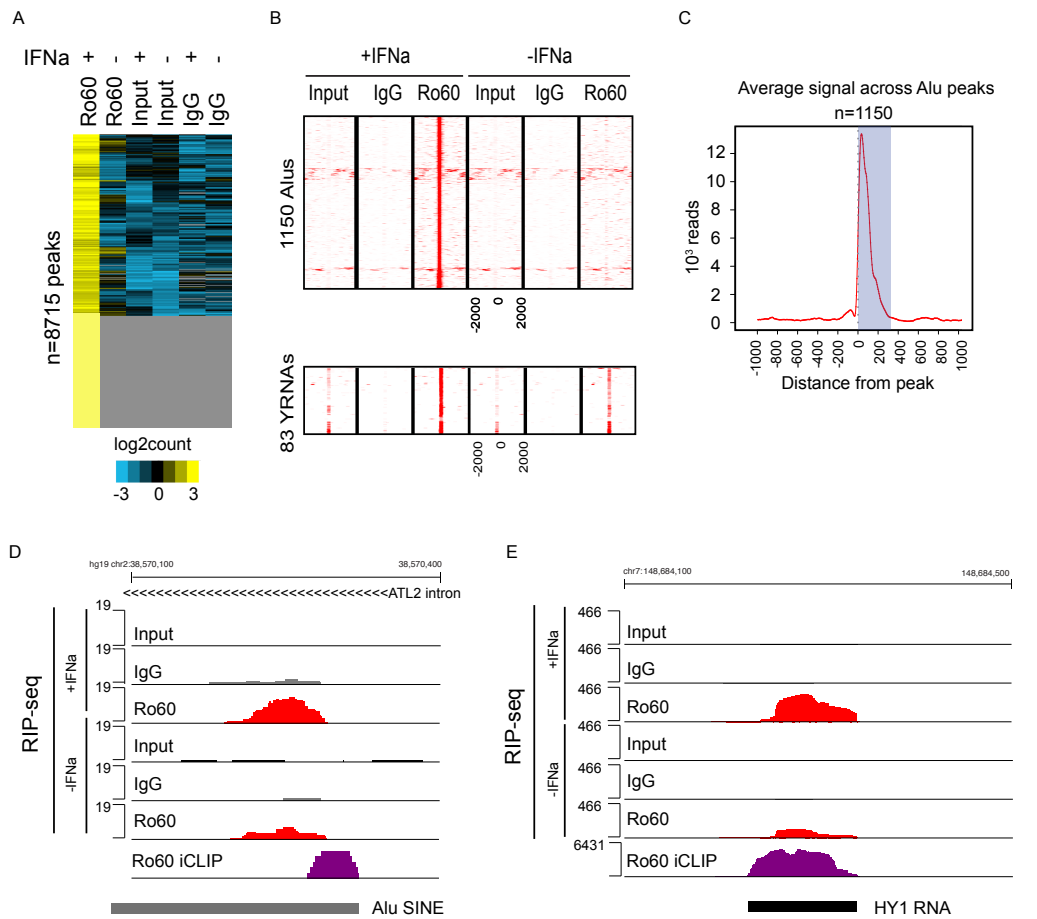


Fig. S3. Ro60 RIP-seq in K562 cells identifies Alu RNAs

(A) Hierarchical clustered heatmap of Ro60 RIP-seq and corresponding signal in Input and IgG control. IPs were conducted +/- IFN α (6 hours). Grey box= Ro60 RIP RNAs with no corresponding signal in other conditions.

(B) Raw RIP-seq signal 2000 bp adjacent to the center of indicated Alu site (n=1150) or Y RNA site (n=83). The Ro60 Alu signal was stronger in the presence of IFN α , and Y RNA signal was detectable +/- IFN α .

(C) Metagene plot of Ro60 RIP-seq reads across all Alu sites reveals no substantial signal beyond the average size of the Alu element.

(D, E) Genome browser view of representative RIP-seq sites alongside iCLIP tags. (D) Alu RNA, (E) Y RNA. Libraries were normalized to 10⁷ reads prior to analysis.

