SUPPLEMENTAL MATERIALS

List of Abbreviations

Supplemental Methods

Supplementary Figures S1-S22

List of abbreviations

AAV, adeno-associated virus

ChIP, chromatin immunoprecipitation

DNMT1, DNA-methyltransferase-1

EMSA, electromobility shift assay

ERS, endogenous retroviral sequence

EWSR1, Ewing sarcoma breakpoint region 1

HIV1, human immune deficiency virus 1

HSBP1, heat shock binding protein 1

IRF1, interferon regulatory factor 1

IRF2, interferon regulatory factor 2

LTR, long terminal repeat

mTOR, mechanistic target of rapamycin

mTORC1, mTOR complex 1

mTORC2, mTOR complex 2

NE, nuclear extract

NRF1, nuclear respiratory factor 1

PBL, peripheral blood lymphocytes

PSIP1, PC4 and SFRS1 interacting protein 1; also known as lens epithelium-derived growth factor of 75 kilodaltons (LEDGF/p75)

PARP1, poly-ADP polymerase 1

SLE, systemic lupus erythematosus

SFRS3, splicing factor, arginine/serine-rich 3

SFRS5, splicing factor, arginine/serine-rich 5

USF1, upstream stimulatory factor 1

Tat, transactivation region

TCR, T-cell receptor

TF, transcription factor

Supplemental Methods

Human subjects. A total of 165 female Caucasian patients with systemic lupus erythematosus (SLE) were investigated. All patients satisfied the criteria for a definitive diagnosis [1;2]. As controls, 111 healthy Caucasian healthy control (HC) females were examined. Age of SLE patients was 42.2±1.1 (range 18-68) years, while those of HC subjects was 41.3±2.3 (range 18-63) years. The study has been approved by the Institutional Review Board for the Protection of Human Subjects.

Separation and culture of human peripheral blood lymphocytes. Peripheral blood mononuclear cells (PBMC) were isolated from heparinized venous blood on Ficoll-Hypaque gradient. Peripheral blood lymphocytes (PBL) were separated from monocytes by adherence to autologous serum-coated Petri dishes [3]. T cells (>95% CD3+) were negatively isolated from PBMC with Dynal magnetic beads conjugated to IgG antibodies for CD14, CD16 HLA class II DR/DP, CD56 and CD235a; Invitrogen Cat No.113-11D). CD4+ T cells (>98% CD4+) were

negatively isolated with magnetic beads conjugated to IgG antibodies for CD8, CD14, CD16, HLA class II DR/DP, CD56, CDw123, and CD235a (Invitrogen Cat No.113-39D). The resultant cell population was resuspended at 10^6 cells/ml in RPMI 1640 medium, supplemented with 10% fetal calf serum (FCS), 2 mM L-glutamine, 100 IU/ml penicillin, and 100 µg/ml gentamicin in 12-well plates at 37ºC in a humidified atmosphere with 5% CO_2. Cross-linking of the CD3 antigen was performed by addition of cells to plates pre-coated with 100 µg/ml goat anti-mouse IgG (Jackson, West Grove, PA) for 2 h and, after washing, pre-coated with 1 µg/ml OKT3 monoclonal antibody (CRL 8001 from ATCC, Rockville, MD) for 1 h at 37ºC. CD28 co-stimulation was performed by addition of 500 ng/ml mAb CD28.2 (Pharmingen, San Diego, CA).

Cell lines. HeLa and HeLa-tat human cervix carcinoma cells (from ATCC) were grown in Dulbecco's Modified Eagle's Medium with 10% fetal bovine serum (FBS), 2 mM L-glutamine, 100 U/ml penicillin, 100 µg/ml streptomycin, and 10 µg/ml amphotericin B. HCT116 cells (American Type Culture Collection) were cultured in McCoy's 5A modified medium supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin/amphotericin B [4]. HCT-166 colon carcinoma cells with wild-type DNMT1 and DNMT3 alleles (line #28) and HCT-166 colon carcinoma cells lacking DNMT1 (line #30), DNMT3b (line #38), or both DNMT1 and DNMT3b (DNMT1/3-DKO, line #343) were kindly provided by Dr. Bert Vogelstein [4]. MCF7 cells expressing wild-type and dominant-negative IRF1 were obtained from Dr. Robert Clarke [5]. Jurkat human T cell leukemia cells were maintained in RPMI 1640 medium supplemented with 10% FBS, 2 mM L-glutamine, and antibiotics. All cell lines were maintained in a humidified atmosphere with 5% CO_2 at 37°C, as earlier described [6]. Control and HIV-1 tat

gene-transfected HeLa and Jurkat cells were obtained from the NIH AIDS Research and Reference Program. Cell culture products were purchased from Cellgro (Mediatech Inc., Herndon, VA).

Construction and transfection of eukaryotic expression vectors. HRES-1/Rab4 was over-expressed in Jurkat cells using a doxycyclin-inducible GFP-encoding dicistronic expression vector system [7]. Jurkat cells were transfected by electroporation with pUHD172-1neo plasmid and selected for neomycin resistance in the presence of 1 mg/ml G418. These Jurkat cells produce the reverse *trans*-acting factor (rtTA) capable of interacting with promoters harboring tetracycline operator sequences only in the presence of tetracycline analogs. They were electroporated with GFP-producing pTR5-DC/GFP*TK/hygro control (construct 4480), wild-type (construct 6678), and dominant-negative (DN) mutant HRES-1/Rab4^{S27N}-containing vector (construct 9035) and doubly transfected cells were selected and maintained in 400 μg/ml G418 and 200 μg/ml hygromycin. Percentage of GFP-positive cells in an aliquot of the transfected cultures was periodically checked by flow cytometry. HRES-1/Rab4A expression and resultant changes in CD4 expression were monitored by western blot and flow cytometry [6].

Preparation of nuclear extracts (NE). Nuclear extracts (NE) were prepared according to a previously described protocol at 4°C [8]. Briefly, cells in log phase were harvested, washed with cold phosphate-buffered saline, and resuspended in 1 × packed cell volume of ice cold hypotonic buffer A (10 mM HEPES pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM phenylmethylsulfonyl fluoride (PMSF), 1 ug/ml aprotinin, 1 ug/ml pepstatin A, 1 mM DTT, and 10 mM leupeptin). The cells were incubated for 20 mins on ice to swell. Then, the cells were lysed by passing

through 25-gauge needle 6-7 times. The free nuclei were centrifuged at 16,000 × g and resuspended in 2/3× original packed cell volume of ice cold buffer C (20 mM HEPES pH 7.9, 20% glycerol, 550 mM NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 0.5 mM PMSF, 1 mM DTT, and 10 mM leupeptin). After intermitted vortexing at highest setting and constant stirring for 30 minutes, samples were centrifuged for 5 min at 16,000 × g. The supernatant thus obtained was nuclear extract (NE), which was quantified by Bradford's method (Bio-Rad Laboratories, Hercules, CA). The NE was aliquoted and stored at -80°C until further use.

Electrophoretic mobility shift assay (EMSA). Single-stranded oligonucleotides were customer-synthesized at Integrated DNA Technologies (Coralville, IA). Double-stranded synthetic oligonucleotides (Fig. 4) were generated by heating primer pairs to 95°C for 5 minutes in 100 mM NaCl, 10 mM Tris (pH 8), 1 mM EDTA and slow cooling to room temperature. They were aliquoted and stored at -20°C until further use. These double-stranded oligonucleotides were 5'-end-labeled with 4,500 Ci/mmol [γ-^{32}P]-ATP (Gamma-32, MP Biomedicals; Cat. No. 0135020, Santa Ana, CA) and T4 polynucleotide kinase (Promega, Madison, WI; Cat. No. M4101) for one hour at 37°C and purified on Sephadex G-25 spin columns for use as EMSA probes (MilliporeSigma, Bedford, MA; GE Healthcare Cat. No. 28-9034-08). 10 μg of NE was incubated with 20 fmol (45,000-60,000 cpm) of double-stranded probe in EMSA binding buffer (20 mM Tris pH 7.6, 50 mM NaCl, 1 mM $MgCl_2$, 0.2 mM EDTA, 5% glycerol, 25 ng/μl poly(dI-dC)•poly(dI-dC) or 25 ng/μl sonicated salmon sperm DNA, 30 μg of BSA, 0.1 mM PMSF, and 0.5 mM DTT) for 30 minutes at room temperature in a total volume of 20 μl in the presence or absence of 10 to 100-fold molar excess cold competitor DNA. In supershift experiments, 2 μl of TF specific or transaldolase (TAL) or actin control antibody was

preincubated with NE in EMSA binding buffer for 1 hour on ice followed by a 30 minute incubation with the ^{32}P-labeled probe at room temperature. The following TF antibodies were used in supershift assays: IRF1 (sc-497x, sc-74530x), IRF2 (sc-498x; sc-374327x), PSIP1 (BD Signal Transduction Cat.No., 611714; Abcam Cat.No. ab 177150, ab49281; Bethyl Cat. No. A300-847A), NRF1 (Abcam Cat. No. ab34682), USF1 (sc-229x), Arnt1 (sc-5580x, sc-8076x), HSBP1 (ProSci, Cat. No. 25-272), SFRS3 (ab73891), SFRS5 (Abcam 67175), PARP1 (sc-74469X), HSP70 (sc-59569), HSP90 (sc-13119X), EWSR1 (sc-28327). DNA-protein complexes were resolved on a pre-run 6% (29:1 acrylamide:bis-acrylamide) polyacrylamide gel run at 10 mA in 0.4 × TBE buffer (10x: 890mM Tris-borate, 890mM boric acid, 20mM EDTA); supershift experiments were electrophoresed in a 5% pre-run (29:1 Acrylamide:Bis-Acrylamide) polyacrylamide gel run at 29 mA in 0.4 × TBE. Following electrophoresis, gels were dried under vacuum at 80°C for 1 hour, and exposed to Molecular Dynamics Storage Phosphor screen. Images were developed Molecular Dynamics Typhoon 9410 Imager. (GE Amersham, Piscataway, NJ).

Site-directed mutagenesis of reporter plasmids. Mutagenesis of HRES-1/Rab4 promoter and LTR enhancer constructs were generated by polymerase chain reaction (PCR) using the QuickChange Site-Directed Mutagenesis Kit as suggested by the manufacturer (Stratagene, La Jolla, California). 25 ng of template was incubated with 125 ng of sense and antisense primers, dNTPs, and subjected to 18 PCR cycles with Pfu Turbo DNA polymerase with denaturation at 95°C for 1 minute, annealing at 55°C for 1 minute, and extension at 68°C for 10 minutes. DpnI was used to digest the parental supercoiled double-stranded methylated DNA for 1 hour at 37°C. Transformations were performed in Escherichia coli XL-1 Blue cells using DpnI-treated DNA.

Plasmids were prepared using Qiagen Plasmid Maxi Kit Columns (Qiagen, Valencia, CA). Introduction of mutations into the 5' promoter constructs were verified by DNA sequencing.

Transient transfections and reporter gene assays. HeLa and HeLa-tat cells were transfected with pGL4.11 (Promega, Cat. No. E6661) firefly luciferase plasmid vector-based reporter constructs using LipofectAMINE Plus reagent (Invitrogen Corporation, Carlsbad, CA). Each reporter plasmid was co-transfected with Renilla luciferase vector, hRluc/TK (pGL4.74, Promega, Cat. No. E6921) to normalize for potential variation in transfection efficiency. Promoter DNA was cloned into Kpn1 and Xho1 sites of pGL4.11 vector. LTR enhancer was inserted into the BamH1 site of pGL4.11. The conversion of genotype I LTR to genotype II LTR was carried out using Site Directed Mutagenesis kit (Stratagene/Agilent Technologies, Cat. No.200519-5). For each transfection, HeLa and HeLa Tat cells were plated at a cell density of 2.5×10^4 cells/well in a 48-well plate 24 h prior to transfection. Cells in each well were transfected with 200 ng of luciferase reporter vector and 50 ng of Renilla control vector using Lipofectamine PLUS Reagent (Invitrogen, Cat. No.15338-100). Forty-eight hr post transfection, cells harvested. Luciferase activity was assayed using the Dual-Glo System (Promega, Cat. No. E2920). Luminescence signals for Firefly and Renilla Luciferase were quantified with a BIOTEK bioluminescence plate reader (Winooski, VT). For all transfections, relative light unit value of Firefly Luciferase signal was normalized to the value of Renilla Luciferase signal to evaluate promoter activity.

Yeast One Hybrid (Y1H) Assay. Matchmaker One Hybrid kit was used for construction and screening of cDNA expression library (Clontech-Takara Cat.No. 630304, version PT3529-1). The LTR sequence between nucleotides +945 - +965 from the transcription start site of HRES-1/Rab4

was used as a DNA bait. Three copies of the DNA bait were inserted between the EcoR1 and Mlu1 restriction sites of the pHIS2.1 vector. Single-stranded oligonucleotides were custom-synthesized by IDT and made into double stranded DNA prior to cloning into pHIS2.1. For cDNA library construction, exponential growing Jurkat E.6.1 cells were used for extracting total RNA and isolating mRNA with oligo-dT. The cDNA construction and PCR amplification of the cDNA library was carried out as suggested by the Matchmaker protocol. The mRNA was reverse transcribed using enzyme MMLV reverse transcriptase and oligonucleotide primers CDSIII Oligo dT (5'ATTCTAGAGGCCGACCGACATG-d(T) 30 VN-3') and SMARTIII oligo (5'AAGCAGTGGTATCAACGCAGAGTGGCCATTATGGCCGGG 3'). The generated cDNA was amplified using 5' PCR Primer- 5'-TTCCACCCAAGCAGTGGTATCAACGCAGAGTGG-3' (homologous to SMART III oligo) and 3' PCR Primer- 5'-GTATCGATGCCCACCCTCTAGAGGCCGAGGCGGCCGACA-3' (homologous to CDSIII oligo dT). The amplified cDNA was purified using Chroma-Spin TE-400 columns supplied by kit. The eluate obtained after purification was precipitated with $NaCH_3COO$ and 100% ethanol and pelleted by centrifugation at 14,000 x g for 10min at 4^oC. The pellet obtained after centrifugation was resuspended in de-ionized water and stored in -20^oC until further use. For Y1H screening, sequential transformation method was used. Yeast strain Y187 was first transformed with pHIS2.1 containing DNA bait. The transformants were selected in SD/-Trp media. The HIS3 (Histidine synthesis) gene in pHIS2.1 plasmid is placed under the control of DNA bait, which drives its expression by binding to DNA binding proteins. To test and inhibit the leaky expression of this HIS3 in the absence DNA binding proteins, the transformants were grown in different concentrations of 3-AT (3-amino-1, 2, 4-triazole), a competitive inhibitor of HIS3 gene product. A 3-AT concentration of 50 mM was selected for further screening. In the second step of sequential transformation, the DNA bait pHIS2.1

containing yeast was grown in selective media SD/-Trp. These yeast cells were used for transformation with PCR amplified cDNA library and Sma1 linearized pGADT7-Rec2 plasmid. Transformants were selected in SD/-Trp/-His/-Leu/+50mM 3-AT media. The in vivo recombination enables amplified cDNA library to recombine with pGADT7-Rec2 plasmid, which contains regions homologous to CDSIII oligo and SMARTIII oligo, and proteins expressed as fusion partners of Gal4 activation domain. The positive colonies were selected in SD/-Trp/-His/-Leu/+50mM 3-AT media 3-5 days after transformation. Plasmids were isolated from the positive yeast colonies after overnight culture in SD/-Leu media. Cells from 1.5 ml culture volumes were centrifuged and resuspended in 500 ul of STES buffer (0.5 M NaCl, 0.2 M Tris pH7.5, 10 mM EDTA, 1% SDS). The yeast cells were lysed by vortexing with 0.5 µm glass beads for 1 min. 2 µl of crude supernatant was used to transform *E. coli* DH5α cells. Transformants were selected in media containing 50 µg/ml ampicillin. Overnight liquid culture of DH5α cells grown with ampicillin was used for plasmid isolation. This plasmid DNA was sequenced with the 5' T7 primer annealing 5' of the cDNA insert in the pGADT7-Rec2 cloning vector at the Genomic Core Facility of SUNY Upstate Medical University.

