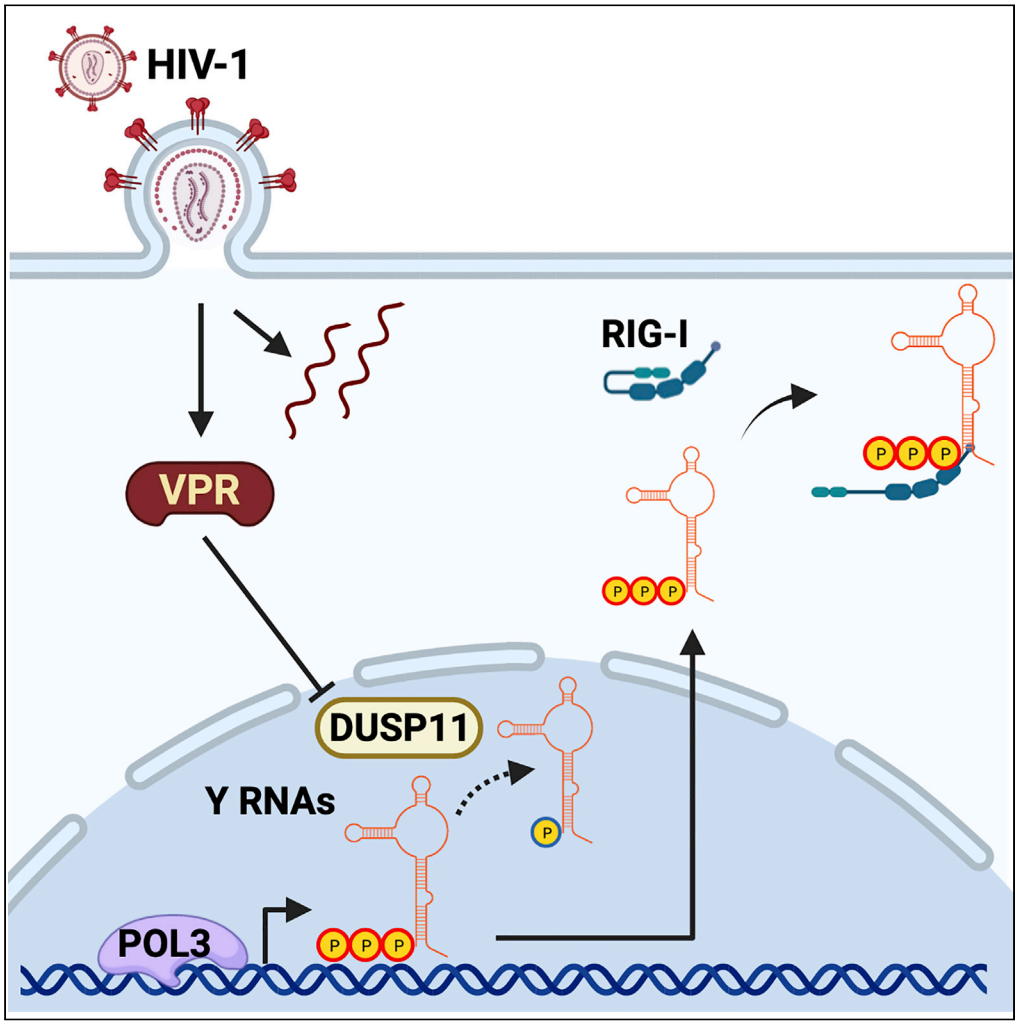
iScience

CellPress
OPEN ACCESS

Article

Y RNAs are conserved endogenous RIG-I ligands across RNA virus infection and are targeted by HIV-1



Nicolas Vabret, Valérie Najburg, Alexander Solovyov, ..., Nina Bhardwaj, Benjamin D. Greenbaum, Anastassia V. Komarova

nicolas.vabret@mssm.edu (N.V.)
anastasia.komarova@pasteur.fr (A.V.K.)

