Supplementary Information for

Human cytomegalovirus induces and exploits Roquin to counteract the IRF1-mediated antiviral state

Jaewon Song, Sanghyun Lee, Dongyeon Cho, Sungwon Lee, Hyewon Kim, Namhee Yu, Sanghyuk Lee, Kwangseog Ahn

Corresponding author:
Kwangseog Ahn
Email: ksahn@snu.ac.kr

This PDF file includes:

Supplementary text
Figs. S1 to S5
Table S1
Captions for Datasets S1 to S3
References for SI reference citations

Other supplementary materials for this manuscript include the following:

Datasets S1 to S3

www.pnas.org/cgi/doi/10.1073/pnas.1909314116

Supplementary Information Text

SI Materials and Methods

qPCR

Cells were harvested, and RNA was isolated using TRIzol reagent (Invitrogen). cDNA was synthesized using a ReverTra Ace kit (Toyobo). Viral genomic DNA was isolated using a QIAamp DNA blood mini kit (Qiagen). Real-time PCR was performed using a TOPreal SYBR Green PCR kit (Enzynomics) with primers listed in Table S1. mRNA levels were normalized against GAPDH mRNA.

Immunoblotting

Cells were lysed using RIPA buffer [50 mM Tris (pH 7.4), 150 mM sodium chloride, 0.5 % sodium deoxycholate, 0.1% SDS, and 1.0% NP-40] supplemented with 10 uM Leupeptin (Sigma-Aldrich) and 1 mM phenylmethanesulfonyl fluoride (PMSF; Sigma-Aldrich) and boiled with SDS sample buffer, and protein samples were subjected to SDS-PAGE. The following antibodies were used: IE1/2 (MAB810R; Millipore), IE2 (sc-69835; Santa Cruz Biotechnology), UL44 (CA006-1; Virusys), pp28 (CA004-1; Virusys), pp65 (P1205; Virusys), Roquin (A300-514a and A300-515a; Bethyl Laboratories), IRF1 (D5E4; Cell Signaling Technology), Regnase-1 (MAB7875-SP; R&D Systems), phospho-TBK1 (#5483; Cell Signaling Technology), phospho-IRF3 S386 (ab76493; Abcam), phospho-STAT1 (#9167; Cell Signaling Technology), GAPDH (AbFrontier), peroxidase-conjugated anti-mouse IgG (115-035-062) and anti-rabbit IgG (111-035-003; Jackson Laboratories). Signals were detected using a SuperSignal West Pico chemiluminescence kit (Thermo Fisher Scientific).

RNA-seq library generation

RNA-seq and CLIP-seq were performed,as previously described, with some modifications (1). For RNA-seq, primary HFFs were infected with HCMV Toledo at an MOI of 3, and sequencing was performed in biological duplicates at four time points of infection (0, 6, 24, and 72 hpi) for sequencing in Roquin-knockdown cells and at three time points of infection (0, 24, and 72 hpi) for reference for CLIP-seq. Total RNA was extracted using TRIzol reagent (Invitrogen) according to manufacturer instructions. For RNA-seq in Roquin-knockdown cells, polyadenylated RNAs were enriched using Dynabeads mRNA direct purification kit (Ambion). For CLIP reference sequencing, ribosomal RNAs were removed using a Ribo-Zero rRNA removal kit (Epicentre). After RNA fragmentation in RNA fragmentation buffer (New England Biolabs), RNAs of 30 to 60 nucleotides in length were purified and ligated with 3′ adaptors (5′-rApp-TGGAATTCTCGGGTGCCAAGG-ddC-3′; Integrated DNA Technologies) using T4 RNA ligase 2 and truncated K227Q (New England Biolabs), and 5′ adaptors (5′-GrGrUrUrCrArGrArGrUrUrCrUrArCrArGrUrCrCrGrArCrGrArUrC-3′; Integrated DNA Technologies) using T4 RNA ligase (Takara).