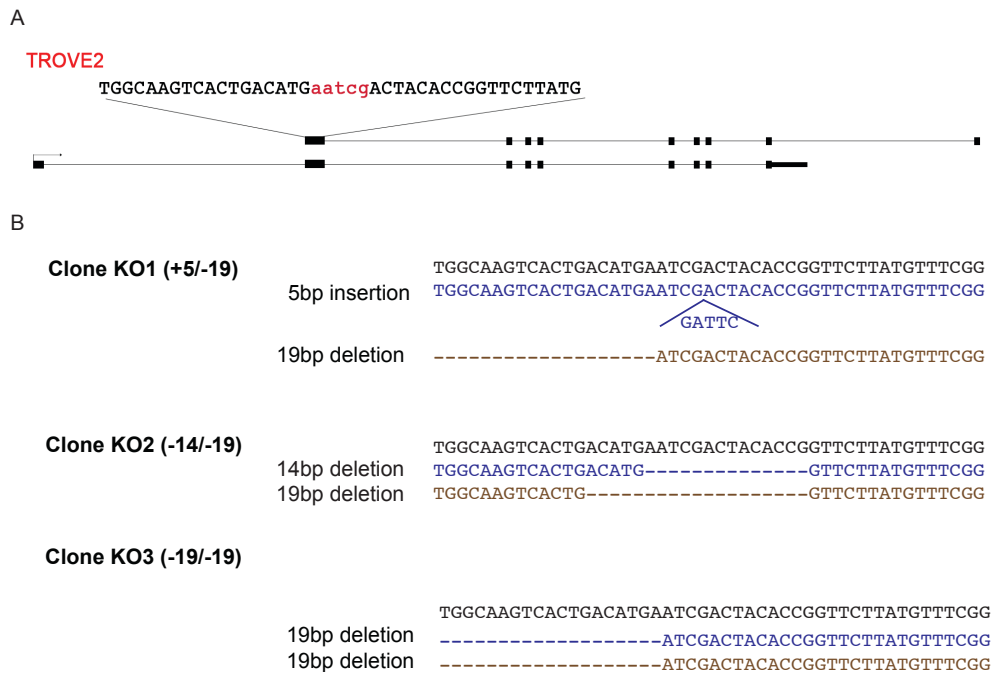


Fig. S5. GM12878 Ro60 null clones.

(A) Zinc finger nucleases targeting exon 2 of TROVE2, the gene encoding Ro60. Black= binding site, red= cutting site

(B) Mutations in each of the null clones. Black= wildtype, blue=allele1, brown=allele2.