Affinity purification and southwestern blotting of TFs binding to the HRES-1 LTR.

Southwestern blotting was carried out to detect TF binding to the LTR nucleotide +945 to +965 with respect to the transcriptional start site. Firstly, proteins resolved in SDE-PAGE after DNA affinity pulldown from Jurkat NE, which was followed by transfer to nitrocellulose membrane. Secondly, membrane-bound proteins were hybridized to ^{32}P-labelled double-stranded oligonucleotides corresponding to bases +945 to +965 of the HRES-1 LTR. For each DNA affinity pulldown reaction, 110 µg of Jurkat NE was incubated with biotinylated double-stranded oligonucleotides in 300µl of LTR EMSA buffer for 1hr at 4^{o}C.The resultant DNA-protein

complexes were added to 50 µl of 50% slurry of streptavidin (SA) beads (Sigma S1638). These SA beads were prior blocked with 0.2 mg/ml BSA and 0.1 mg/ml salmon sperm DNA in 1 x EMSA buffer. DNA-protein complex was pulled down by centrifuging at 1000 rpm for 5 mins at 4^{o}C. The pelleted beads were washed in 1 ml of 1X LTR EMSA buffer 3 times. The firmly SA bead-bound proteins were eluted into 50 µl of 4 X SDS-PAGE loading dye and heated at 95^{o}C for 5mins. Subsequently, the proteins were resolved in 12% SDS-PAGE and later transferred onto nitrocellulose membrane. The proteins on membrane were hybridized with radioactive probe. The synthesis of radioactive probe was carried out as described earlier for EMSA. Before hybridization, the proteins on membrane were renatured. Both, the renaturing of the proteins on the membrane and hybridizing with radiolabelled probe were carried out in 1X binding buffer (25mM Hepes-NaOH pH 7.9, 60mM KCl, 1mM EDTA pH 8.0, 5mM $MgCl_2$, 0.5mM DTT). For renaturing of proteins on membrane, the nitrocellulose membrane was sequentially exposed to 1X binding buffer containing decreasing concentrations of guanidine HCl. In first step, binding buffer contained 6 M guanidine-HCl (GuHCl), followed by 3M GuHCl, 1.5M GuHCl, 0.75M GuHCl, 0.375M GuHCl, 0.187M GuHCl and 0.094M HCl. For each step, the membrane was incubated in 20 ml of GuHCl solution in 1 x binding buffer at 4^{o}C under rotating conditions. Last wash was carried out in 1 x Binding buffer without the GuHCl. Subsequently, the membrane was transferred into blocking solution (1x binding buffer containing 0.5% Gelatin and 10 µg/ml salmon sperm DNA) for 1hr at 4^{o}C under rotating conditions. Then, the membrane was incubated with ^{32}P-radiolabelled double-stranded oligonucleotides (1 x 10^{6} cpm/ml) in 10ml of 1 x binding buffer with 0.25% gelatin and 10 µg/ml salmon sperm DNA. The hybridization was carried out overnight at room temperature under rotating conditions. The unbound probe was washed in 20 ml of 1 x binding buffer for 5 mins four times. The membrane was air-dried and

exposed to Storage Phosphor Screen under Saran wrap. The image was developed using software Image Quant 5.1 using Typhoon 9410 Imaging System (GE).

.

Trypsin digestion and mass spectrometry. Mass spectrometry was carried out on peptides generated through trypsin digestion of proteins that had exhibited binding to NRF1 sites 1 (nucleotides -50 to -67) and 2 (nucleotides -114 to -124) of the 5' promoter and to the HRES1 LTR enhancer (nucleotides +945 to +965). Following affinity pulldown, using biotinylated DNA and SA beads as described for Southwestern blotting, the pulldown protein products were analyzed in SDS page using silver staining compatible with mass spectrometry [9].
The bands of interest were cut from the gel and transferred into an Eppendorf tube. The gel slices were washed in 0.1% trifluoroacetic acid and 50% acetonitrile. The bands were destained twice in 200 µl of destaining solution (100 mM NH_4HCO_3, 50% Acetonitrile) for 45 min. The gel was completely dried with 100 % acetonitrile for 10 mins at room temperature. The gel was then placed into 20 ul of trypsin digestion buffer (40mM NH4HCO3/10% Acetonitrile) with 0.5ug/0.5ul of trypsin (Trypsin Gold, Mass Spectrometry Grade; Promega Cat. No. V5280). Then, 50 ul of additional trypsin digestion buffer was added to cover the gel slice completely with buffer.

Trypsin was reconstituted at a concentration 1µg/µl in 50 mM acetic acid. This stock solution of trypsin was diluted into working stock solution to 100 ng/µl (10-fold) by taking 1µl of stock solution into 9 µl of protein buffer B (5 mM B-mercapto-ethanol, 0.1% SDS, 50 mM Tris-HCl pH 8.0). The DNA affinity pulldown reactions were carried out as described earlier in Southwestern blotting. The pulldown reactions were resuspended in 40µl of Protein buffer B (5 mM B-mercapto-ethanol, 0.1% SDS, 50 mM Tris-HCl pH 8.0). The proteins in the sample were

denatured by heating at 95^{o}C for 15 mins. The beads were allowed to cool before adding trypsin. Trypsin solution of 1 µl from working stock (100ng/µl) was added into each sample tube. As a control for each sample, a sample without trypsin was also included. The trypsin digestion was carried out overnight at 37^{o}C under rotating conditions at 200 rpm. The beads were spun down at 2000 rpm for 5mins at 4^{o}C. The supernatant was removed in a separate tube; 1/3rd of the sample was run in 15% SDS-PAGE and silver stained to check for complete digestion. To make the digested samples mass spectrometry compatible, the detergent (0.1% SDS) from the samples was cleared by using ZIPTIP (Millipore, Cat. No. ZTSC XS0-96). Initially, the ZIPTIPs were equilibrated in washing solution I (0.1%TFA). The trypsin digested samples were made to bind to ZIPTIP by pipetting several times. The detergent from the samples was removed by pipetting several times with washing solution II (0.1% TFA/30% methanol). Then, the samples were eluted in elution buffer (5% ammonium hydroxide/ 30% methanol). Lastly, the samples were air-dried and resuspended in 25 µl of 1% TFA solution and proceeded to mass spectrometry using Thermo LTQ Orbitrap instrument.

Chromatin Immunoprecipitation (Chip). Chip assay was carried out using the Zymo-Spin ChIP Kit (Zymo Research, Irvine CA) according to the manufacturer's instructions. Jurkat or HeLa cells were cross-linked for 7 minutes in PBS with 1% (v/v) formaldehyde, after which they were quenched by incubation with 125 mM glycine for 5 minutes. Following successive 5 minute incubations with the Nuceli Prep and Chromatin Shearing buffers, cells were sonicated with a Sonic Dismembrator (Fisher Scientific, Waltham MA) at amplitude setting 3, using 4 cycles of 30s 'On' and 30s 'Off'. The resultant lysates were then cleared and diluted 5-fold in Chromatin Dilution buffer. All buffers were supplemented with 1 mM PMSF, 1 µg/ml aprotinin,

1 µg/ml pepstatin, 1 µg/ml leupeptin, 1 mM sodium orthovanadate, 0.1 mM sodium molybdate, and 10 mM sodium pyrophosphate. Lysates corresponding to an input of 10^6 cells were incubated overnight at 4°C with 3.0 µg of the following antibodies: IRF1, IRF2, USF1, ARNT, EWSR1, PARP1 (sc497x, sc498x, sc229x, sc5580x, sc28327, sc74669, Santa Cruz Biotechnology, Dallas TX), NRF1, PSIP1, SFSR3, SFSR5 (ab34682, ab49281, ab73891, ab67175, Abcam, Cambridge UK), HSBP1 (Aviva Biosciences, San Diego CA), actin (MAB1501 Millipore, Billerica MA) and transaldolase [10] . As loading control, actin antibody was placed onto each membrane to accurately authenticate changes in expression of each specific gene. Zymo-Mag Protein A beads were then added, and the lysates were incubated for an additional hour at 4°C. After washing, chromatin-protein complexes were eluted by resuspending the beads in Chromatin Elution buffer and incubation at 75°C for 5 minutes, after which the eluates were reverse cross-linked through incubation with proteinase K for 90 minutes at 65°C. After addition of ChIP DNA Binding Buffer, the eluates were added to Zymo-Spin IC columns, and ChIP DNA was eluted in DNA Elution Buffer. ChIP efficacy was evaluated through end-point PCR using the following primers: 5'-CGTCTCGGTGGGGAAGGACC-3' and 5'-GTGCGCACGCGCGAACG-3' (NRF1-site 1), 5'-CGGGGCGCGTGAGGACTT-3' and 5'-CTACCCGACCCTCCCTCAGG-3' (NRF1-site2), 5'-CCTCTCTCTGGGACTGTAAAACT-3' and 5'-CGTTTATTTCTCCCTCCCTCT-3' (HRES-1 LTR-enhancer). In ChIP assays where cells were transfected beforehand, HeLa cells were seeded into 6-well dishes, and transfections were carried out using 2.5µg plasmid DNA, 7.0 µl lipofectamine LTX, and 2.5µl PLUS reagent (Invitrogen, Carlsbad CA). Cells were harvested and processed for ChIP 48 hr post-transfection. Genomic DNA was used as positive control for PCR reactions.

Cloning of TF cDNA. Total RNA from exponentially growing Jurkat cells was isolated using RNeasy Protected Mini kit (Qiagen, Cat. No.74124). From total RNA, mRNA was isolated using Oligotex mRNA Mini kit (Qiagen, Cat. No.70022). 100ng of mRNA was reverse-transcribed using Superscript RT II (Invitrogen, Cat. No.11904-018). Individual transcription factors were amplified using gene specific primers, which also contained the site for restriction enzyme digestion. The ORFs were cloned into the EcoRI and SalI sites of the multiple cloning site within pAAV-IRES-hrGFP plasmid. The ORF for IRF2 was purchased from Origene (Cat. No. SC118745, pCMVXL5 containing IRF2 cDNA). After ligation, plasmids were transfected into *E.coli* cells (Stratagene, XL-Gold, Cat. No. 200314), transformants were selected in media containing 50 µg/ml ampicillin. HEK-293-derfived AAV-293 cells were transfected with plasmids pRC, pHelper and pAAV-IRES-hrGFP containing TF ORFs to produce AAV.

Construction and production of adeno-associated virus (AAV). Transcription factors (TF) were transduced into PBL by infection with AAV carrying cDNA upstream of the internal ribosomal entry site (IRES) within the pAAV-IRES-hrGFP vector (Stratagene/Agilent Technologies, Cat. No. 240071, La Jolla, CA), as described previously [6]. cDNAs for IRF1, HSBP1, EWSR1, SFRS3, SFRS5, PARP1 NRF1, and USF1 were cloned from RNA extracted from HeLa cells. The ORF for IRF2 was purchased from Origene (Cat. No. SC118745, pCMVXL5 containing IRF2 cDNA). PSIP1 cDNA was purchased from Genecopoeia (Rockville MD). cDNAs were ligated between the EcoR1 and Sal1 sites of pAAV-IRES-hrGFP. Recombinant pAAV-IRES-hrGFP plasmids were co-transfected with pHelper and pRC-AAV plasmids into AAV-293 cells using the CalPhos Mammalian Transfection Kit (Clontech, Mountain View, CA) according to the

manufacturer's directions. Transfections were scaled up using T-175 flasks and 20 µg of each plasmid. The media was replaced with fresh DMEM supplemented with 10% FBS, 2mM L-glutamine, 1 mM sodium pyruvate, and antibiotic-antimycotic 6 hr post-transfection. Cells were harvested by scraping 72hr after transfection and lysed by 4 freeze-thaw cycles with a dry-ice ethanol bath. After centrifugation, the supernatant was passed sequentially through Millex-SV 5.00 µm and Vented Millex-AA 0.8 µm syringe-driven filters (Millipore). The filtered, AAV-containing supernatants were then concentrated by centrifugation through a 15 ml of Amicon Ultrafree CL filter units (Millipore Cat No. UFC4BHK25, Burlington, MA). Concentrated virus preparations were aliquoted and stored frozen at -80°C until use. Concentrated AAV supernatants were used to infect primary human PBL or negatively-isolated T cells, Jurkat and Jurkat-tat cells. Typically, 200µl of concentrated viral stock was incubated with $2x10^6$ cells in 200µl of serum-free RPMI supplemented with 2mM L-glutamine. The incubation was allowed to proceed for 4hr at 37°C in 5% CO_2, after which 1.6ml of RPMI with 10% FBS, 2mM L-glutamine, and antibiotic-antimycotic was added. Transduction was verified by flow cytometry of 10^5 per sample (LSRII, BD Biosciences, San Jose CA). The remainder of the cells were lysed directly at a density of $2x10^6$ live cells per 100 µl of 4 x SDS-PAGE sample buffer. GFP expression was monitored by flow cytometry and equilibrated among constructs (excitation: 488 nm, emission: 509 nm; FL1 channel). Optimal GFP expression occurred at 48 h post infection of PBL with >99% of cells infected (**Figure S22**). Primary PBL and T cells were analyzed 24 h post infection, due to a decline in cell viability at 48 h. Over-expression of TFs was confirmed by western blotting.

Transfection of HeLa and HeLa-tat cells with TFs. HeLa and HeLa tat cells were transfected with pAAV-IRES-hrGFP containing TF ORFs. The cells were plated at a cell density of $1x10^5$

cells/well in a 6 well plate 24 hr prior to transfection. The cells were transfected with Lipofectamine PLUS Reagent. 24 hr and 48 hr post–transfection, the cells were harvested through trypsinization and an aliquot was used to detect GFP signal through flow cytometry to confirm transfection. Protein lysates were prepared as described for PBL.