CLIP-seq was performed at three time points of infection (0, 24, and 72 hpi). Primary HFFs (1×10^8 cells) were irradiated for RNA-protein cross-liking with 254 nm UV at 300 mJ/cm^2 using a Spectrolinker (Spectroline) for Roquin CLIP. The RNAs bound to Roquin were immunoprecipitated with beads conjugated with two different anti-Roquin antibodies (A300-514a and A300-515a; Bethyl Laboratories). After ligation with the 3′ adaptors described for the RNA-seq library, RNAs were labeled with [γ-^{32}P]ATP using the T4 polynucleotide kinase (Takara). The RNA–protein complex was separated by SDS-PAGE and transferred to a nitrocellulose membrane (Whatman), after which the RNA was extracted with phenol/chloroform (Ambion), followed by ethanol precipitation for RNA isolation. The 5′ adaptors described for use in the RNA-seq library were subsequently ligated.

RNAs harboring the ligated 5′ and 3′ adaptors were reverse transcribed using the RNA RT primer (5′-GCCTTGGCACCCGAGAATTCCA-3′; Integrated DNA Technologies). PCR was performed to generate libraries for high-throughput sequencing with the 5′-end Illumina RNA PCR Primer (RP1) and the 3′-end Illumina RNA PCR Primer with index sequences (Indexes 1–9). Sequencing was performed on a HiSeq2500 system (Illumina).

RNA-seq analysis

Analysis of the sequencing data was performed as previously described, with slight modifications (2). For preprocessing, we removed the adapter sequences and low-quality ends from RNA-seq and CLIP-seq reads using Cutadapt version 1.10 (command-line parameters: -m 17 --match-read-wildcards -O 10 -e 0.1 -q 30,30 -g AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATC -a TGGAATTCTCGGGTGCCAAGGAACTCCAGTCAC). Artifact reads were eliminated by the fastx_artifacts_filter command in the FASTX-Toolkit (v0.0.13.2; http://hannonlab.cshl.edu/fastx_toolkit/). Additionally, we discarded reads mapped to human rRNA or tRNA by Bowtie2 version 2.2.7 (command-line parameters: -t - k 2–very-sensitive). Tophat version 2.1.1 [command-line parameters: with–no-coverage-search–b2-very-sensitive (−b2- score-min L, −0.6, −0.9 only for CLIP-seq due to the high error rate)] was used to align these preprocessed sequencing reads against the human reference genome (GRCh37.p13) and the HCMV Toledo genome (GU937742.2). We applied Fisher's exact test to detect significant peaks (CLIP-seq enriched regions over reference seq), as previously described (1), except that we calculated p-values for every genomic position in the whole genome background. The p-values were adjusted using the qvalue package in R (https://www.r-project.org/). RSEM (3) version 1.2.30 was used to align sequencing reads against all transcripts of the human reference genome (GRCh37.p13) and estimate gene-level transcript abundance. Differentially expressed genes were assessed using the limma package in R (4). We first filtered genes with total read counts <12 for each sample. We applied trimmed mean of M values (TMM) normalization to read counts for the estimation of scale factors among samples. The voom transformation was applied to the filtered and normalized counts, and the usual limma procedure for differential-expression analysis was followed. We estimated the fold changes and standard errors by fitting a linear model for each gene, and applied empirical Bayes smoothing to the standard errors. Moderated t statistics and corresponding p-values were computed, with the Benjamini–Hochberg method used to adjust p-values for multiple testing.

CLIP reads mapped on the genome were visualized using Integrative Genomics Viewer (5). Heatmaps and cumulative-distribution plots were created using the ggplot2 R package (6), and p-values for cumulative distributions were determined by two-sided Mann–Whitney *U* test.

GSEA

Based on the fold changes in expression of target genes between Roquin-knockdown and control cells, we generated a ranked list for GSEA (7), which was performed using REACTOME pathways or GO terms. The top 10 significant terms were selected and demonstrated with p-value and the number of genes included.

Stem-loop enrichment analysis

To predict the secondary structure associated with each peak from Roquin CLIP-seq and CLIP-seq analyses from a previous study (1), we used RNAfold from the Vienna RNA package (8). According to the predicted folding results, the number of each stem-loop-stem motifs were counted for the indicated datasets and compared with the number of the same motifs retrieved from 100 dinucleotide shuffles of the peak sequences using Dishuffle (9) in order to determine fold enrichment. The nucleotide composition of each stem-loop was determined using WebLogo (10).