Highlights

Y RNAs and other POL3 transcripts are RLRs ligands during RNA virus infections

Y RNA immunogenicity depends on their 5′-PPP and viral-mimicry secondary structure

Development of a novel sequencing approach to measure 5′-PPP status of RNA

HIV-1 VPR triggers DUSP11 downregulation and modulates host 5′-PPP transcriptome

Vabret et al., iScience 25, 104599
July 15, 2022
https://doi.org/10.1016/j.isci.2022.104599

iScience

CellPress
OPEN ACCESS

Article

Y RNAs are conserved endogenous RIG-I ligands across RNA virus infection and are targeted by HIV-1

Nicolas Vabret,[1,2,10,15,*] Valérie Najburg,[3] Alexander Solovyov,[12] Ramya Gopal,[1,2,10] Christopher McClain,[1,2,10] Petr Šulc,[4] Sreekumar Balan,[1,2,10] Yannis Rahou,[5] Guillaume Beauclair,[3] Maxime Chazal,[3] Hugo Varet,[6,7] Rachel Legendre,[6,7] Odile Sismeiro,[6] Raul Y. Sanchez David,[3] Lise Chauveau,[8] Nolwenn Jouvenet,[3] Martin Markowitz,[9] Sylvie van der Werf,[5] Olivier Schwartz,[8] Frédéric Tangy,[3] Nina Bhardwaj,[1,2,10,11,14] Benjamin D. Greenbaum,[12,13,14] and Anastassia V. Komarova[3,5,14,15,*]

SUMMARY

Pattern recognition receptors (PRRs) protect against microbial invasion by detecting specific molecular patterns found in pathogens and initiating an immune response. Although microbial-derived PRR ligands have been extensively characterized, the contribution and relevance of endogenous ligands to PRR activation remains overlooked. Here, we characterize the landscape of endogenous ligands that engage RIG-I-like receptors (RLRs) upon infection by different RNA viruses. In each infection, several RNAs transcribed by RNA polymerase III (Pol3) specifically engaged RLRs, particularly the family of Y RNAs. Sensing of Y RNAs was dependent on their mimicking of viral secondary structure and their 5′-triphosphate extremity. Further, we found that HIV-1 triggered a VPR-dependent downregulation of RNA triphosphatase DUSP11 *in vitro* and *in vivo*, inducing a transcriptome-wide change of cellular RNA 5′-triphosphorylation that licenses Y RNA immunogenicity. Overall, our work uncovers the contribution of endogenous RNAs to antiviral immunity and demonstrates the importance of this pathway in HIV-1 infection.

INTRODUCTION

Pattern recognition receptors (PRRs) were initially described as innate immune sensors of molecular patterns commonly found in pathogens but rarely, if ever, found in their hosts. In recent years, this view has been challenged by evidence that ligands originating from self can engage these same PRRs. Notably, sensing of self-RNA by innate receptors has been observed in various settings such as autoimmune disorders (Hung et al., 2015; Lehmann et al., 2012; Zhao et al., 2018a), tumorigenesis (Choi et al., 2020; Ishizuka et al., 2019; Nabet et al., 2017; Tanne et al., 2015), cancer therapies (Chiappinelli et al., 2017; Heidegger et al., 2019; Leonova et al., 2013; Ranoa et al., 2016; Roulois et al., 2015), or infection by DNA viruses (Chiang et al., 2018; Zhao et al., 2018b). As the importance of endogenous ligands in priming immune responses is progressively uncovered, little is known about the breadth of biological processes in which they engage PRRs or their functional interplay and co-evolution with immune sensors. Furthermore, we lack a complete understanding of what features confer self-RNA the ability to activate innate sensors and whether this activation is a general response to aberrant transcription or is dominated by specific RNA species.