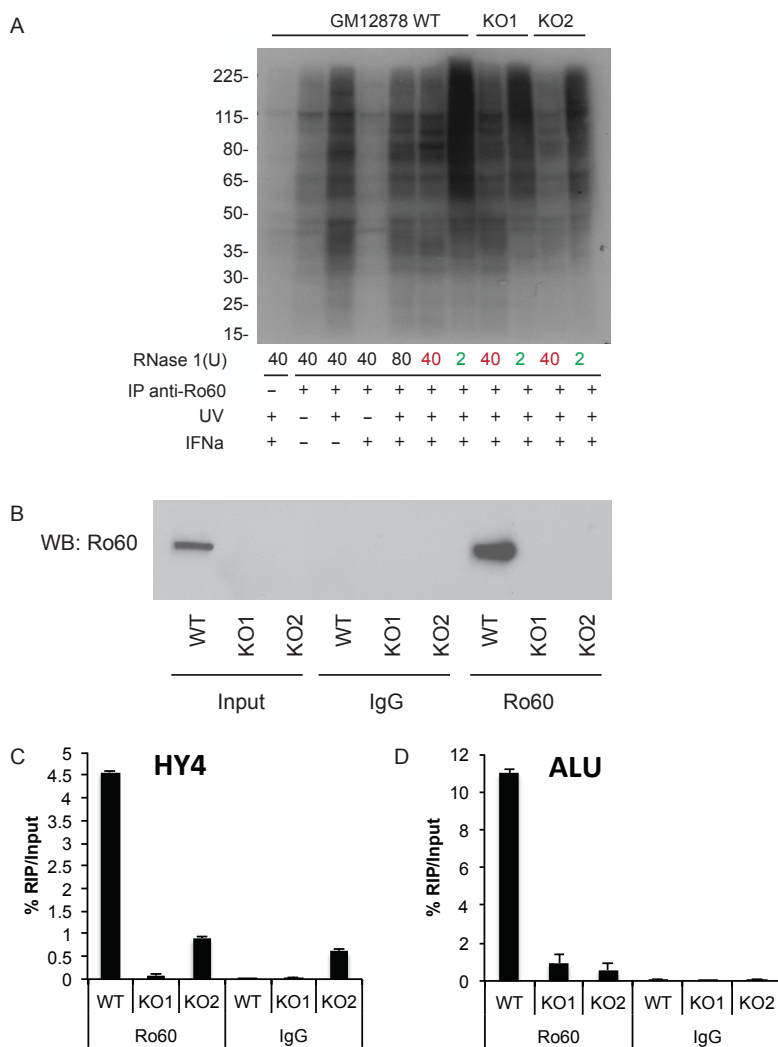


Fig. S6. Ro60 iCLIP in Ro60 knockout cell lines.

(A) Autoradiogram of Ro60 iCLIP in Ro60 wildtype (WT) and null (KO1, KO2) cell lines. UV crosslinking, antibody, and IFN α treatment (6 hrs) indicated below each lane.

(B) Ro60 western blot of Ro60 or IgG immunoprecipitates in wildtype (WT) and Ro60 KO cell lines (KO1, KO2).

(C, D) RT-PCR quantification of indicated RNA from immunoprecipitates from (B).