Multiplex cytokine assay

To measure the protein levels of secreted cytokines, HFFs were transfected with siCon or siRoq and infected with HCMV at 2 MOI. Cell-culture supernatants harvested at 0, 24, or 72 hpi were centrifuged to remove cell debris and analyzed with the Bio-Plex 200 system (Bio-Rad) at the Seoul National University Hospital Biomedical Research Institute (Seoul, Korea) using a customized assay panel.

Conditioned media experiment

HFF cells (donor cells) were treated with siCon or siRoq and infected with HCMV at 2 MOI. Cell culture supernatant (conditioned media; CM) was harvested at 72 hpi, centrifuged to remove cell debris, and UV-irradiated to inactivate the virus in the media. Naïve HFF cells (recipient cells) were incubated with CM for 12 h, infected with HCMV (MOI = 0.5 or 2), and incubated again with CM until harvested at 72 hpi. UL99 mRNA level was measured by qRT-PCR and normalized to GAPDH mRNA level.

RNA IP

HFFs were infected with HCMV Toledo at 2 MOI and harvested at 0, 24, and 72 hpi. Cells were added along with 400 µL of RNA IP lysis buffer [20 mM Tris (pH 7.4), 150 mM NaCl, 1.5 mM $MgCl_2$, and 1 mM DTT supplemented with 10 uM Leupeptin, 1 mM PMSF, and RNase inhibitor (Enzynomics)]. Lysate was incubated with Dynabeads Protein A (40 µL; Invitrogen) conjugated with IgG control or Roquin antibodies at 4℃. After washing with RNA IP lysis buffer, Roquin-bound RNAs were isolated using TRIzol reagent (Invitrogen) and used for qRT-PCR.

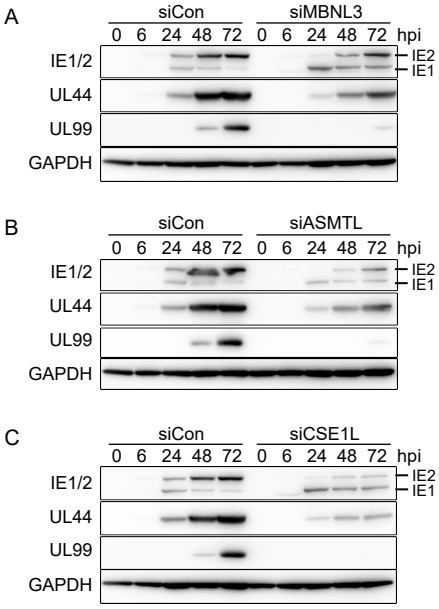
ASO transfection

ASOs complementary to the IRF1 5′ UTR were synthesized with 2′-O-methyl RNA harboring three phosphorothioate linkages on the 5′ and 3′ ends (Integrated DNA Technologies). HFF cells were seeded 1 day before transfection of ASOs at a final concentration of between ~0.1 µM and 2 µM using Dharmafect-1 reagent (Dharmacon). The following sequences were used for each ASO: ASO-5'UTR, 5'-GUUGCCGGGUUCUUAAG-3'; ASO-Stem-loop, 5'-GGGCUGCAGUGAGGGCG-3'.