Further confounding matters, we previously determined that conventional RNA sequencing approaches fail to capture the full spectrum of RNA expression in tumors (Solovyov et al., 2018). In particular, we showed that repetitive RNAs, which can harbor immunostimulatory features (Vabret et al., 2017), require specific computational analysis for unbiased characterization of their transcription (Solovyov et al., 2018). Here, we apply these computational approaches to a novel sequencing platform to identify self-derived RNA ligands of RIG-I-like receptors (RLRs). RLRs are a family of ubiquitously expressed cytosolic RNA sensors composed of three members: RIG-I, LGP2, and MDA5, acting upstream of the signaling adapter protein MAVS in the initiation of expression and secretion of type I and type III interferons (IFN-I/III) and other proinflammatory cytokines (Chow et al., 2018). Their intracellular localization and proximity with host

[1]Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA

[2]Precision Immunology Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

[3]Viral Genomics and Vaccination Unit, Department of Virology, Institut Pasteur, Université de Paris, CNRS UMR-3569, 75015 Paris, France

[4]Center for Molecular Design and Biomimetics at the Biodesign Institute and School of Molecular Sciences, Arizona State University, Tempe, AZ 85287, USA

[5]Molecular Genetics of RNA Viruses, Department of Virology, Institut Pasteur, Université de Paris, CNRS UMR-3569, 75015 Paris, France

[6]Transcriptome and EpiGenome Platform, BioMics, Center of Innovation and Technological Research, Institut Pasteur, Université de Paris, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France

[7]Hub Informatique et Biostatistique, Centre de Bioinformatique, Biostatistique et Biologie Intégrative (C3BI, USR 3756 IP-CNRS), Institut Pasteur, Université de Paris, 28 Rue du Docteur Roux, 75724 Paris Cedex 15, France

[8]Virus & Immunity Unit, Department of Virology, Institut Pasteur, Université de Paris, CNRS UMR-3569, 75015 Paris, France

[9]Aaron Diamond AIDS Research Center, The

Continued

RNA implies a delicate balance between a need to develop high affinity for microbial features and the possibility to encounter self-RNAs that display similar structures. However, despite a growing knowledge of the role of RLRs during RNA virus infection and the microbial-derived ligands they recognize (Chow et al., 2018), the contribution of endogenous RNAs to their activation and the mechanisms controlling their immunogenicity remain elusive.

In this work, we study the specificity of host-derived RLR ligands across infections by multiple RNA viruses: measles, dengue, and HIV-1. We identify a conserved role for endogenous Y RNAs as RIG-I ligands, including upon HIV-1 infection where every RNA engaging RLRs are self-RNAs. We further found that this pathway is triggered by an HIV-1 viral protein R (VPR)-dependent downregulation of cellular dual-specificity protein phosphatase 11 (DUSP11), leading to a surge in cellular transcriptome triphosphorylation that licenses host RNAs immunogenicity.

RESULTS

Y RNAs and other RNA Pol3 transcripts are cellular RIG-I ligands mobilized upon RNA virus infection

We recently developed a riboproteomic approach based on tagged protein affinity purification that measures and compares receptor affinity of RNA molecules with improved statistical evaluation of specific binding (Chazal et al., 2018; Sanchez David et al., 2016). We use this approach to perform an unbiased quantification of RLR-bound self-RNAs during RNA virus infection. We used human HEK293 (293) cells stably expressing the 1-STrEP-tagged RLRs RIG-I, MDA5, or LGP2, or the protein Cherry as nonbinding control. We infected each cell line with either positive-sense RNA virus dengue virus 4 (DV-4) or negative-sense RNA virus measles virus (MV). As a model of retroviral infection, we co-cultivated HIV-1-infected MT4 T cells with 293 cells overexpressing HIV-1 receptors CD4 and CXCR4 (293-4x4), as cell-free HIV-1 particles are poor stimulators of IFN-I (Lepelley et al., 2011). We performed total RNA sequencing on each RLR- or Cherry-purified fraction and on total cellular RNA to compare specific viral RNA-binding profiles on RLR with nonspecific binding (Cherry) upon MV, DV-4, and HIV-1 infections (Figure S1A).