Western blot analyses. Whole cell protein lysates were prepared by lysis in radio-immunoprecipitation assay buffer (150 mM NaCl, 2% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, 50 mM Tris pH 8.0, 1 mM PMSF, 1 μg/ml aproptinin, 1 μg/ml pepstatin, 1 μg/ml leupeptin, 1 mM NaF, 1 mM sodium orthovanadate, 0.1 mM sodium molybdate, 10 mM sodium pyrophosphate) at a density of 4 x 10^7 cells/ml on ice, followed by addition of equal volumes of Laemmli protein sample buffer (60 mM Tris-Cl pH 6.8, 2% SDS, 10% glycerol, 5% β-mercaptoethanol, 0.01% bromophenol blue) and heated to 95°C for 5 minutes prior to separation on SDS-PAGE gels and transfer to 0.45 μm nitrocellulose membranes. HRES-1/Rab4 was detected by primary rabbit antibodies directed to the C-terminus (Santa Cruz sc-312) or the full-length native protein (Abcam ab13407 or Abcam ab108974) [6]. Transcription factors were detected with primary antibodies to NRF1 (Abcam 34682), IRF1 (sc-497X), IRF2 (sc-498X), PSIP1 (Abcam-177159), HSBP1 (ProSci, Cat. No. 25-272), USF1 (sc-229X), EWSR1 (sc-28327), SFRS3 (ab73891), SFRS5 (Abcam 67175), and HSBP1 (ProSci, Cat. No. 25-272). Actin was detected with monoclonal antibody clone C4, MAB 1501R, from MilliporeSigma. 4E-BP1 (#9644), phospho-4E-BP1 (Thr 37/46, #2855), pan-Akt (#4685), and phospho-Akt (Ser 473, #4058) antibodies were obtained from Cell Signaling. p70 S6 kinase (sc-8418), and phospho-p70 S6 kinase (Thr 389, sc-8416) antibodies were from Santa Cruz Biotechnology. Human CD4 was detected with monoclonal mouse antibody 4B12 from Novocastra/Leica (NCL-L-CD4-368).

Reactivities to primary antibodies were detected with horseradish peroxidase-conjugated secondary antibodies (Jackson, West Grove, PA) and visualized by enhanced chemiluminescence (Western Lightning Chemiluminescence Reagent Plus, GE Health Care/PerkinElmer Life Sciences, Inc., Boston, Massachusetts). Automated densitometry was used to quantify the relative levels of protein expression using a Kodak Image Station 440CF with Kodak 1D Image Analysis Software (Eastman Kodak Company, Rochester, NY).

Statistical analysis. Statistical analyses were performed with Prism Version 5.0 for Windows (GraphPad Software, San Diego, CA). Data were expressed as the mean ± standard error of the mean (SE) of individual experiments. Changes were considered significant at p-value < 0.05.

Reference List

1. Tan EM, Cohen AS, Fries JF *et al*. The 1982 revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum.* 1982;**25**:1271-7.

2. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum.* 1997;**40**:1725.

3. Perl A, Gonzalez-Cabello R, Lang I *et al*. Effector activity of OKT4+ and OKT8+ T-cell subsets in lectin- dependent cell-mediated cytotoxicity against adherent HEp-2 cells. *Cell.Immunol.* 1984;**84**:185-93.

4. Rhee I, Bachman KE, Park BH *et al*. DNMT1 and DNMT3b cooperate to silence genes in human cancer cells. *Nature* 2002;**416**:552-6.

5. Bouker KB, Skaar TC, Fernandez DR *et al*. Interferon Regulatory Factor-1 Mediates the Proapoptotic but Not Cell Cycle Arrest Effects of the Steroidal Antiestrogen ICI 182,780 (Faslodex, Fulvestrant). *Cancer Res.* 2004;**64**:4030-9.

6. Nagy G, Ward J, Mosser DD *et al*. Regulation of CD4 Expression via Recycling by HRES-1/RAB4 Controls Susceptibility to HIV Infection. *J.Biol.Chem.* 2006;**281**:34574-91.

7. Gossen M, Freundlieb S, Bender G *et al*. Transcriptional activation by tetracyclins in mammalian cells. *Science* 1995;**268**:1766-9.

8. Grossman CE, Qian Y, Banki K *et al*. ZNF143 Mediates Basal and Tissue-specific Expression of Human Transaldolase. *J.Biol.Chem.* 2004;**279**:12190-205.

9. Shevchenko A, Wilm M, Vorm O *et al*. Mass Spectrometric Sequencing of Proteins from Silver-Stained Polyacrylamide Gels. *Anal.Chem.* 1996;**68**:850-8.

10. Banki K, Halladay D, Perl A. Cloning and expression of the human gene for transaldolase: a novel highly repetitive element constitutes an integral part of the coding sequence. *J.Biol.Chem.* 1994;**269**:2847-51.

11. Thurman RE, Rynes E, Humbert R *et al*. The accessible chromatin landscape of the human genome. *Nature* 2012;**489**:75-82.

12. Quandt K, Frech K, Karas H *et al*. MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucl.Acids Res.* 1995;**23**:4878-84.

13. Cartharius K, Frech K, Grote K *et al*. MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics* 2005;**21**:2933-42.

14. Wang J, Zhuang J, Iyer S *et al*. Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors. *Genome Res.* 2012;**22**:1798-812.

15. Absher DM, Li X, Waite LL *et al*. Genome-Wide DNA Methylation Analysis of Systemic Lupus Erythematosus Reveals Persistent Hypomethylation of Interferon Genes and Compositional Changes to CD4+ T-cell Populations. *PLoS Genet* 2013;**9**.

16. Falli T, Le Dantec C, Thabet Y *et al*. DNA methylation modulates HRES1/p28 expression in B cells from patients with lupus . *Autoimmunity* 2014;**47**:265-71.

17. Perl A, Rosenblatt JD, Chen IS *et al*. Detection and cloning of new HTLV-related endogenous sequences in man. *Nucl.Acids Res.* 1989;**17**:6841-54.

18. Magistrelli C, Samoilova E, Agarwal RK *et al*. Polymorphic genotypes of the HRES-1 human endogenous retrovirus locus correlate with systemic lupus erythematosus and autoreactivity. *Immunogenetics* 1999;**49**:829-34.

19. Tanaka N, Kawakami T, Taniguchi T. Recognition DNA sequences of interferon regulatory factor 1 (IRF-1) and IRF-2, regulators of cell growth and the interferon system. *Mol.Cell.Biol.* 1993;**13**:4531.

20. Escalante CR, Yie J, Thanos D *et al*. Structure of IRF-1 with bound DNA reveals determinants of interferon regulation. *Nature* 1998;**391**:103-6.

21. Turlure F, Maertens G, Rahman S *et al*. A tripartite DNA-binding element, comprised of the nuclear localization signal and two AT-hook motifs, mediates the association of LEDGF/p75 with chromatin in vivo. *Nucl.Acids Res.* 2006;**34**:1653-65.

22. Sgarbanti M, Borsetti A, Moscufo N *et al*. Modulation of Human Immunodeficiency Virus 1 Replication by Interferon Regulatory Factors. *J.Exp.Med.* 2002;**195**:1359-70.

23. Fernandez DR, Telarico T, Bonilla E *et al*. Activation of mTOR controls the loss of TCR. in lupus T cells through HRES-1/Rab4-regulated lysosomal degradation. *J.Immunol.* 2009;**182**:2063-73.

24. Caza TN, Fernandez D, Talaber G *et al*. HRES-1/RAB4-Mediated Depletion of DRP1 Impairs Mitochondrial Homeostasis and Represents a Target for Treatment in SLE. *Ann.Rheum.Dis.* 2014;**73**:1887-97.

25. Talaber G, Miklossy G, Oaks Z *et al*. HRES-1/Rab4 promotes the formation of $LC3^{+}$ autophagosomes and the accumulation of mitochondria during autophagy. *PLoS ONE* 2014;**9**:e84392.

26. Moulton VR,.Tsokos GC. T cell signaling abnormalities contribute to aberrant immune cell function and autoimmunity. *J.Clin.Invest.* 2015;**125**:2220-7.

Legends of Supplemental Figures

Figure S1. Regulatory DNA was delineated within the entire 545 bp promoter by DNase I hypersensitive sites [11], MatInspector Program [12;13] and selected for analyses on the basis of ChIP-seq annotations generated by the ENCODE Consortium [14]. Blue highlighted region corresponds to nucleotide positions 229406661-229407517 which were found to be undermethylated (less than 10%) in Build 37 of the human genome (GRCh37/19; http://genome.ucsc.edu). The promoter DNA and the transcription factors found to bind in 125 cell types are bracketed in red.

Figure S2. DNA sequence of base positions 229406534-229407258 in human chromosome 1 (Build 37, GRCh37/19) that contains the promoter region of HRES-1/Rab4. NRF1 binding sites 1 and 2 and altered methylation state of flanking CpG motifs are indicated. Regions with hypomethylated and hypermethylated CpG motifs in CD4 T cells of 49 SLE patients relative to 58 controls are highlighted in black and red, respectively [15]. This earlier genome-wide DNA methylation analysis utilized the Infinium 450K BeadChip Kit [15]. Based on our analysis of raw methylation data [15], we identified seven CpG motifs spanning 857 bases between nucleotide positions 229406661-229407517 in the HRES1/Rab4 locus, corresponding to nucleotide positions -220 to +636 relative to the transcriptional start site[6], which were methylated at 3.6±1.1% and 3.7±1.1% in DNA of CD4 T cells from healthy and lupus subjects, respectively. Relative to this undermethylated region, 11 adjacent CpG motifs spanning 122,860 bases upstream (controls: 71.8±5.6%, $p=2.4x10^{-8}$; SLE: 71.8±5.7%, $p=3.8x10^{-8}$) and 51167 bases downstream displayed markedly greater methylation (controls: 86.1±3.1%, $p=2.8x10^{-11}$; SLE: 86.0±2.9%, $p=1.3x10^{-11}$). However, no differences were noted between SLE and controls subjects within these

hypomethylated or hypermethylkated regions flanking the HRES-1/Rab4A locus. Similar to our study, the CpG at 229406711, which is 170 bases upstream from the transcription start site, was found to be 5.18-fold hypermethylated in $CD4^+$ lupus T cells ($p=8.6x10^{-4}$; **Figure S2**).

Figure S3. DNA sequence of base positions 229407358-229408308 in human chromosome 1 (Build 37, GRCh37/19) that contains the LTR region of HRES-1/Rab4. Inverted repeats demarcating the LTR and CpG motifs analyzed for methylation state by bisulfite sequencing in this study are highlighted in green. CpG motif found to be hypomethylated in SLE by differential HpaII/MspI digestion is highlighted in yellow [16]. The HIV1 trans-activating region (TAR) is bolded and underlined with broken lines. SNP rs451401 within polymorphic Hind III site is shaded in grey.

Figure S4. Genotyping of SNP rs451401 in the HRES-1 LTR. **A**, Detection of heterozygous rs451401 alleles in HCT-166 colon carcinoma cells by Hind III digestion of the HRES-1 LTR following PCR amplification from genomic DNA. PBL genomic DNA of human subjects with homozygous CC, homozygous GG and heterozygous GC alleles were used as controls. **B**, Genotyping of HeLa, MCF-7 and HEK-293 cells. **C**, Genotyping of Jurkat human T cells carrying doxycycline-inducible vectors (control, 4480) that overexpress wild-type HRES-1/Rab4 (construct 6678) or dominant-negative HRES-1/$Rab4^{S27N}$ (HRES-1/Rab4-DN; construct 9035).

Figure S5. Effect of nitric oxide (NO) on activity of the 5' minimal promoter of HRES-1/Rab4. HeLa cell were transfected with pGL4.11 firefly luciferase reporter plasmid with or without 223 bp (construct 9185) or 127 bp 5' promoter DNA (construct 9184). Each transfection reaction

contained pGL4.74 renilla luciferase control plasmid to equalize for transfection efficiency. Culture were incubated for 24 hours before harvest with or without 400 μM NOC-18, which was used as source of NO. Data represent mean ± SEM. P values reflect comparison with two-tailed paired t-test.

Figure S6. Effect of cytosine methylation at six CpG motits, M1-M6, in NRF1 site 1 of the HRES-1 promoter. Bar charts represent mean ± SEM of shifted band intensities in the absence (left panel) and presence of NRF1 antibody (right panel). P values reflect comparison using two-tailed paired t-test.

Figure S7. Effect of cytosine methylation at two CpG motits, M1 and M2, in NRF1 site 2 of the HRES-1 promoter on binding of NRF1 and USF1 by EMSA using Jurkat NE.

Figure S8. SNP rs451401 influences TF binding to the HRES-1 LTR. **A,** Sequences of oligonucleotides used to assess transcription factor (TF) binding to the HRES-1 LTR region harboring rs451401 SNP that creates a G→C transition at base position +959 relative to the transcription start site [6]. This SNP at chromosome 1 base position 229407840 generates a polymorphic HindIII restriction site [17;18]. LTR genotypes I and II are designated based on the absence or presence of the polymorphic Hind III restriction site, respectively. The position of recognition motifs for IRF1 and IRF2 [19;20] and AT-hook target sequence of PSIP1 [21] are bracketed within the oligonucleotides used for EMSA. **B**, EMSA of ^{32}P-labelled HRES-1 LTR DNA using Jurkat nuclear extracts (NE). Specificity of shifted complexes were controlled with inclusion of 10-fold and 100-fold molar excess of unlabeled (cold) oligonucleotides. Seven shifted complexes

(C1-C7) were detected. C6 was formed with genotype II LTR only. **C**, Supershift analysis of TF binding to the HRES-1 LTR using antibodies to IRF1, IRF2, and transaldolase (TAL). Following the pretreatment of NE with antibodies, EMSA revealed that DNA-protein complexes 2 (C2) and 3 (C3) are formed by IRF1 and IRF2, respectively. Neither IRF1 nor IRF2 antibodies affected the migration of genotype II-specific DNA-protein complex C6. IRF2 antibody generated a new, shifted complex, C8.

Figure S9. Isolation of TFs that bind the HRES-1 genotype II LTR. Proteins forming genotype II-specific complex C6 were excised from polyacrylamide gels following EMSA. Bands isolated and polled for protein sequencing are bracketed in red.

Figure S10. Isolation of TFs through binding to biotinylated HRES-1 LTR oligonucleotides following DNA affinity pulldown. Western blot detection of IRF1 and IRF2 following DNA affinity pulldown with HRES-1 LTR oligonucleotides. Pulldown of IRF1 and IRF2 was increased over avidin negative control and competed out with 100-fold access of non-biotinylated oligonucleotides.

Figure S11. Methylation of cytosine at position +954 in intron 1 affects TF binding to the HRES-1 LTR in allele-specific manner. **A,** Sequences of oligonucleotides used to assess TF binding to the HRES-1 LTR region harboring rs451401 SNP. This SNP originates from a G→C transition and results in a polymorphic Hind III restriction site [17;18]. LTR genotypes I and II are designated based on the absence or presence of Hind III restriction site, respectively. Effect of methylation of cytosine^{+954}, which is highlighted in red, was tested on TF binding to genotype I and genotype II

LTR. **B,** Methylation of genotype II LTR blocks the formation of C6 and elicits the formation of a new complex, designated, CII_m. **C**, Effect of methylation on TF binding to genotype I LTR. Nuclear extracts (NE) were pre-incubated with and without antibodies directed to TFs detected by mass spectrometry of trypsin-digested proteins extracted from C6. None of the antibodies affected binding to methylated genotype I LTR. **D,** Effect of methylation on TF binding to genotype II LTR. Formation of CII_m with genotype II LTR was inhibited by antibodies to IRF2 and PSIP1; PSIP1 also elicited the formation of a novel complex with methylated genotype II LTR.