Statistical analysis

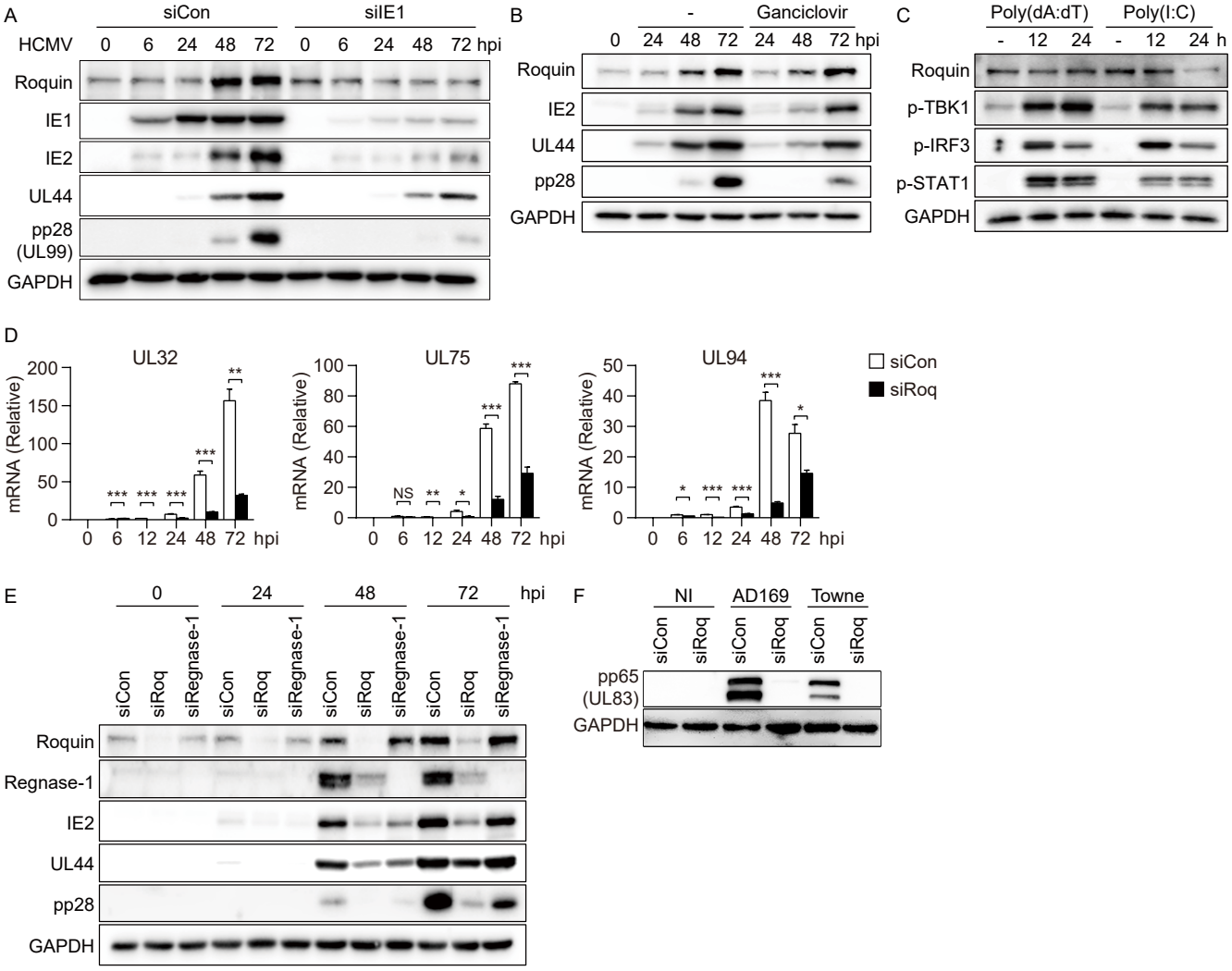
All analyses except for RNA-seq data and cumulative-distribution plot were performed using Graphpad Prism v.7. An unpaired two-tailed *t*-test or Analysis of variance (ANOVA) test was used to evaluate differences between two groups.

Figure S1. HCMV Requires Multiple RNA-binding Proteins for Efficient Lytic Replication