We previously reported viral RNA-binding profiles on RLRs compared to non-specific binding (Cherry) upon MV and DV-4 infection (Chazal et al., 2018; Sanchez David et al., 2016) (Figures S1B and S1C). However, in the case of HIV-1 infection, no enrichment of viral RNA was observed on any receptors (Figure S1D). Importantly, we also confirmed that the RLR-MAVS pathway was critical for sensing each viral infection in this model (Figures 1A and S1E). We then aligned RLR-bound RNAs to the human genome and measured specific cellular RNA enrichment in infected and noninfected conditions (NI). We found a strong enrichment of Pol3-transcribed RNAs to RIG-I and LGP2 during each RNA virus infection and in particular Y RNAs (Figures 1B, S1F, S1I, and S1L; Table S1). We observed that the enrichment of RNY4 to RIG-I was conserved across each RNA virus infection (Figure 1C), contrary to LGP2 or MDA5, where no shared enrichment of individual RNA was measured (Figures S1G and S1J). RIG-I binding of vault RNAs (vtRNAs), another family of Pol3-transcribed RNAs, has been previously observed during the lytic reactivation of Kaposi's sarcoma-associated herpesvirus (KSHV), a DNA virus (Zhao et al., 2018b). Interestingly, we detected a specific enrichment of vtRNA1-2 and vtRNA2-1 to RIG-I upon DV infection (Figure 1B).

Y RNAs constitute a family of highly conserved small noncoding RNAs composed of four canonical Y RNA (RNY1, RNY3-5) and several hundreds of repeat pseudogenes (Kowalski and Krude, 2015; Perreault et al., 2005; Wolin and Steitz, 1983). As multiple copies of Y RNAs can impair the identification of the exact origin of each transcript, we measured RLR enrichment of each Y RNA family rather than individual genes (Figures 1D, S1H, and S1K; Table S2). Specifically, the subfamily of HY4, which contains RNY4 and its pseudogenes, and to a lesser extent HY3, showed significant binding enrichment to RIG-I in the three RNA virus infections compared with NI. Hence, for all viruses tested in this study, self-RNAs, and in particular Y RNAs, played a common role in RIG-I engagement. Importantly, in the case of HIV, every RIG-I-bound RNA came from self, distinguishing it from other RNA viruses.

5′-PPP and viral-like specific secondary structures are required for RNY4 RIG-I agonism

To analyze the immunostimulatory properties of Y RNAs, we generated *in vitro* transcribed (IVT) molecules of each canonical Y RNA and measured IFN-I signaling after stimulation of individual RLR knockouts generated in the haploid cell line HAP1. Each individual Y RNA elicited an IFN-I response (induction of promoter ISRE and promoter IFN-β) after transfection, which was dependent on the presence of RIG-I and MAVS but

Rockefeller University, New York, NY, USA

[10]Department of Medicine, Hematology and Medical Oncology, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

[11]Extra-mural Member, Parker Institute of Cancer Immunotherapy, USA

[12]Computational Oncology, Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA

[13]Physiology, Biophysics, & Systems Biology, Weill Cornell Medicine, New York, NY 10065, USA

[14]Senior author

[15]Lead contact

*Correspondence: nicolas.vabret@mssm.edu (N.V.), anastasia.komarova@pasteur.fr (A.V.K.)