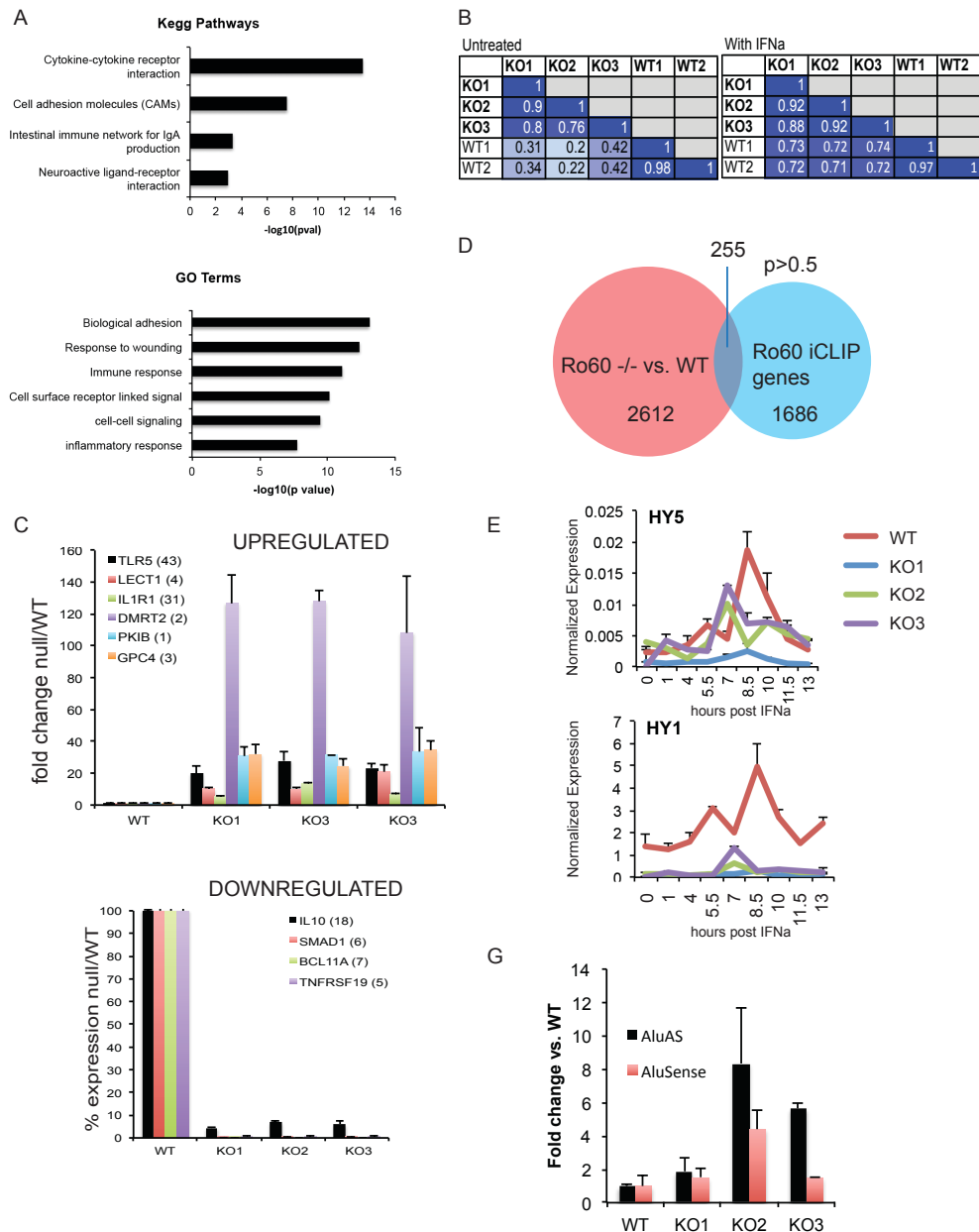


Fig. S7. Characterization of Ro60 knockout cell lines.

(A) KEGG pathways and GO terms enriched in genes differentially expressed between Ro60 KO relative to wildtype GM12878 cells.

(B) Pearson correlation R values between RNA-seq libraries of the indicated GM12878 wildtype or KO clones +/- 6 hours IFN α . Each comparison represents RPKMs of all Refseq annotated genes, n=26,469. There is heterogeneity amongst the Ro60 KO clones

but the global gene expression correlation is similar; these differences are lessened upon IFN α treatment.

(C) RT-PCR confirmation of gene expression changes in Ro60 KO clones. Upregulated genes plotted as fold change relative to wildtype. Downregulated genes plotted as percent of wildtype. Parenthesis indicates rank of gene within total list of RNA-seq upregulated or downregulated genes.

(D) Venn diagram of overlap between genes with Ro60 iCLIP tags and genes differentially expressed (>1.5 fold up or downregulated) upon Ro60 knockout.

(E) RT-PCR of Y RNAs across a time course of IFN α treatment in WT or Ro60 null clones.

(G) Strand specific PCR of Alu RNA (Alu4 primers) in WT or Ro60 null clones. Sense or Antisense (AS) indicates orientation to canonical Alu element.

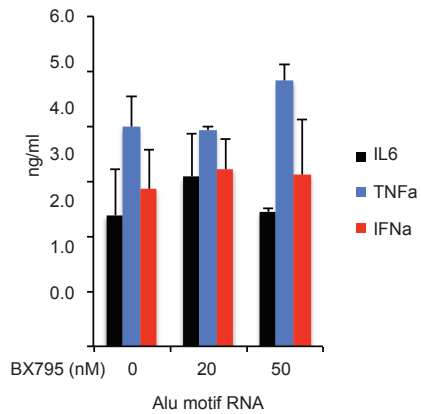


Fig. S8. TBK/IKKe inhibitor BX795 does not inhibit Alu motif mediated cytokine production. Treatment of PBMCs with BX795 at the indicated concentrations during Alu RNA transfection (15 hours) followed by ELISA of the culture supernatant. Higher concentrations of BX795 resulted in cell death.

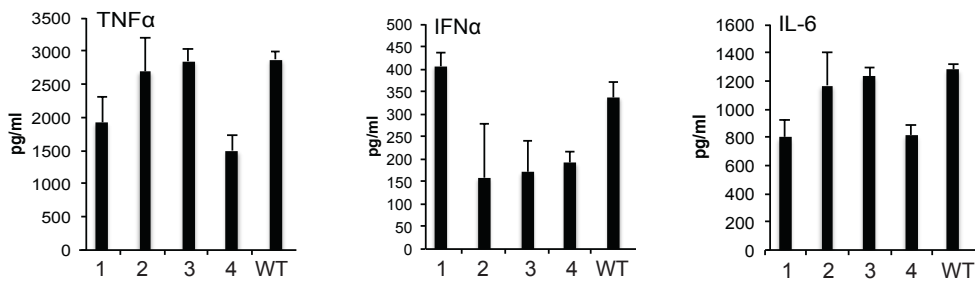


Fig. S9. ELISA of cytokines from culture supernatants of PBMCs transfected with RNA Alu motif or motif variants (see Fig. S4A for oligos used).