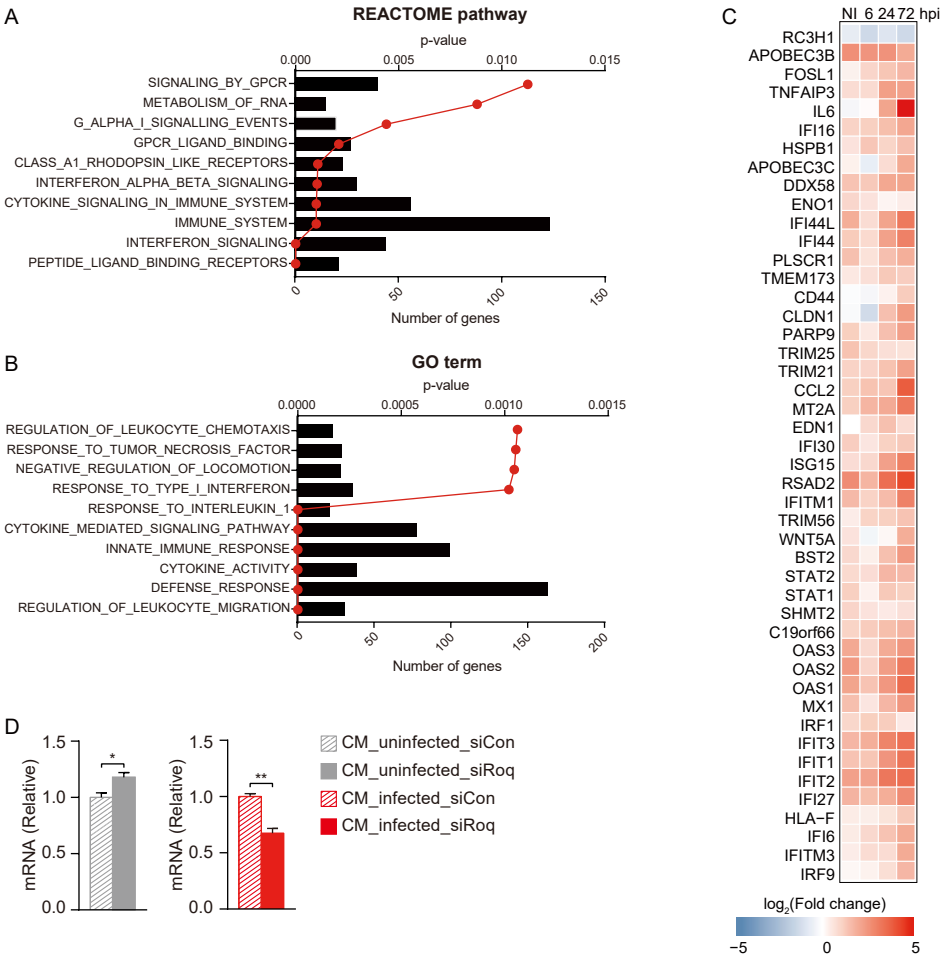
(A–C) HFF cells were treated with siCon or siRNA targeting MBNL3, ASMTL, or CSE1L and infected with HCMV (MOI = 2). Cells were harvested at the indicated time points, and the protein levels of viral genes were measured by immunoblot.

Figure S2. Roquin expression is increased during viral infection and is required for viral gene expression