Figure S12. Effect of methylation on TF binding affinity to genotype I and genotype II HRES-1 LTR. **A**, Effect of antibodies on TF binding to HRES-1 LTR oligonucleotides shown in Figure 6A, which have been unmethylated or methylated at cytosine^{+954}. EMSA was performed after incubation of Jurkat NE with or without the indicated antibodies in the presence of 1 µg/ml poly-(dIdC) as non-specific DNA competitor. **B**, Bar chart represents the formation of complex 8 with HRES-1 LTR oligonucleotides following incubation with IRF2 antibody. The relative affinity of IRF2 to form C8 complex with unmethylated genotype II LTR was compared to that with unmethylated genotype I LTR. Data represent mean ± SEM. P values represent comparison using two-tailed paired t-test.

Figure S13. Chromatin immunoprecipitation (ChIP) analysis of TF binding to the 5' promoter and the LTR-enhancer *in vivo* [8]. As a negative control, antibodies to β-actin (Actin) or transaldolase (TAL) were utilized. As a positive control, genomic DNA (input) was used for PCR amplification. **A**, Chip analysis of TF binding to the NRF1 and NRF2 sites in Jurkat and Jurkat-tat

cells. **B**, Chip analysis of TF binding to the NRF1 and NRF2 sites in HeLa cells. **C**, Chip analysis of TF binding to the LTR enhancer in HeLa cells.

Figure S14. Chip analysis of TF factor to the promoter of proton-coupled folate transporter (**PCFT**). Selective binding of NRF1 to the PFCT promoter was confirmed in HeLa cells, as a positive control for in vivo TF binding assessment by Chip analysis. As a negative control, antibody to actin was utilized. Genomic DNA (input) was used as positive control for PCR amplification.

Figure S15. Suppression of HRES-1/Rab4 expression by IRF1 is relieved by HIV-tat in HeLa cells. Transfection of pAAV-IRES-GFP expression vectors was confirmed by GFP fluorescence and IRF1 expression. Top panels show representative western blots. Bottom panel show cumulative analysis of independent experiments. Means are normalized to 1.0 in HeLa and HeLa-tat cells. P values reflect comparison with paired t-test.

Figure S16. Suppression of HRES-1/Rab4 expression by IRF-1 in MCF7 cells. Rab4A expression was is reduced in MCF7 cells constitutively expressing IRF1. Overexpression of dominant-negative IRF1 (dnIRF1) increased the expression of HRES1/Rab4 in MCF7 cells.

Figure S17. Differential regulation of Rab4A expression by IRF2 in HeLa and Jurkat cells. **A**, IRF2 stimulated the expression of Rab4A in HeLa cells, which is unaffected by HIV-tat. **B**, Suppression of Rab4A expression by IRF2 is relieved by HIV-tat in Jurkat cells. Transduction of pAAV-IRES-GFP expression vectors was confirmed by GFP fluorescence and IRF2 expression.

Top panels show representative western blots. Bottom panel show cumulative analysis of independent experiments. P values reflect comparison with unpaired t-test.

Figure S18. Effect of PSIP1 deletion on expression of Rab4A in HEK-293t cells that carry homozygous genotype I (rs451401/GG) LTR. **A**, Western blot analysis of Rab4A expression in the absence of PSIP1. **B**, Effect of AAV on expression of Rab4A in the absence of PSIP1. Western blot analysis was performed 24 hours after infection with virus. **C**, Cumulative analysis of 3 independent experiments. P values reflect comparison with paired t-test.

Figure S19. Effect of AAV-mediated transduction of SFRS3 on the expression of HRES-1/Rab4 in normal human PBL. Left panels show representative western blots. Right panel shows cumulative analysis of four independent experiments. P value reflect comparison with paired t-test.

Figure S20. Effect of transfected SFRS5 on the expression of HRES-1/Rab4 in HeLa and HeLa-tat cells. Left panels show representative western blots. Right panel shows cumulative analysis of 4-6 independent experiments. Means are normalized to 1.0 in HeLa cells. P values reflect comparison with unpaired t-test.

Figure S21. Effect of transfected HSBP1 on the expression of HRES-1/Rab4 in HeLa and HeLa-tat cells. Top panels show western blots. Bottom panel shows cumulative analysis of 3 independent experiments. Means are normalized to 1.0 in HeLa and HeLa-tat cells. P values reflect comparison with paired t-test.

Figure S22. Flow cytometry of GFP fluorescence following transduction of adeno-associated virus (AAV) carrying pAAV-IRES-GFP vector alone or PSIP1 inserted upstream of the IRES sequence. Representative forward scatter and side scatter dot plot and histogram overlays of mock-transfected, control virus (GFP) and PSIP1- and GFP-expressing virus (PSIP1) are shown.

Figure S1

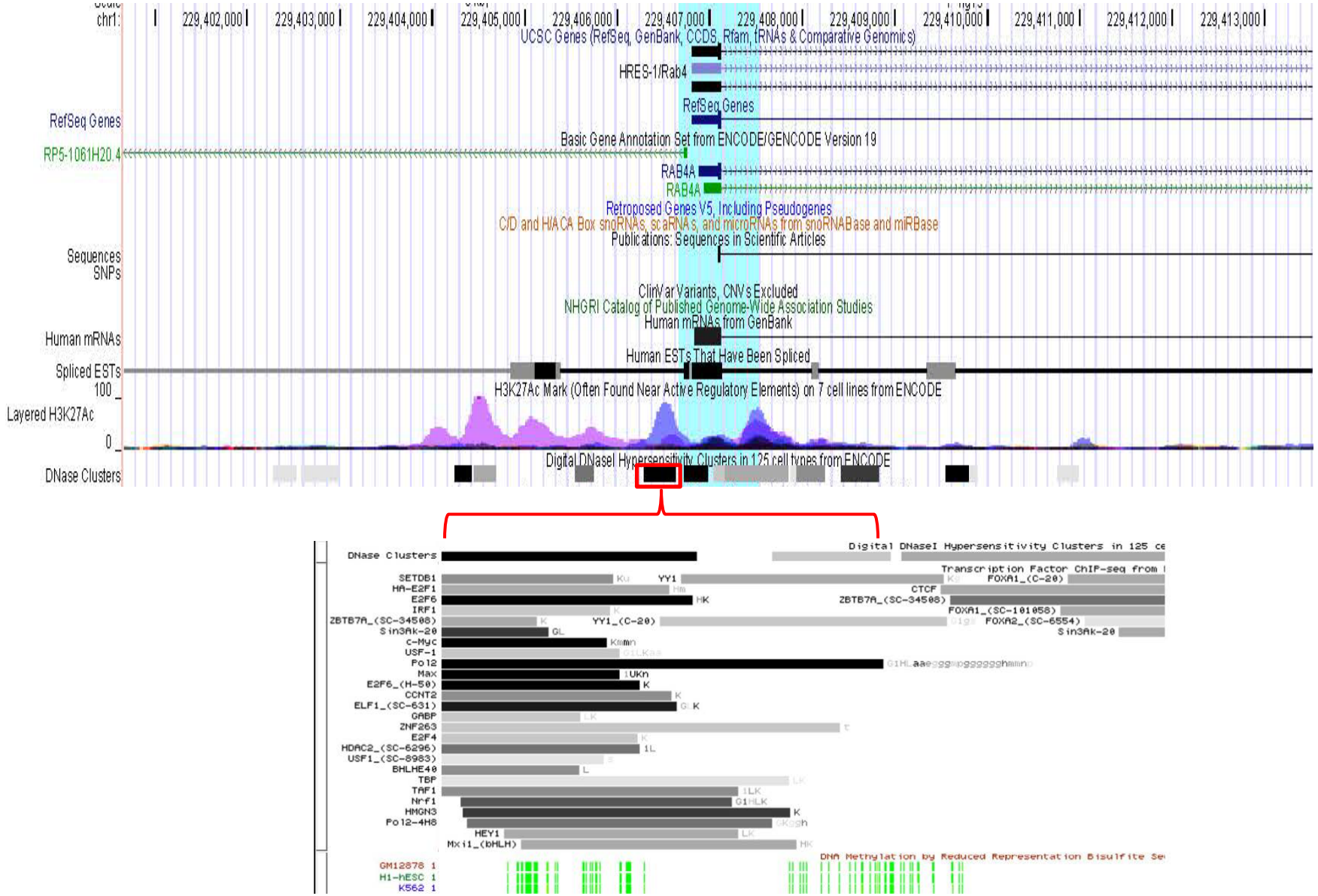

chr1:
229,402,000
229,403,000
229,404,000
229,405,000
229,406,000
229,407,000
229,408,000
229,409,000
229,410,000
229,411,000
229,412,000
229,413,000
UCSC Genes (RefSeq, GenBank, CCDS, Rfam, tRNAs & Comparative Genomics)
HRES-1/Rab4
RefSeq Genes
RefSeq Genes
Basic Gene Annotation Set from ENCODE/GENCODE Version 19
RP5-1061H20.4
RAB4A
RAB4A
Retroposed Genes V5, Including Pseudogenes
C/D and H/ACA Box snoRNAs, scaRNAs, and microRNAs from snoRNABase and miRBase
Publications: Sequences in Scientific Articles
Sequences
SNPs
ClinVar Variants, CNVs Excluded
NHGRI Catalog of Published Genome-Wide Association Studies
Human mRNAs
Human mRNAs from GenBank
Human ESTs That Have Been Spliced
Spliced ESTs
100
H3K27Ac Mark (Often Found Near Active Regulatory Elements) on 7 cell lines from ENCODE
Layered H3K27Ac
0
Digital DNaseI Hypersensitivity Clusters in 125 cell types from ENCODE
DNase Clusters
DNase Clusters
Digital DNaseI Hypersensitivity Clusters in 125 ce
Transcription Factor ChIP-seq from
SETDB1
HA-E2F1
E2F6
IRF1
ZBTB7A_(SC-34508)
Sin3Ak-20
c-Myc
USF-1
Pol2
Max
E2F6_(H-50)
CCNT2
ELF1_(SC-631)
GABP
ZNF263
E2F4
HDAC2_(SC-6296)
USF1_(SC-8983)
BHLHE40
TBP
TAF1
Nrf1
HMGN3
Pol2-4H8
HEY1
Mxi1_(bHLH)
YY1
YY1_(C-20)
FOXA1_(C-20)
CTCF
ZBTB7A_(SC-34508)
FOXA1_(SC-101058)
FOXA2_(SC-6554)
Sin3Ak-20
GM12878 1
H1-hESC 1
K562 1
DNA Methylation by Reduced Representation Bisulfite Se

Figure S2

Genomic chr1 :

Absher low CpG methylation region: 229406661-229407517: c
Absher high CpG methylation borders:229406534-229409358

rs451401: aagctt

chr1:229406534-

```
                          cgggc caatttggca aatggactgc  229406558
tttcctcccc agtgcgggca ccttcctcct ctgtccggca actggacccc  229406608
actgtcccca gccctctgtc acccccaaac ctgcccaccg cccgcctcta  229406658
cccgaccctc cctcagggaa atcgggactc agcccccaac gcccccacct  229406708
```

5.18-fold hypermethylated in CD4+ lupus T cells ($p=8.6x10^{-4}$)

'NRF1' site 2

```
gccgctctgc ccacctcagc gcacaccgcc agagcgccca gcccgcgcac  229406758
gtgcttccgg cgcgcgccga agtcctcacg cgccccgccc cggcctcccc  229406808
```

NRF1 site 1

```
GTGCGCACGC GCGAACGCGG GGTGGGCCGG GCCTCGCGTA GCCCATCTCC  229406858
```

Transcription start site

```
TCTTCCTCCT CGCGGTCGCG GCCGGACGGA GGGTGGAGGG CCCTGCGCCT  229406908
GCGCGGAGCT GGAGTCCGGC TGGGCCGCAG CCGCTGGGAG ACCGGCGGTT  229406958
GCCGTGGGGA CCGGTCGGGC CCCTCCCTCC TCCGGTCCCC CGCCCCAGGT  229407008
CCTTCCCCAC CGAGACGCGC CGGCGGACCG CGGGCGAGTG CAGCCGGTGA  229407058
```

Translation start site

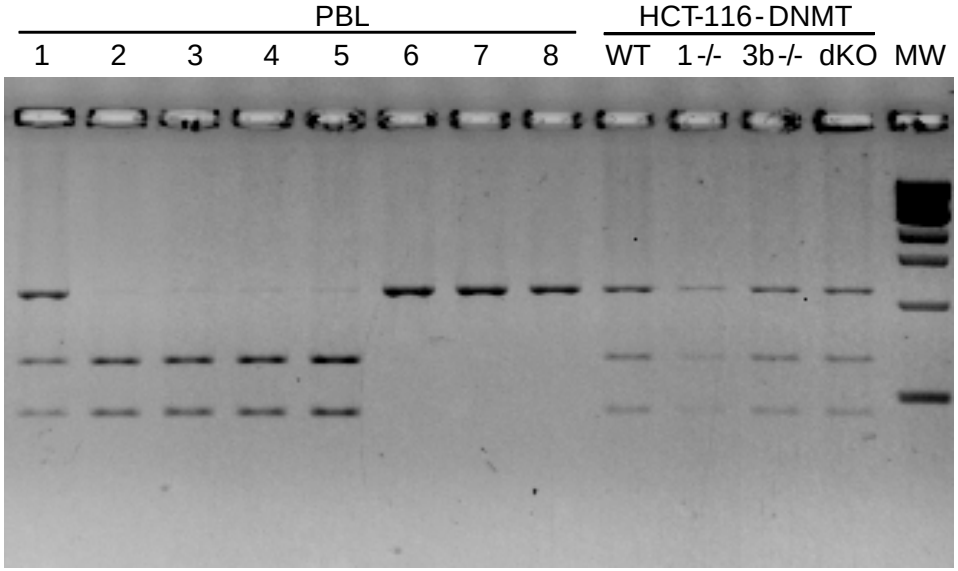
```
CCCGGCGAGA GGCGGCGCCG CTCCCAAGAT GTCGCAGACG GCCATGTCCG  229407108
AAACCTACGg tacgaggccc gggctggcgg ggcgcgcggg tcgggccgcg  229407158
gggggcgtcg gtggcgcggg gccgggcctg ggctgcgggg gtcttgaggg  229407208
tggcggtggc gccggccgga gccgggggac ggatctcgga ggtgtctggg  229407258
```

Figure S3

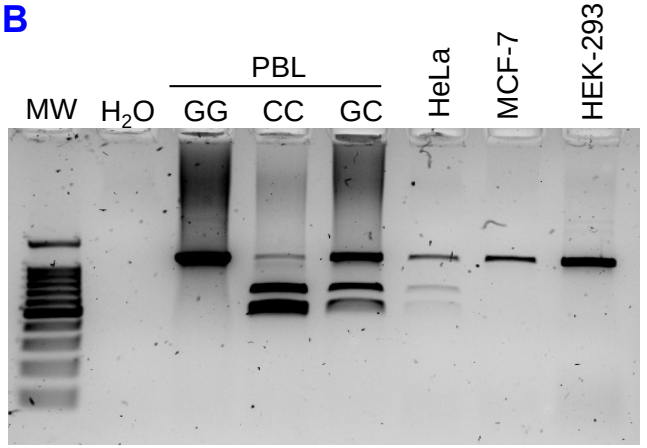

LTR inverted repeat
229407630 CpG
229407737 CpG #3
229407834 CpG #2
2229407896 CpG #1
LTR inverted repeat