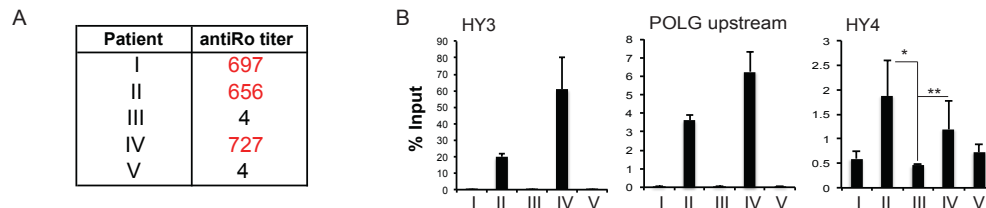


Fig. S10. SLE patient sera

(A) Titers of anti-Ro autoantibodies in the serum of five SLE subjects determined by clinical laboratory testing with a classical Ro ELISA containing a mix of Ro52 and Ro60 antigens.

(B) RT-PCR of RNAs precipitated by SLE IgGs (SLE-RIP). Error bars represent s.d. of 3 technical replicates. * $p=0.03$, ** $p=0.08$.

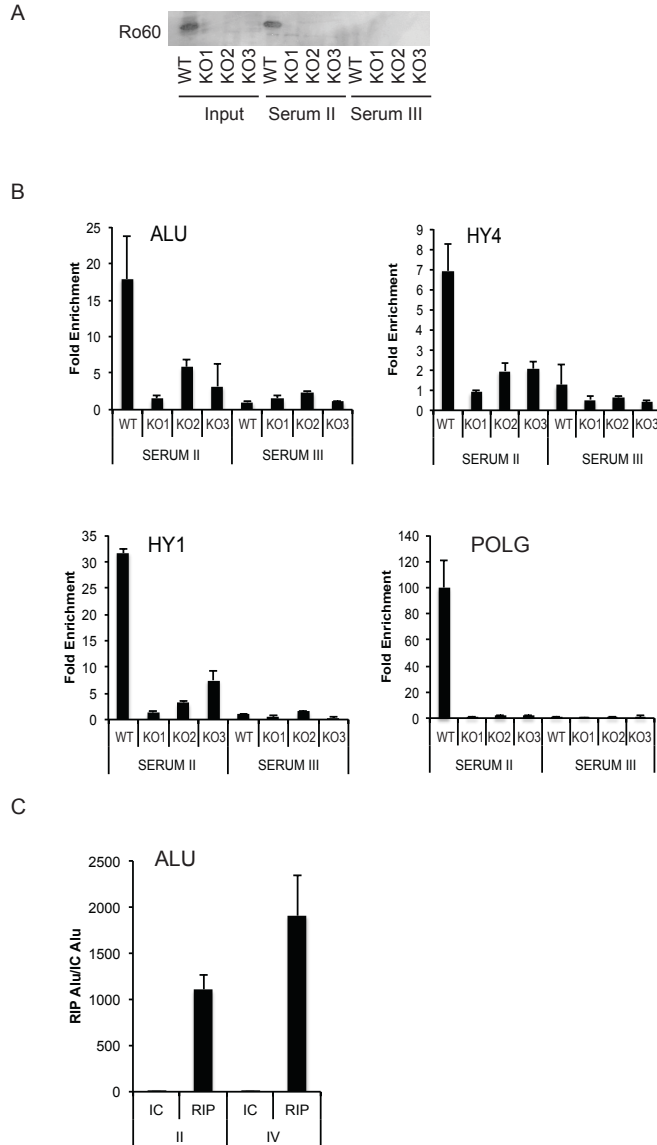


Fig. S11. SLE-RIP in Ro60 KO cells

(A) Western blot of Ro60 protein immunoprecipitated by SLE IgGs in anti Ro60 positive (II) or negative (III) SLE patients. SLE-IPs performed in wildtype or Ro60 $-/-$ GM12878 cells.

(B) RT-PCR of indicated RNAs within the SLE-immunoprecipitates from (A).

(C) Fold enrichment of RIP immunoprecipitated Alu RNA from GM12878 cells relative to endogenous immune complex Alu RNA from the two anti-Ro60 positive SLE patients II and IV.