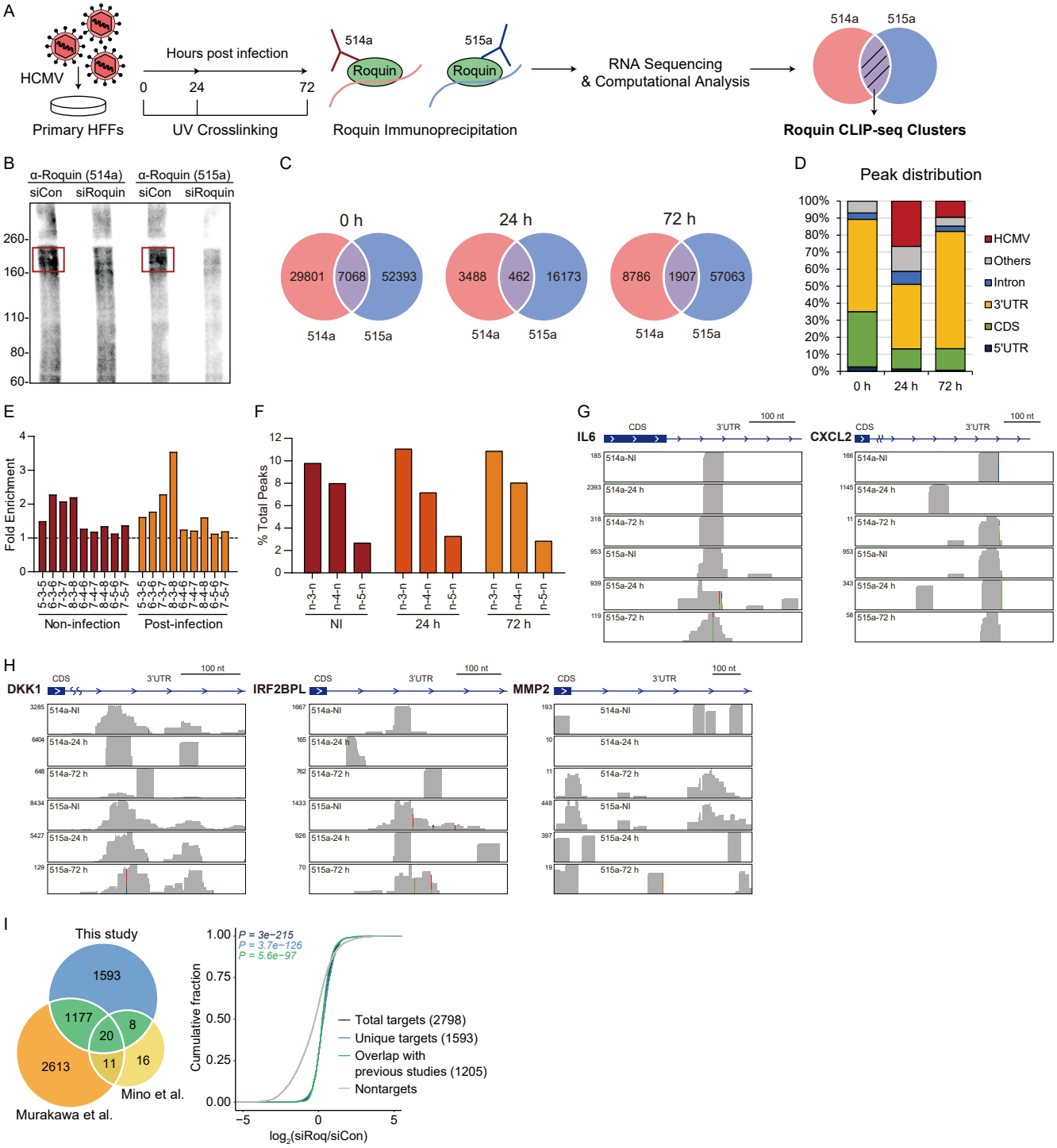
(A) HFF cells were treated with siCon or siIE1, then infected with HCMV Toledo (MOI = 2) and harvested at the indicated time points for immunoblot. (B) HFF cells were infectd with HCMV (MOI = 2), administered with 5 μM ganciclovir at 1 hpi, and incubated until harvested at indicated time points. (C) HFF cells were transfected with poly(dA:dT) or poly(I:C) and harvested at 12 or 24 h post-transfection for immunoblot. (D) HFF cells were treated with siCon or siRoq, infected with HCMV (MOI = 2), and harvested at the indicated time points. mRNA levels of viral genes UL32, UL75, and UL94 were measured by qRT-PCR. mean ± SEM, n = 3; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$ according to two-tailed Student's -- test; NS, not significant. (E) HFF cells were treated with siCon, siRoq, or siRegnase-1, infected with HCMV Toledo (MOI = 2), and harvested at indicated time points for immunoblot. (F) HFF cells were treated with siCon or siRoq and infected with the HCMV AD169 or HCMV Towne strain (MOI = 2). At 72 hpi, viral protein level was measured by immunoblot.