Figure S4

A



B



C

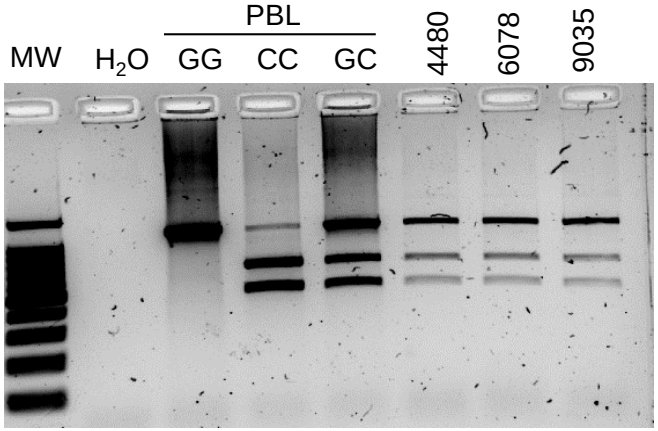


Figure S5

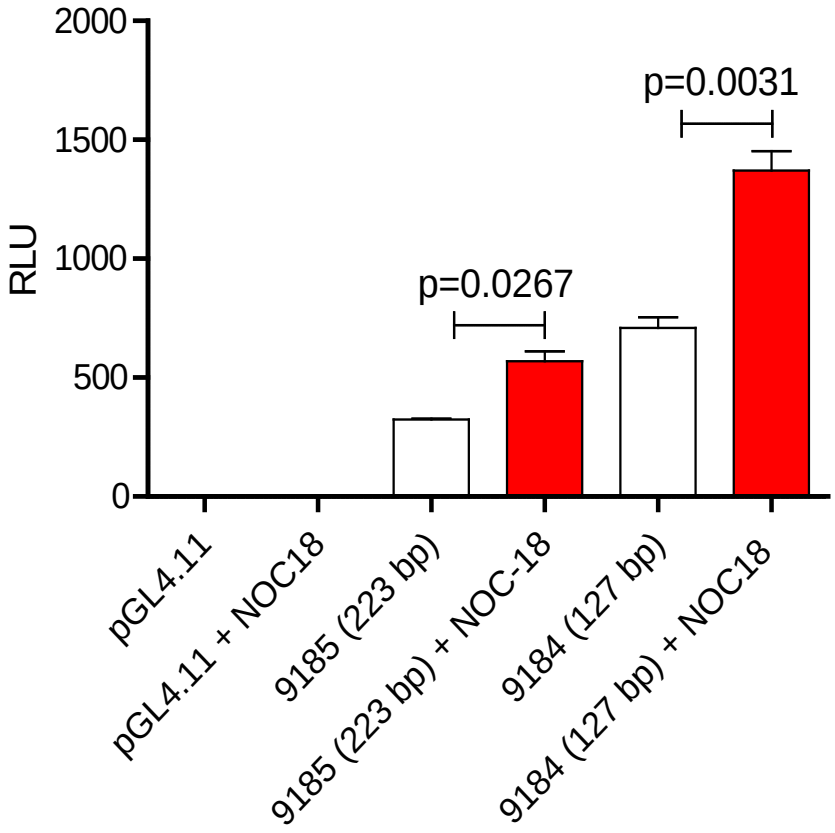

2000
1500
1000
500
0
RLU
p=0.0031
p=0.0267
pGL4.11
pGL4.11 + NOC18
9185 (223 bp)
9185 (223 bp) + NOC-18
9184 (127 bp)
9184 (127 bp) + NOC18

Figure S6

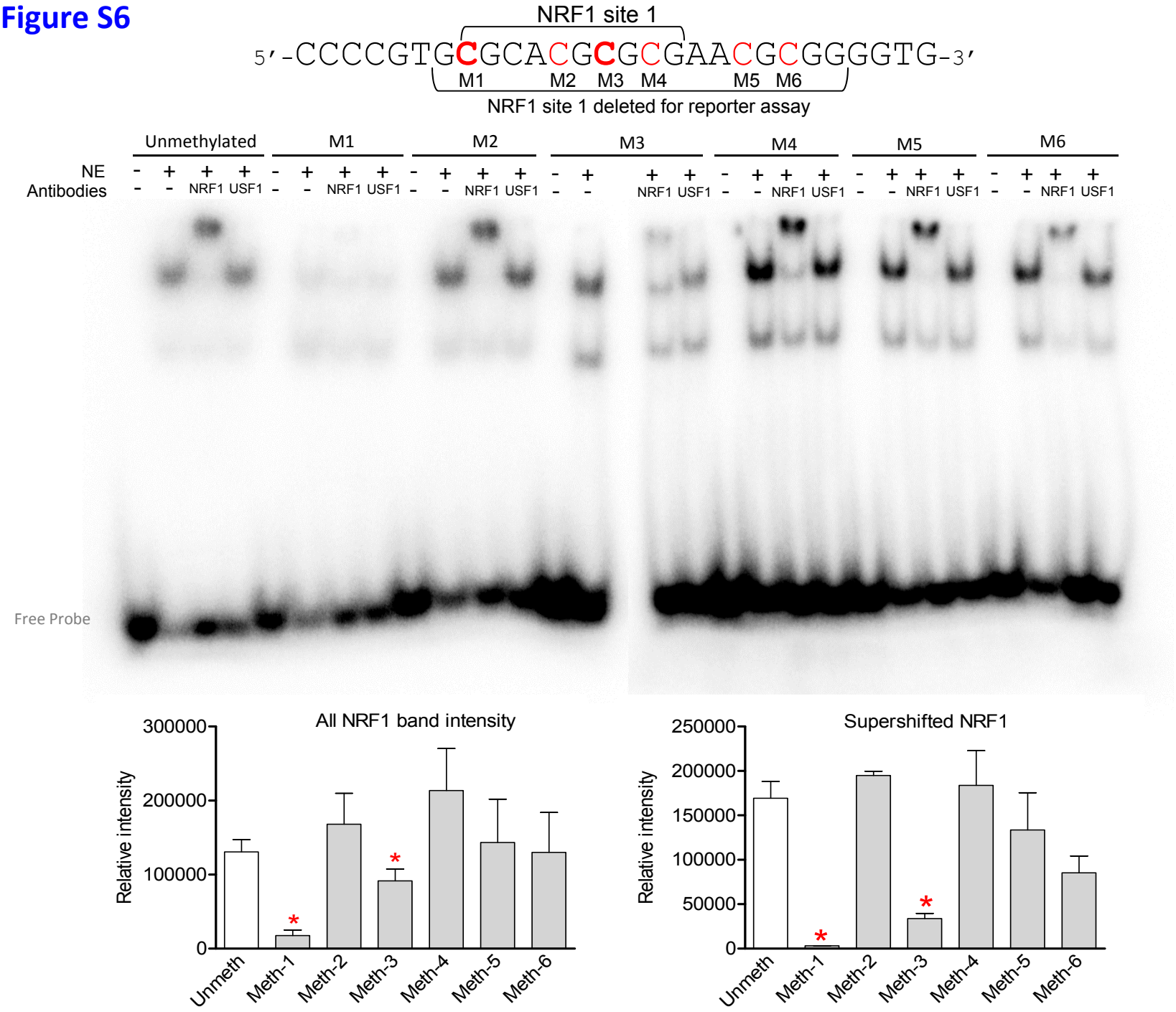
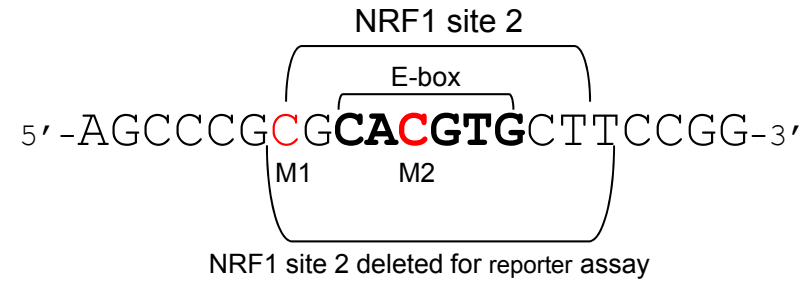

NRF1 site 1
5′-CCCCGTGCGCACGCGCGAACGCGGGGTG-3′
M1 M2 M3 M4 M5 M6
NRF1 site 1 deleted for reporter assay
Unmethylated
M1
M2
M3
M4
M5
M6
NE
Antibodies
NRF1
USF1
Free Probe
All NRF1 band intensity
Relative intensity
300000
200000
100000
0
Unmeth
Meth-1
Meth-2
Meth-3
Meth-4
Meth-5
Meth-6
Supershifted NRF1
250000
200000
150000
100000
50000
0

Figure S7


NRF1 site 2
E-box
5′-AGCCCGCGCACGTGCTTCCGG-3′
M1
M2
NRF1 site 2 deleted for reporter assay

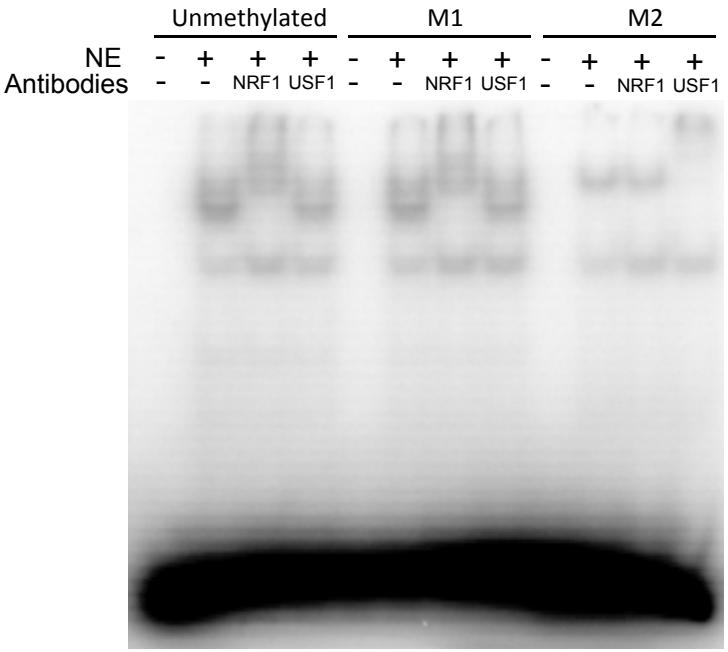

Unmethylated
M1
M2
NE - + + + - + + + - + + +
Antibodies - - NRF1 USF1 - - NRF1 USF1 - - NRF1 USF1

Figure S8

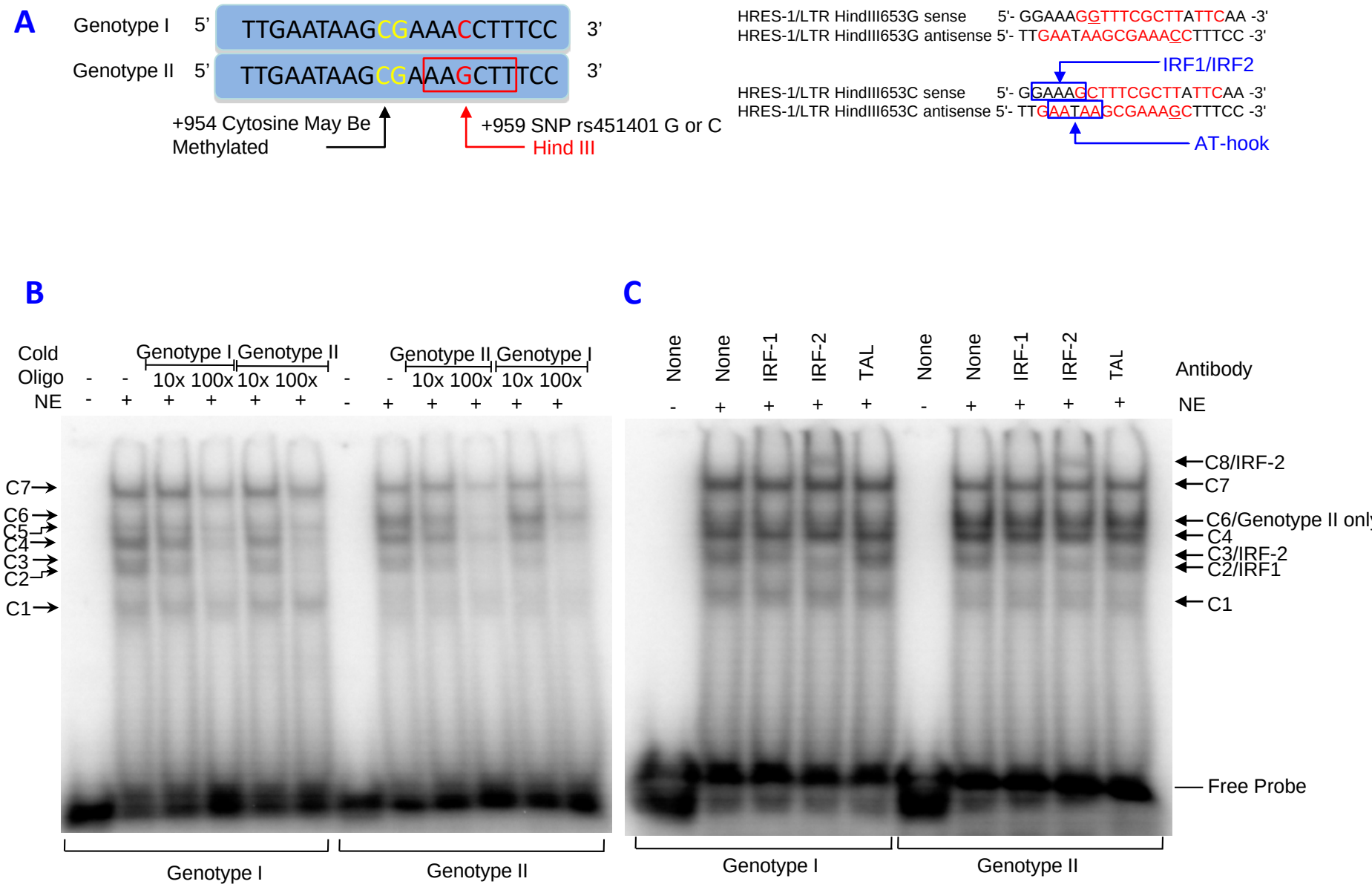

A
Genotype I 5' TTGAATAAGCGAAACCTTTCC 3'
Genotype II 5' TTGAATAAGCGAAAGCTTTCC 3'
+954 Cytosine May Be Methylated
+959 SNP rs451401 G or C
Hind III
HRES-1/LTR HindIII653G sense 5'- GGAAAGGTTTCGCTTATTCAA -3'
HRES-1/LTR HindIII653G antisense 5'- TTGAATAAGCGAAACCTTTCC -3'
IRF1/IRF2
HRES-1/LTR HindIII653C sense 5'- GGAAAGCTTTCGCTTATTCAA -3'
HRES-1/LTR HindIII653C antisense 5'- TTGAATAAGCGAAAGCTTTCC -3'
AT-hook
B
Cold Oligo NE
Genotype I Genotype II
Genotype II Genotype I
10x 100x 10x 100x
C7
C6
C5
C4
C3
C2
C1
Genotype I
Genotype II
C
None None IRF-1 IRF-2 TAL None None IRF-1 IRF-2 TAL
Antibody
NE
C8/IRF-2
C7
C6/Genotype II only
C4
C3/IRF-2
C2/IRF1
C1
Free Probe
Genotype I
Genotype II