Disease status		
control		18
SLE		99

Gender		
control	female	18
	male	0
SLE	female	93
	male	6

Ethnicity		
control	Hispanic	2
	Caucasian	14
	African	2
SLE	Hispanic	33
	Caucasian	44
	African	11
	Native American	6
	Other	5

ISM status		
control	NA	
SLE	ISM_low	29
	ISM_high	70

Fig. S12. Disease status, gender, ethnicity, and ISM status of SLE and control subjects profiled by RNA-seq

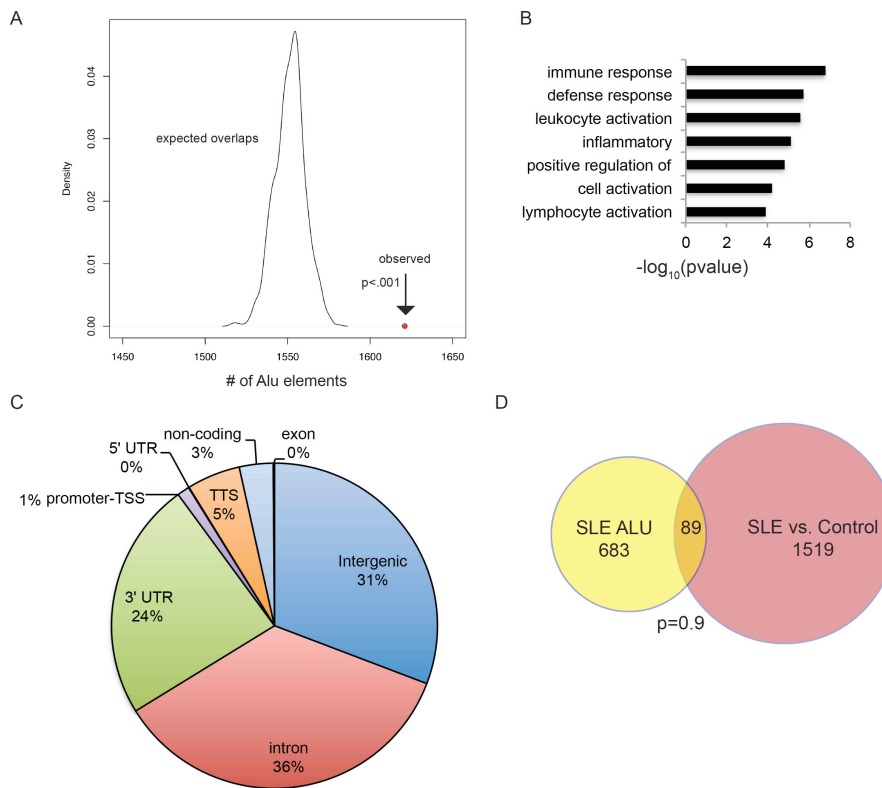


Fig. S13. Top SLE Alu RNAs

(A) Ro60 Alu Motif enriched within the 1,641 Alu elements highly expressed in SLE. Red marker= observed Alus with Ro60 Alu motif (1,621/1,641). Expected distribution of Alus with the motif plotted in black (1000 permutations of 1,641 random Alu elements).

(B) GO terms enriched in the SLE Alu elements plotted in Fig. 4F.

(C) Annotation of SLE Alu elements plotted in Fig. 4F (n=1,641) to nearest gene structure.

(D) Highly expressed Alu RNAs did not derive from SLE differentially expressed genes. Intersection of genes with SLE ALUs and genes differentially expressed in SLE vs. control patients. The 1,641 SLE ALU RNAs collapsed to 772 genes. p value= hypergeometric test. Differentially expressed genes were >1.5 fold up or downregulated.