Figure S3. Roquin Regulates Cellular Genes Involved in Immune Responses



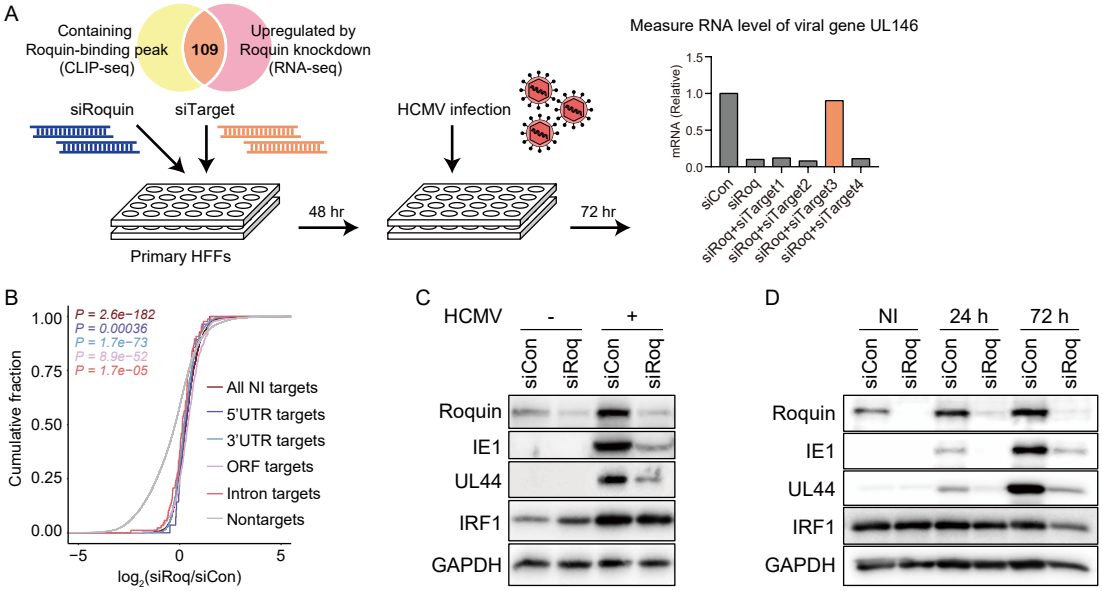
(A, B) GSEA based on the REACTOME pathway or GO term was performed for upregulated differentially expressed genes identified by RNA-seq in cells treated with siCon or siRoq. (C) Heatmap of fold changes [$\log_2$(siRoq/siCon)] in immune-related gene expression. (D) Naïve HFFs were incubated with CM which was prepared as described in Fig. 3D. Then the cells were infected with HCMV at 2 MOI and harvested at 72 hpi. UL99 mRNA level was measured by qRT-PCR. mean ± SEM, n = 3; *$P < 0.05$; **$P < 0.01$ according to two-tailed Student's *t* test.

Figure S4. Analysis of Roquin CLIP-seq Data



(A) CLIP-seq data preprocessing and peak-call scheme. (B) Phosphorimage of SDS-PAGE of ^{32}P-labed RNA-Roquin IP compelxes. (C) Venn diagram showing the overlap of Roquin-binding sites between two antibodies. (D) Distribution of Roquin-binding clusters between the human and HCMV genomes. (E) Enrichment of stem-loop structures in the Roquin-binding clusters from non-infected or infected (24 h and 72 h) CLIP peak clusters. (F) Percent ratio of Roquin-binding clusters containing stem-loop structures with 3, 4, or 5-nt loop. (G, H) Roquin CLIP reads mapped on the indicated genes. (I) Venn diagram comparing the RNA-seq data from this study with those from two previous studies, and the cumulative-distribution function plot of the indicated Roquin target genes. Log_2 fold changes were calculated using RNA-seq data from uninfected cells.

Figure S5. Identification of Roquin Target Genes Involved in HCMV Control



(A) Experimental scheme of siRNA screening. HFFs were treated with siRNA targeting each gene in the library along with siRoq. At 2-days post-transfection, cells were infected with HCMV (MOI = 2), transfected with another set of siRNAs, and the mRNA levels of UL146 were measured at 72 hpi and normalized against GAPDH mRNA level. (B) Cumulative-distribution function plot of fold-changes in CLIP target gene expression following Roquin silencing according to the Roquin-binding site at the indicated region. (C) PMA-differentiated THP-1 cells were treated with siCon or siRoq, infected with HCMV (MOI = 10), and harvested at 48 hpi for immunoblot. (D) U373MG cells were treated with siCon or siRoq, infected with HCMV (MOI = 2), and harvested at indicated time points for immunoblot.