Figure S9

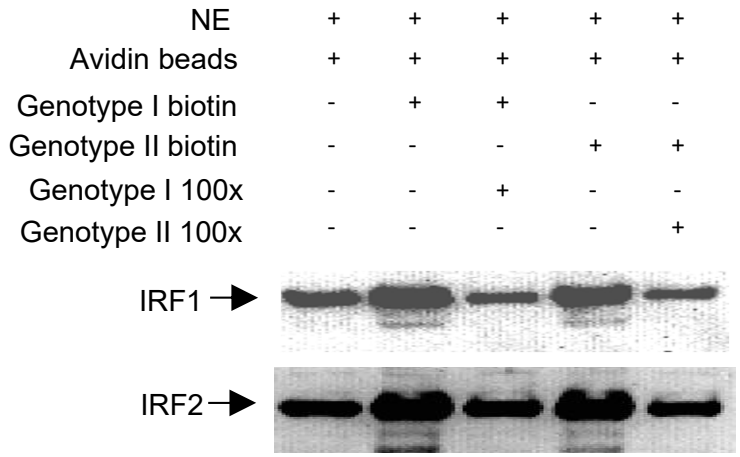

NE + + + + +
Avidin beads + + + + +
Genotype I biotin - + + - -
Genotype II biotin - - - + +
Genotype I 100x - - + - -
Genotype II 100x - - - - +
IRF1
IRF2

Figure S10

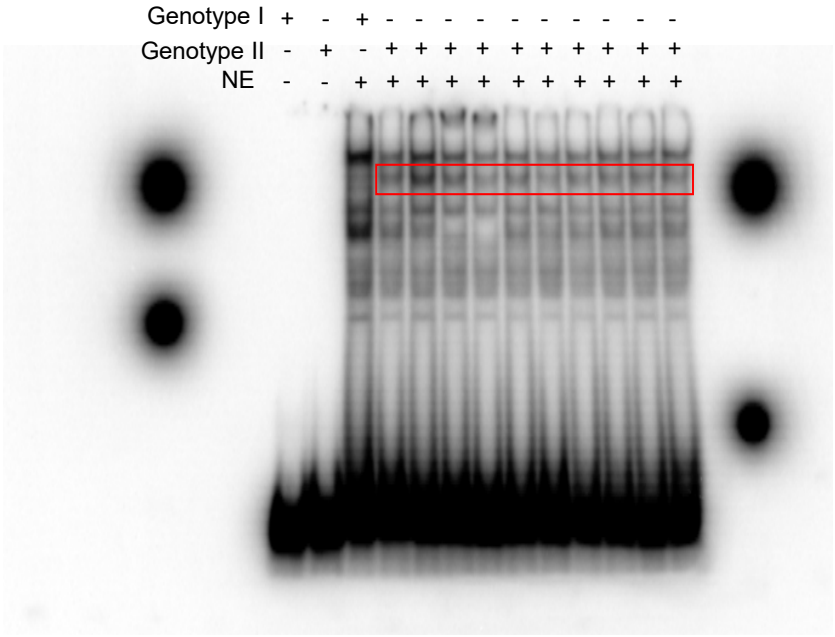


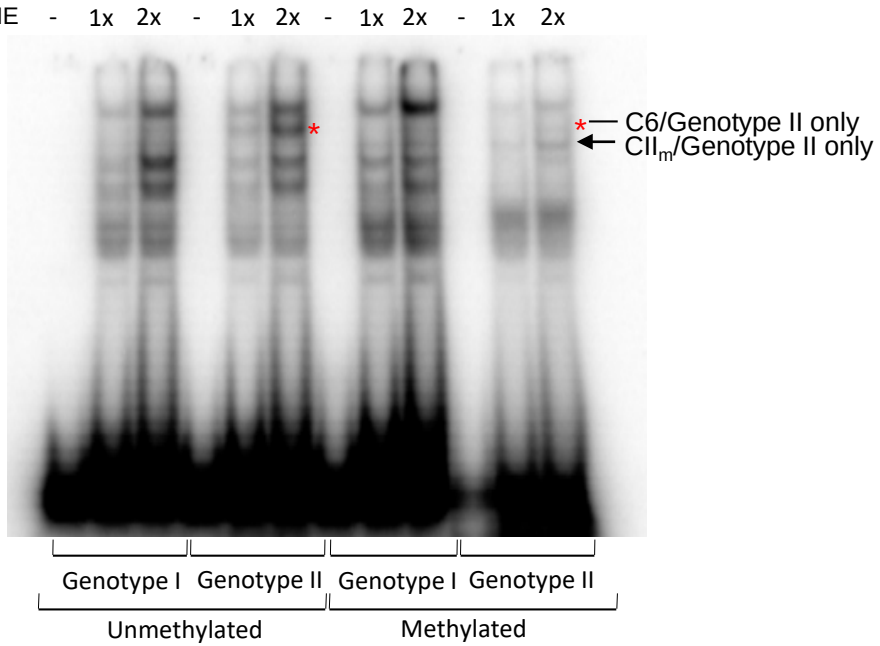
Figure S11

A

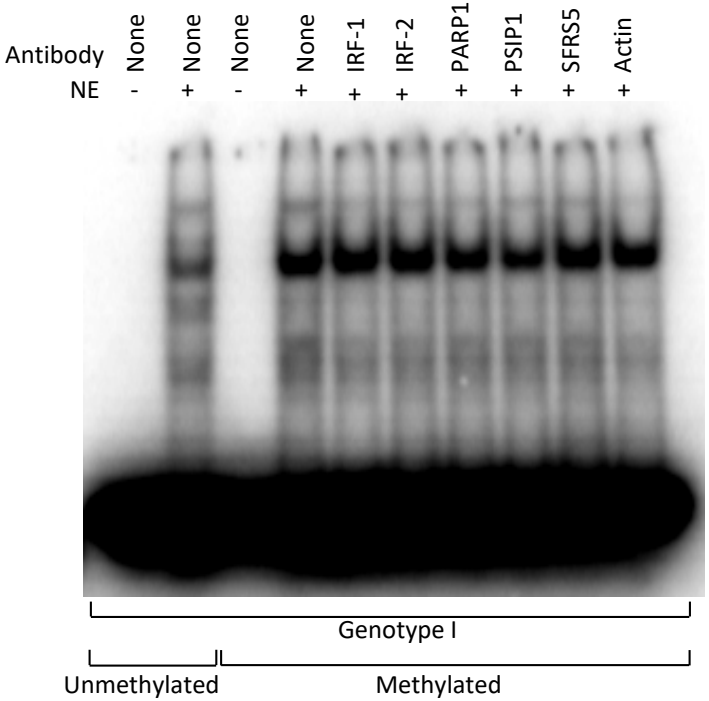
HRES-1/LTR HindIII653G sense 5'- GGAAAGGTTTCGCTTATTCAA -3'
HRES-1/LTR HindIII653G antisense 5'- TTGAATAAGCGAAACCTTTCC -3'

HRES-1/LTR HindIII653C sense 5'- GGAAAGCTTTCGCTTATTCAA -3'
HRES-1/LTR HindIII653C antisense 5'- TTGAATAAGCGAAAGCTTTCC -3'

B



C



D

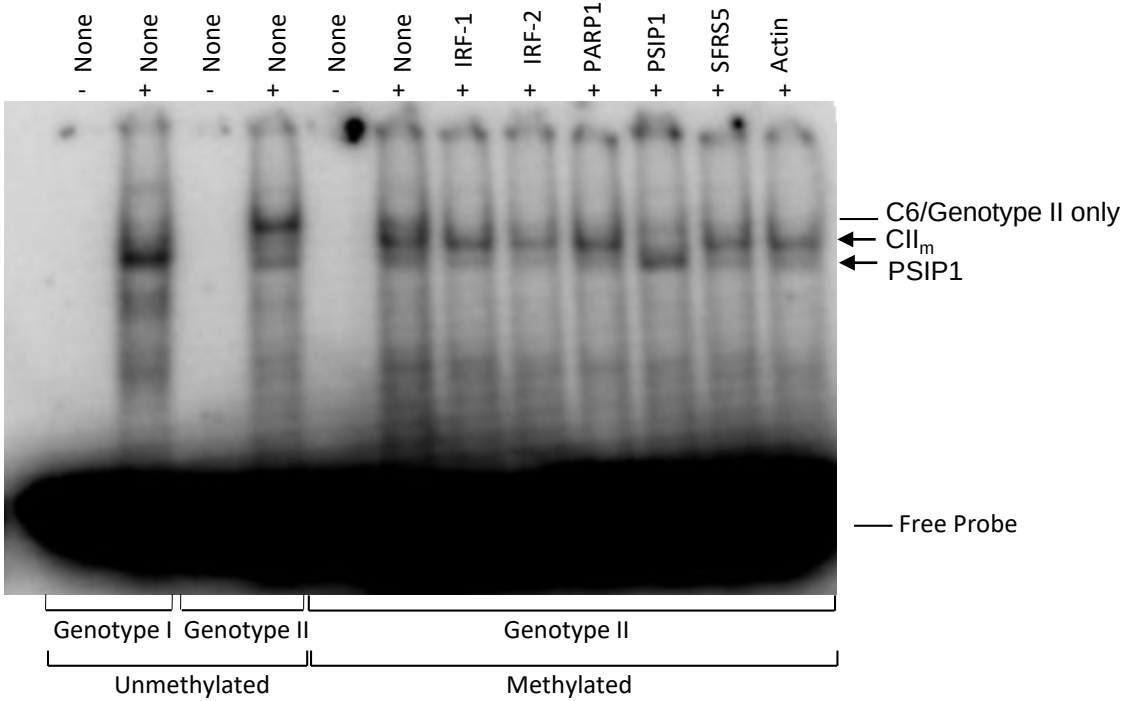


Figure S12

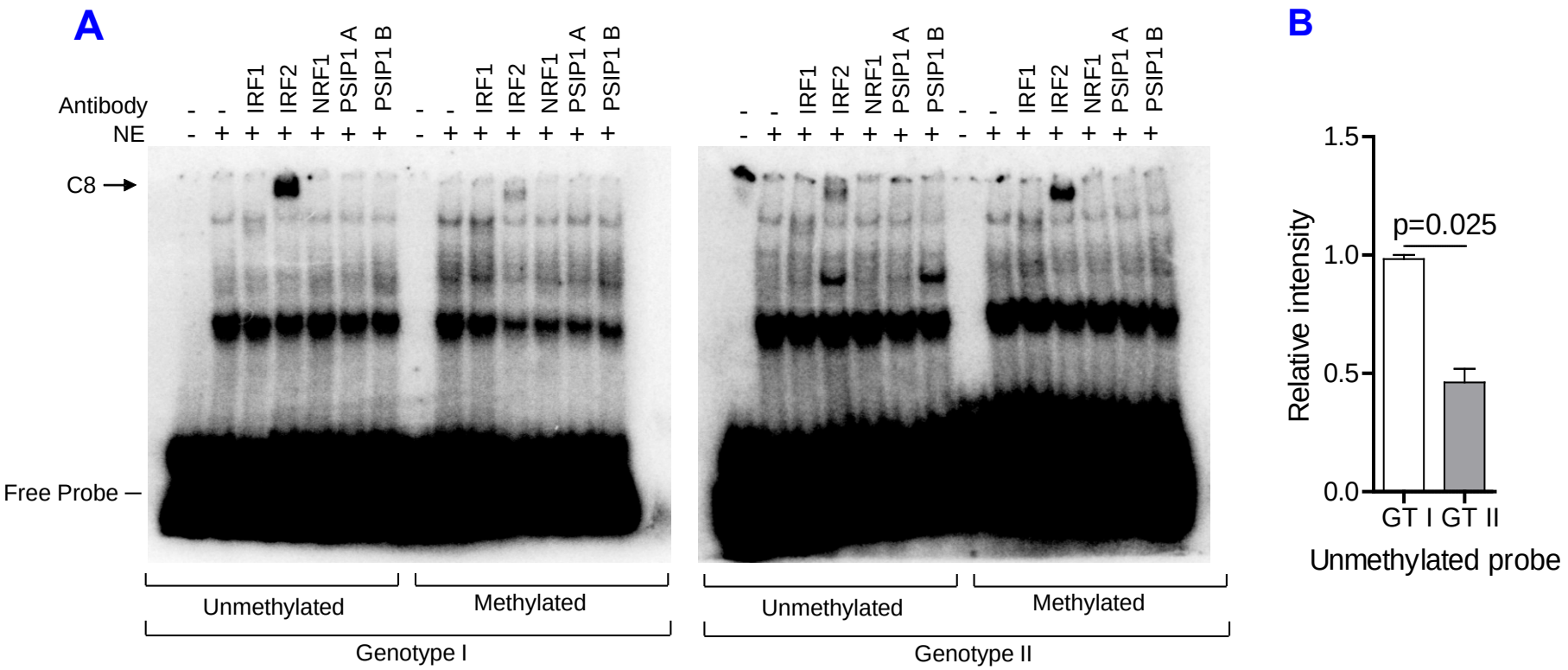
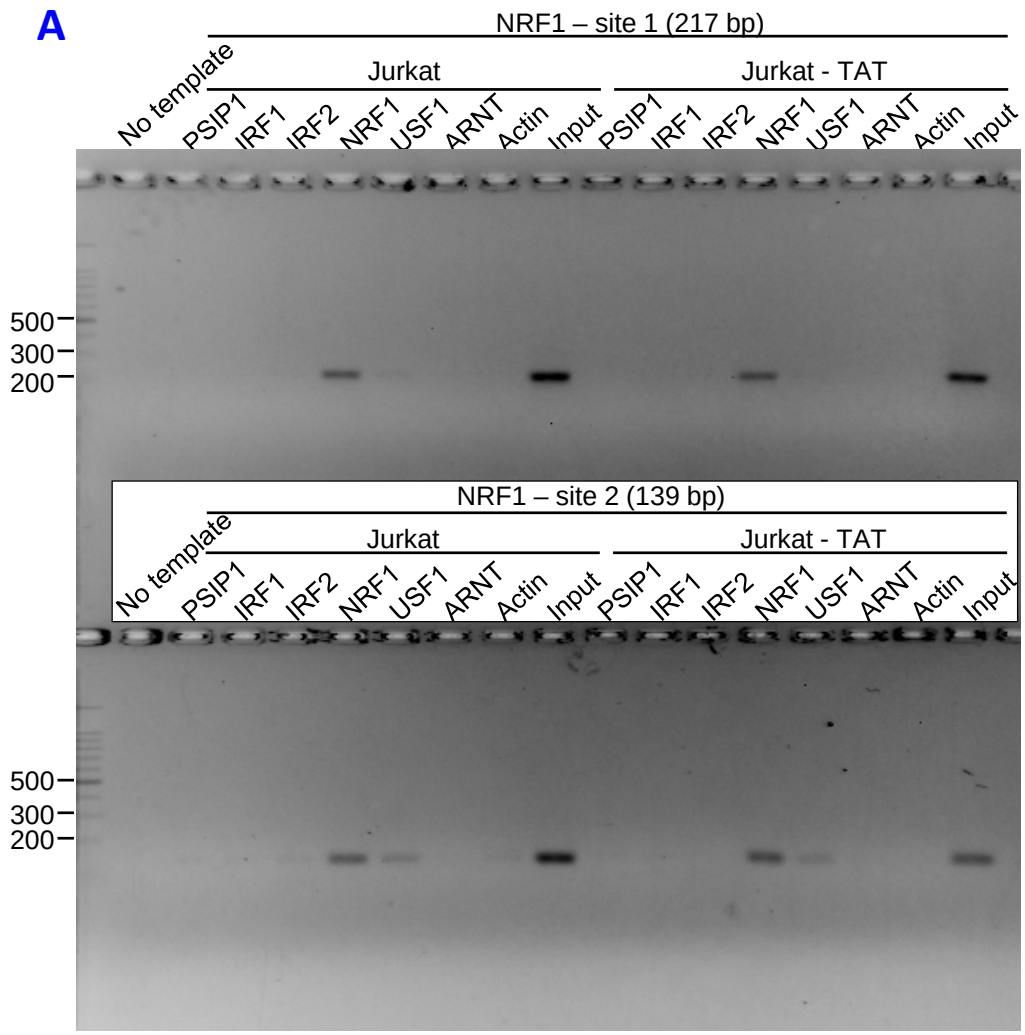
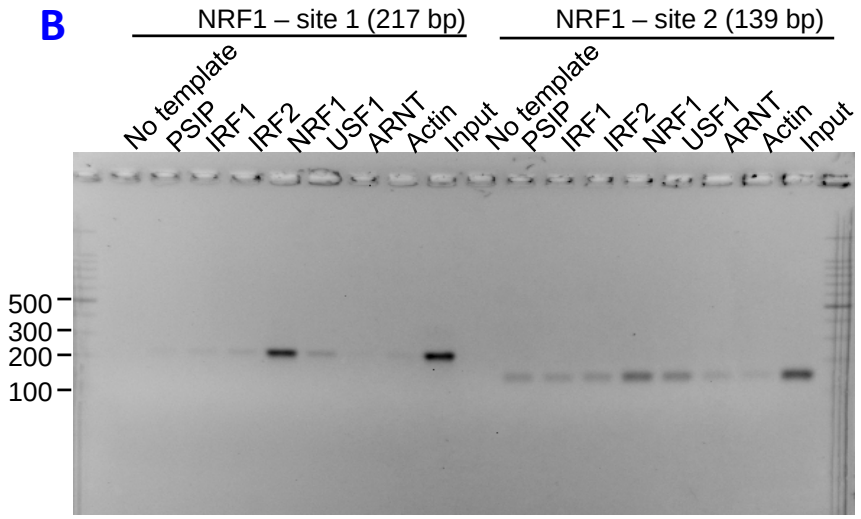