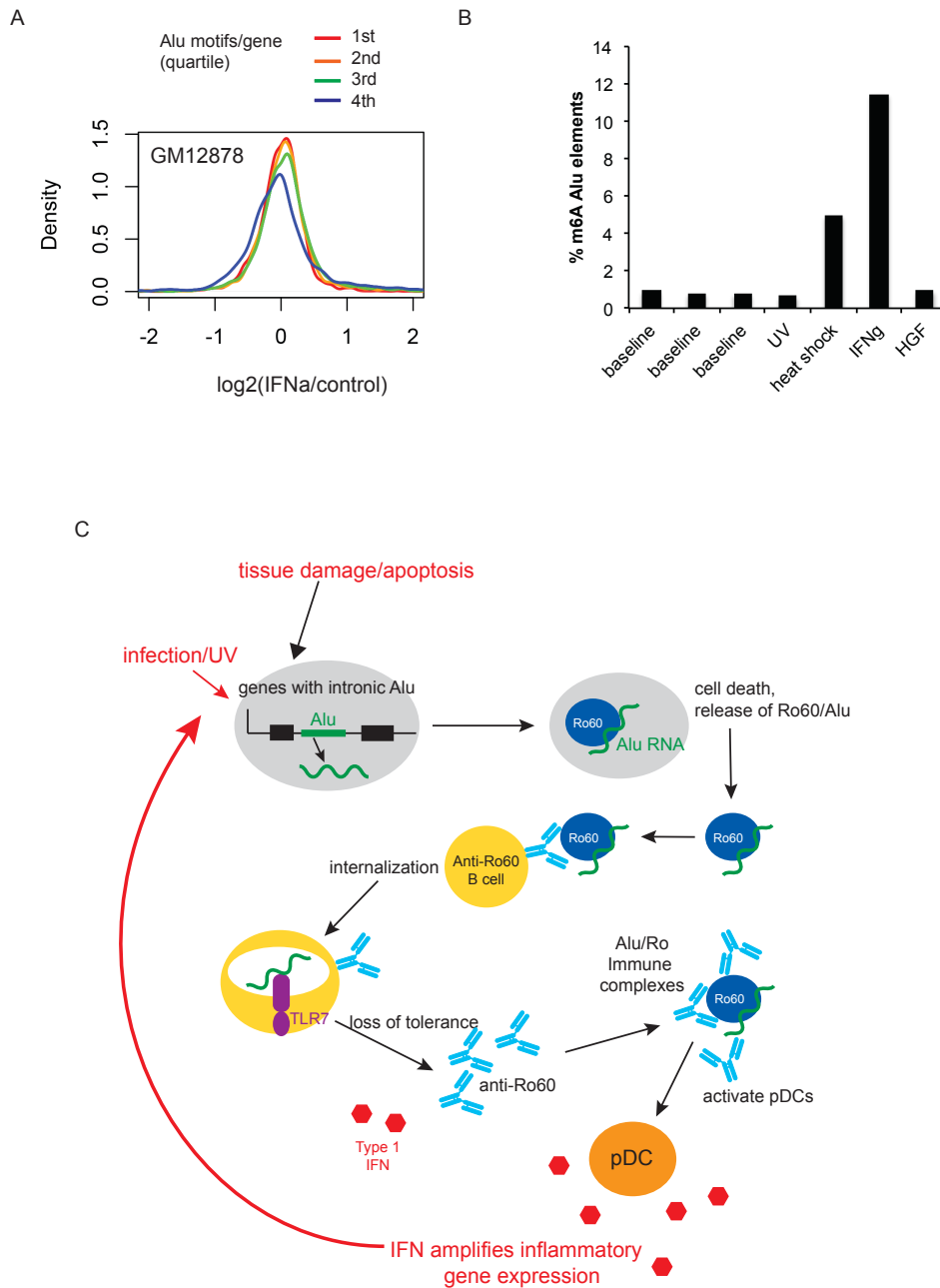


Fig. S14.

(A) Gene expression alteration upon 6 hours of IFN α treatment in GM12878 cells. All expressed genes (RPKM>5) binned by density of Alu motifs=(Ro60 Alu motifs/gene/10 Kb). Q1= >16, Q2= 9.9-16, Q3= 5.4-9.9, Q4= 0-5.4. p value Q4 vs. Q3 <.001. Each

quartile represents ~2000 genes. Gene expression is the average of 2 biological replicates. **(B)** Interferon-stimulated modifications to Alu RNA, such as m6A, may enhance affinity for Alu RNAs to Ro60 in response to interferon activation. Quantification of Alu transcripts in published RNA-seq dataset of m6a transcriptome in HEPG2 cells at baseline or treated as indicated. IFN γ was applied to cells overnight. Dataset and additional treatment details available at GEO GSE37002 (PMID 22575960). **(C) Model.** Endosomal uptake of Ro60/Alu RNA complexes by either naïve or anergic B cells with B cell receptor specificity for Ro60 epitopes results in the delivery of both an autoantigen – Ro60 - and a potent adjuvant – Alu RNA - to intracellular RNA sensors, including TLR7. This leads to aberrant activation of B cells, production of inflammatory cytokines and a break in normal tolerance. Activated B cells generating autoantibodies to Ro60 facilitate the production of anti-Ro60 Ab/Ro60/Alu RNA immune complexes, further fueling the cycle of Ro60/Alu RNA uptake and B cell and dendritic cell activation. Strategies to therapeutically block RNA sensor signaling or downstream production of cytokines in SLE and Sjögren's patients with anti-Ro60 autoantibodies are warranted.