Table S1. List of Primer Sequences Used for Quantitative PCR.

	Forward	Reverse
IE1	CACGACGTTCCTGCAGACTA	TTTTCAGCATGTGCTCCTTG
IE2	AACCCCGAGAAAGATGTCCT	CCGGGGAGAGGAGTGTTAGT
UL44	CGTGTCGTGCTCCGTAACTA	AGCTGGAATTCACGGCCAAT
UL99	CGGGGGAAACGACAGTAGTA	CTGATGGTGGTGACGTTTTG
UL146	GGCCCGGATGCGATAAAAATG	TCGTCTCGGTCCTGGTGATT
RC3H1	ACACGGGAACTGAGTATGGAAA	GTAAGGTCCTCAGCCGGAAC
IRF1	TTTGTATCGGCCTGTGTGAATG	AAGCATGGCTGGGACATCA
IL6	CAAATTCGGTACATCCTCGACGGC	GGTTCAGGTTGTTTTCTGCCAGTGC
CSF3	GCTGTGCCACCCCGAGG	TGCAGGAGCCCCTGGTAGAGG
IL1B	TGCGAATCTCCGACCACCACTACA	TGGAGGTGGAGAGCTTTCAGTTCATAT
CXCL2	GGGCAGAAAGCTTGTCTCAA	GCTTCCTCCTTCCTTCTGGT
CCL2	AGGTGACTGGGGCATTGAT	GCCTCCAGCATGAAAGTCTC
IFNA1	ACCTGGTTCAACATGGAAATG	ACCAAGCTTCTTCACACTGCT
IFNB	AGTAGGCGACACTGTTCGTG	GCCTCCCATTCAATTGCCAC
MXA	AGGTCAGTTACCAGGACTAC	ATGGCATTCTGGGCTTTATT
RSAD2	AGGTTCTGCAAAGTAGAGTTGC	GATCAGGCTTCCATTGCTC
DKK1	CTCGGTTCTCAATTCCAACG	GCACTCCTCGTCCTCTG
IRF2BPL	CCCCAAAACATTCCGGATTC	AAGGGCACTGAACGAAATGC
MMP2	CTTCCAAGTCTGGAGCGATGT	TACCGTCAAAGGGGTATCCAT
GAPDH	ATCATCCCTGCCTCTACTGG	GTCAGGTCCACCACTGACAC

Additional data

Dataset S1. List of Genes in the siRNA Library (Separate file).
Genes screened for their effect on viral replication are listed. Genes from three categories are shown in separate worksheets.

Dataset S2. Differentially Expressed Genes Following Roquin Knockdown According to RNA-seq Data (Separate file).
Genes exhibiting a >2-fold increase or decrease in expression at any time point were defined as differentially expressed genes. Fold changes on a logarithmic scale are shown.

Dataset S3. Roquin Target Sites According to CLIP-seq Data (Separate file).
Roquin-related CLIP peak clusters from each time point are shown in separate worksheets.

References

1. Kim S, et al. (2015) Temporal landscape of microRNA-mediated host-virus crosstalk during productive human cytomegalovirus infection. Cell Host Microbe 17(6):838-851.
2. Oh C, et al. (2018) A central role for PI3K-AKT signaling pathway in linking SAMHD1-deficiency to the type I interferon signature. Sci. Rep. 8(1):84.
3. Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12:323.
4. Ritchie ME, et al. (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 43(7):e47.
5. Robinson JT, et al. (2011) Integrative genomics viewer. Nat. Biotechnol. 29(1):24-26.
6. Wickham H (2009) ggplot2: Elegant Graphics for Data Analysis (Springer-Verlag, New York).
7. Subramanian A, et al. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. U. S. A. 102(43):15545-15550.
8. Lorenz R, et al. (2011) ViennaRNA Package 2.0. Algorithms Mol. Biol. 6:26.
9. Clote P, Ferre F, Kranakis E, Krizanc D (2005) Structural RNA has lower folding energy than random RNA of the same dinucleotide frequency. RNA 11(5):578-591.
10. Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: a sequence logo generator. Genome Res. 14(6):1188-1190.