A
Antibody
NE
- - IRF1 IRF2 NRF1 PSIP1 A PSIP1 B - - IRF1 IRF2 NRF1 PSIP1 A PSIP1 B
- + + + + + + - + + + + + +
C8
Free Probe
Unmethylated
Methylated
Genotype I
Unmethylated
Methylated
Genotype II
B
Relative intensity
1.5
1.0
0.5
0.0
p=0.025
GT I GT II
Unmethylated probe

Figure S13


A
NRF1 – site 1 (217 bp)
Jurkat
Jurkat - TAT
No template
PSIP1
IRF1
IRF2
NRF1
USF1
ARNT
Actin
Input
500
300
200
NRF1 – site 2 (139 bp)


B
NRF1 – site 1 (217 bp)
NRF1 – site 2 (139 bp)
No template
PSIP
IRF1
IRF2
NRF1
USF1
ARNT
Actin
Input
500
300
200
100

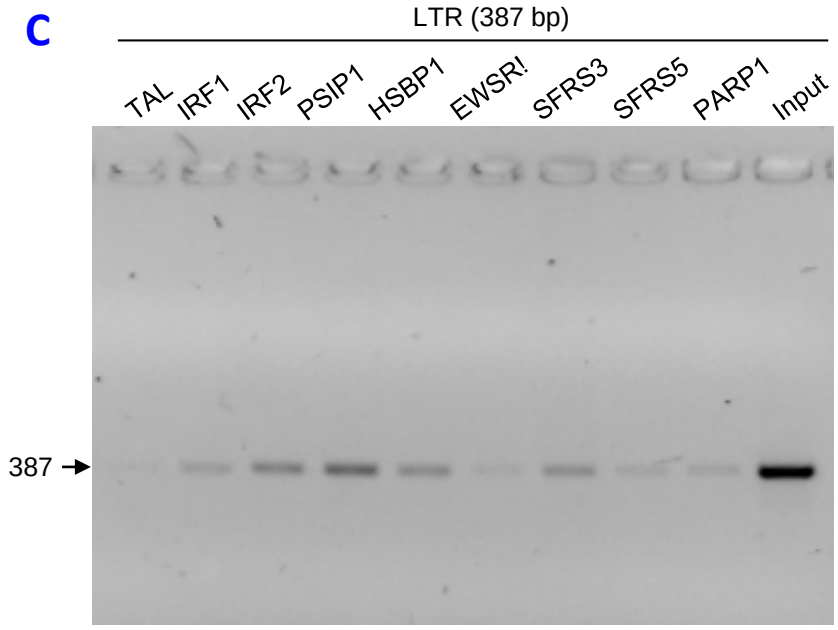

C
LTR (387 bp)
TAL
IRF1
IRF2
PSIP1
HSBP1
EWSR!
SFRS3
SFRS5
PARP1
Input
387

Figure S14

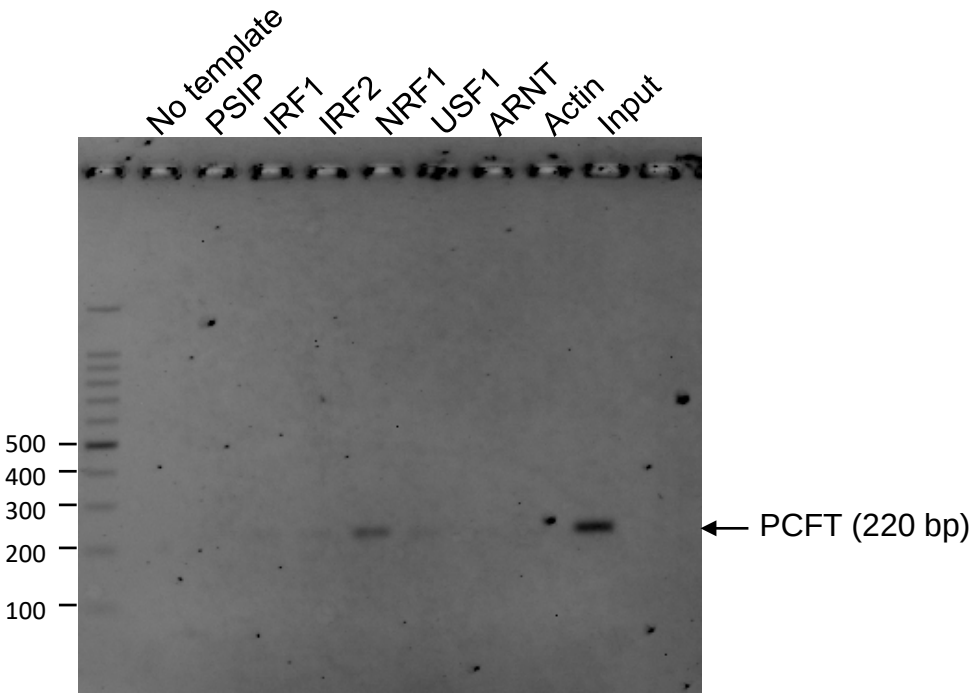
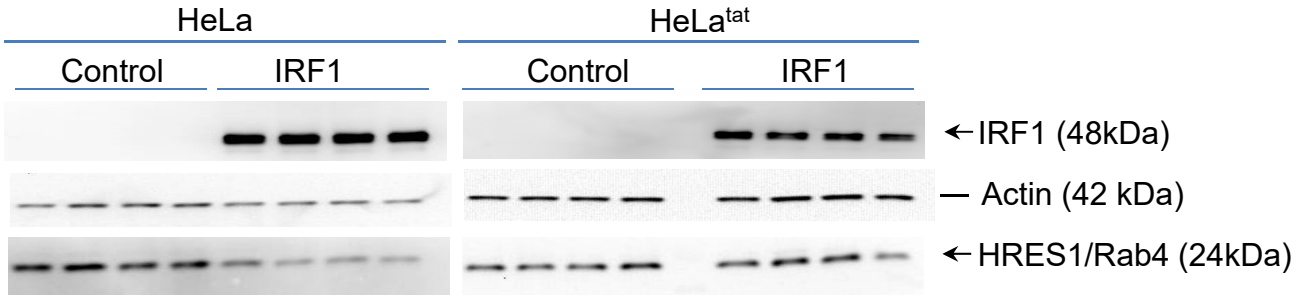

No template
PSIP
IRF1
IRF2
NRF1
USF1
ARNT
Actin
Input
500
400
300
200
100
PCFT (220 bp)

Figure S15


HeLa
HeLa^tat
Control
IRF1
Control
IRF1
←IRF1 (48kDa)
— Actin (42 kDa)
←HRES1/Rab4 (24kDa)

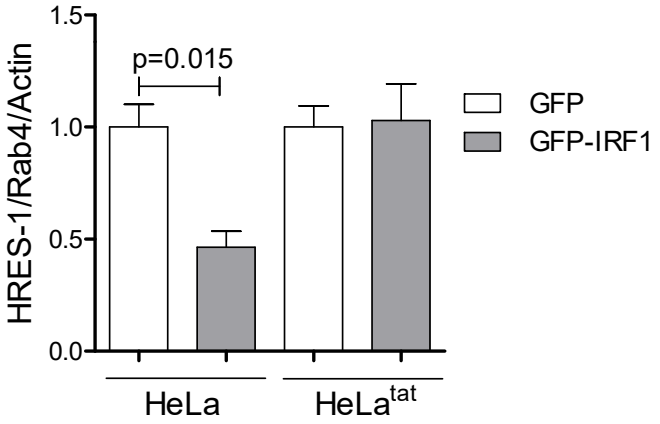

1.5
1.0
0.5
0.0
HRES-1/Rab4/Actin
p=0.015
GFP
GFP-IRF1
HeLa
HeLa^tat

Figure S16

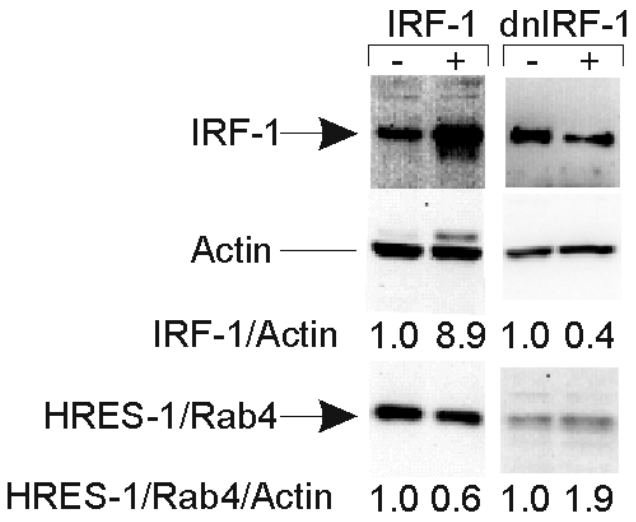

IRF-1
dnIRF-1
- + - +
IRF-1
Actin
IRF-1/Actin 1.0 8.9 1.0 0.4
HRES-1/Rab4
HRES-1/Rab4/Actin 1.0 0.6 1.0 1.9

Figure S17

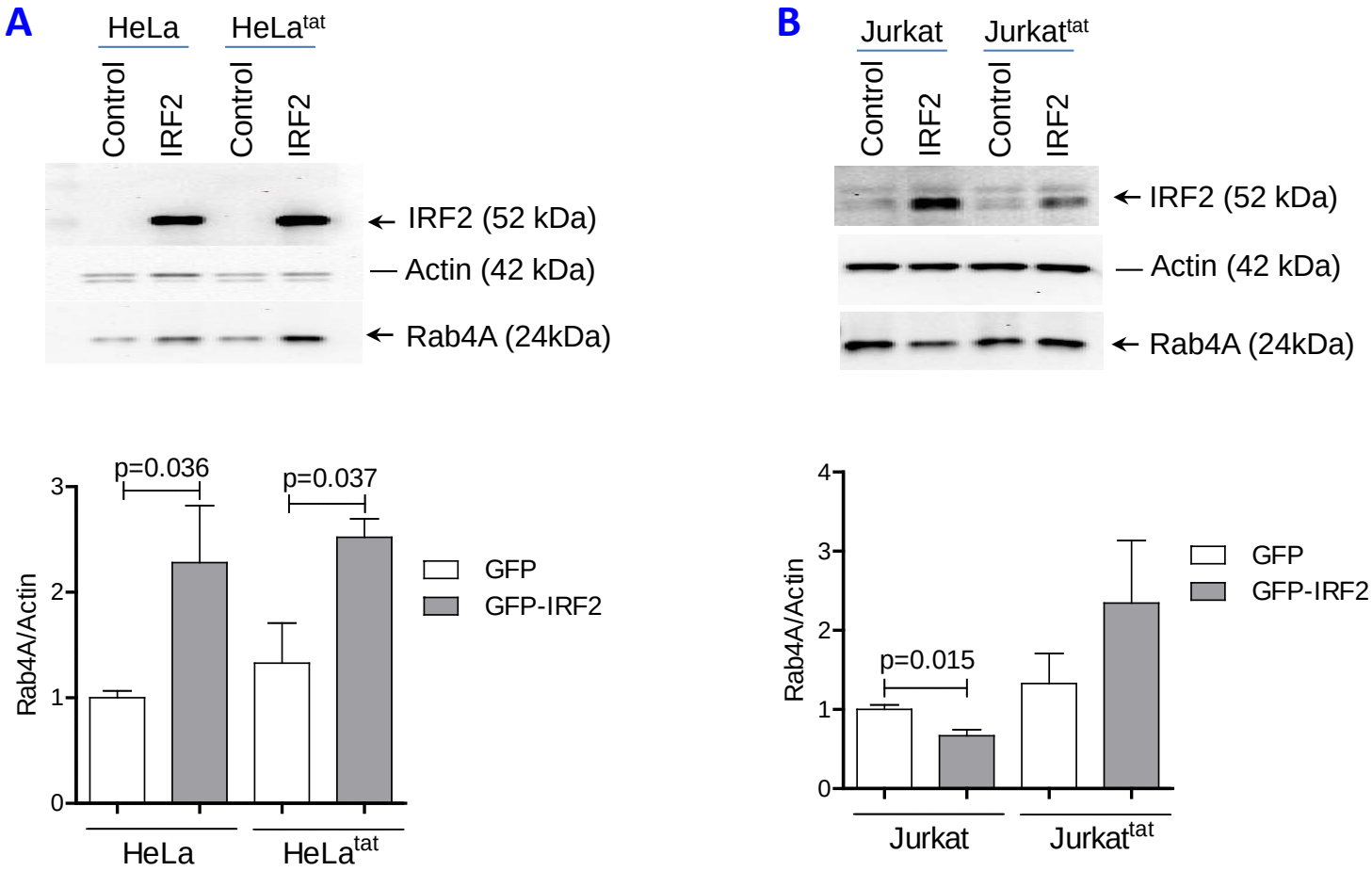

A
HeLa
HeLa^tat
Control
IRF2
Control
IRF2
← IRF2 (52 kDa)
— Actin (42 kDa)
← Rab4A (24kDa)
p=0.036
p=0.037
Rab4A/Actin
3
2
1
0
GFP
GFP-IRF2
HeLa
HeLa^tat
B
Jurkat
Jurkat^tat
Control
IRF2
Control
IRF2
← IRF2 (52 kDa)
— Actin (42 kDa)
← Rab4A (24kDa)
Rab4A/Actin
4
3
2
1
0
p=0.015
GFP
GFP-IRF2
Jurkat
Jurkat^tat

Figure S18

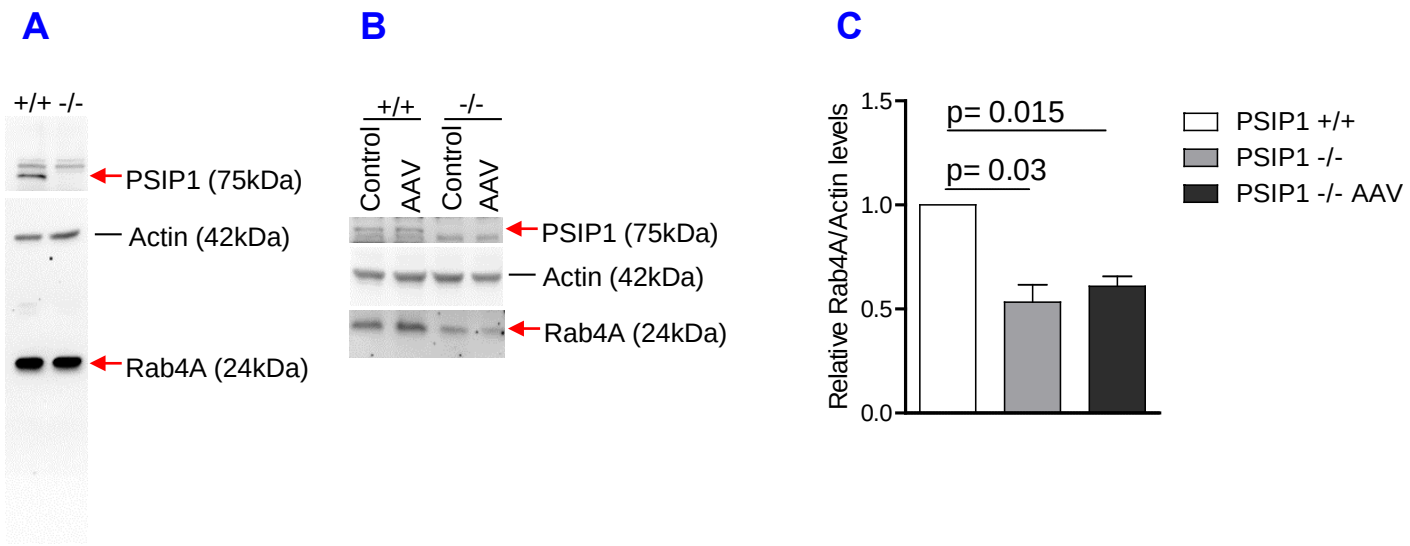
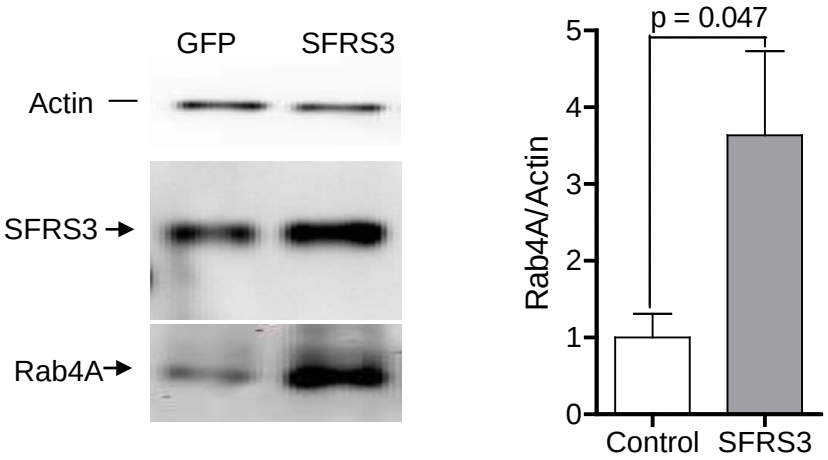

A
+/+ -/-
PSIP1 (75kDa)
Actin (42kDa)
Rab4A (24kDa)
B
+/+
-/-
Control
AAV
Control
AAV
PSIP1 (75kDa)
Actin (42kDa)
Rab4A (24kDa)
C
Relative Rab4A/Actin levels
1.5
1.0
0.5
0.0
p= 0.015
p= 0.03
PSIP1 +/+
PSIP1 -/-
PSIP1 -/- AAV

Figure S19


GFP
SFRS3
Actin
SFRS3
Rab4A

p = 0.047
Rab4A/Actin
5
4
3
2
1
0
Control
SFRS3

Figure S20

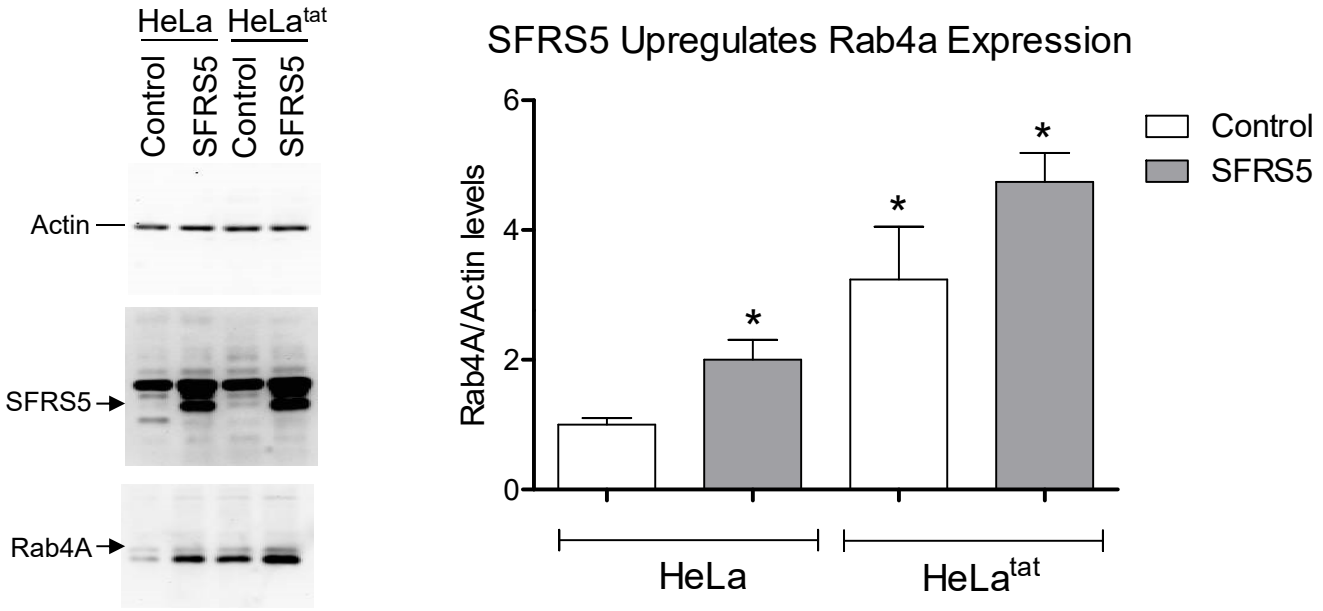
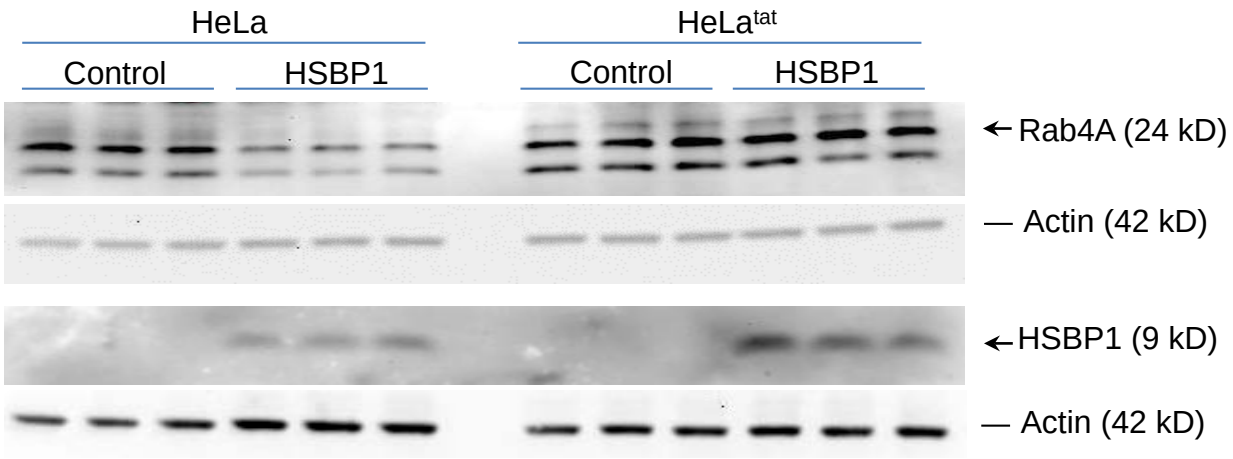

HeLa
HeLa^tat
Control
SFRS5
Control
SFRS5
Actin
SFRS5
Rab4A
SFRS5 Upregulates Rab4a Expression
Rab4A/Actin levels
6
4
2
0
*
*
*
Control
SFRS5
HeLa
HeLa^tat

Figure S21


HeLa
HeLa^tat
Control
HSBP1
Control
HSBP1
←Rab4A (24 kD)
— Actin (42 kD)
←HSBP1 (9 kD)
— Actin (42 kD)

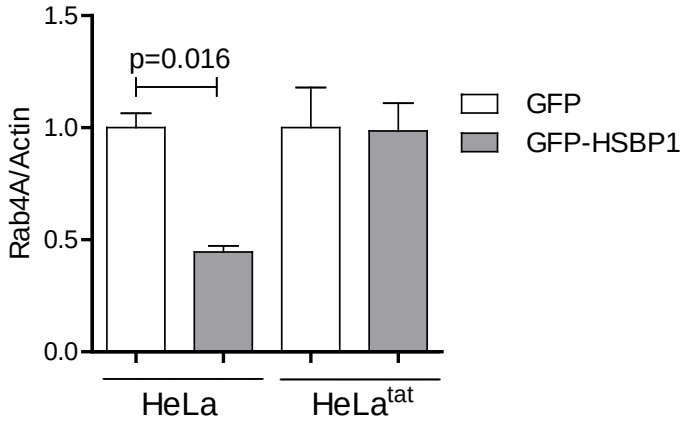
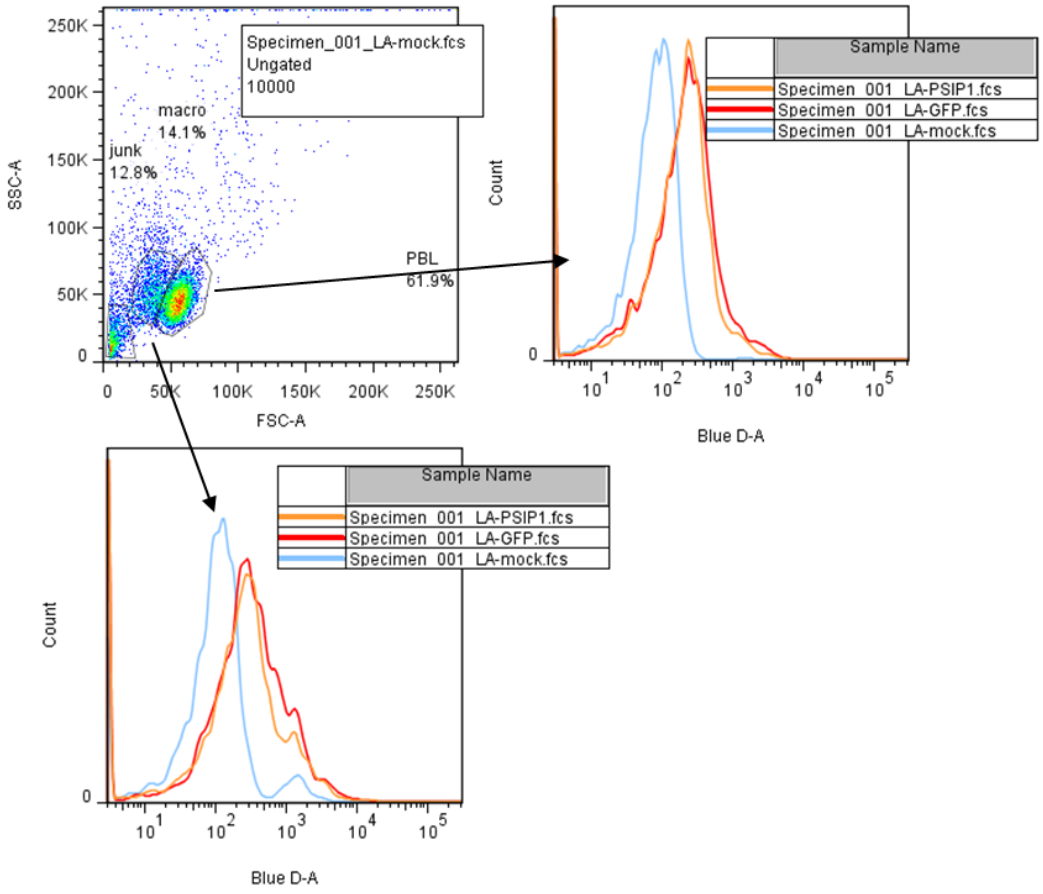

1.5
1.0
0.5
0.0
Rab4A/Actin
p=0.016
GFP
GFP-HSBP1
HeLa
HeLa^tat

Figure S22


Specimen_001_LA-mock.fcs
Ungated
10000
macro
14.1%
junk
12.8%
PBL
61.9%
SSC-A
FSC-A
250K
200K
150K
100K
50K
0
Count
Blue D-A
Sample Name
Specimen_001_LA-PSIP1.fcs
Specimen_001_LA-GFP.fcs
Specimen_001_LA-mock.fcs