https://doi.org/10.1016/j.isci.2022.104599

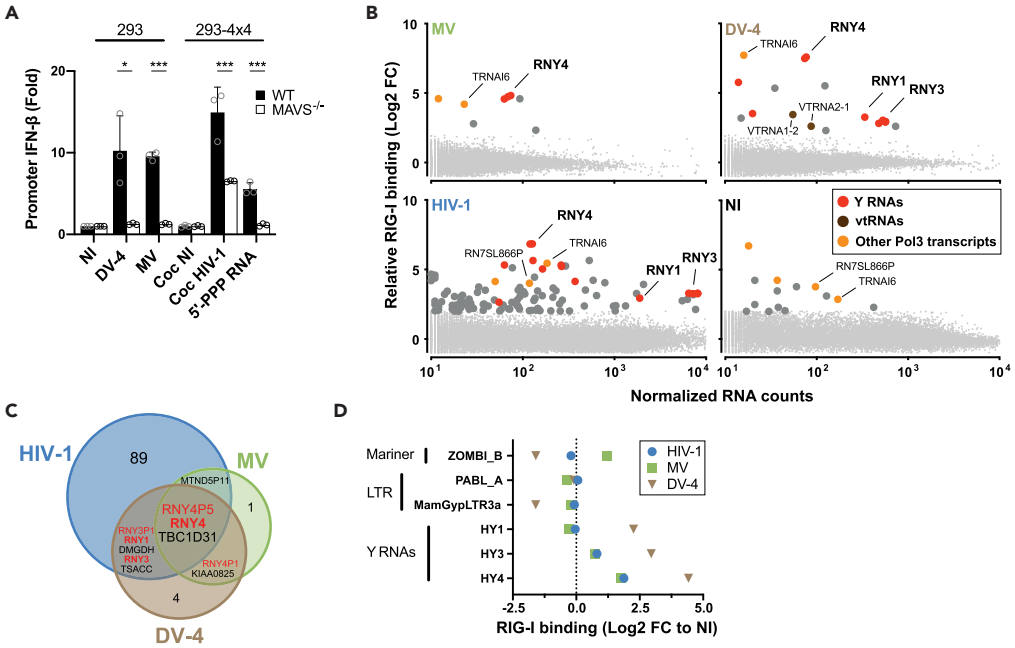


Figure 1. A differential affinity screen identifies Y RNAs and other POL3 RNAs as RIG-I ligands mobilizable during RNA virus infection

(A) Promoter IFN-β-luciferase reporter activity in WT or MAVS$^{-/-}$ (left) 293 cells infected with measles virus (MV) or dengue virus 4 (DV-4) at MOI of 1 and 0.5, respectively, or (right) 293-4x4 co-cultivated with HIV-1-infected MT4C5 at a ratio of MT4C5:293-4x4 of 1:1. 5′-PPP is a short *in vitro* transcribed RNA RIG-I agonist transfected at a concentration of 10 ng/mL.

(B) Twenty-four hours post-infection with MV, DV-4, after co-culture with HIV-1-infected MT4 or in noninfected (NI) control, sequencing reads were mapped to human genome Hg38. Differential analyses were performed between RIG-I/RNA and Cherry/RNA samples. Genes are represented following their normalized count in cellular RNA (x axis) and their fold enrichment (log2) to RIG-I compared with Cherry control (y axis) (average of three independent replicates). Genes that showed a log2(FC) > 2 and adj-pval<0.05 are represented with larger dot size. Among these, Y RNAs are labeled with red dots vtRNA in brown and other Pol3 transcripts are shown in orange. Canonical Y RNAs and Pol3 transcripts that show enrichment in more than two conditions (virus or NI) are specifically annotated.

(C) Venn diagram representing individual transcripts specifically enriched to RIG-I compared with Cherry in any of MV, DV-4, or HIV-1-infected conditions but absent in NI condition.

(D) Families of RNA repeats that show specific affinity to RIG-I compared with Cherry in at least one infected or NI condition, computed according to their relative enrichment compared with NI. The name of the repeat (e.g. *Zombi_B*) and the subfamily to which it belongs (e.g. *Mariner*) are indicated.

(A) Data representative of n = 3 independent experiments. Bars show mean ± SEM of technical triplicates.

Student's t test; *p < 0.05; ***p < 0.001. (B–D) Enrichment calculated from the mean of n = 3 infection/RLR-purification/sequencing experiments.

independent of MDA5 or to a large extent LGP2 (Figures 2A, S2A, and S2B). RIG-I recognizes RNA ligands based on a level of specificity in terms of sequence composition, length, double-stranded structures, and presence of 5′-triphosphate (5′-PPP) or diphosphate moieties (Chow et al., 2018). We generated fragments derived from RNY4 missing specific molecular substructures (Figures S2B and S2C). As shown in Figure 2B, both 5′-PPP and stem S3 were required to confer upon RNY4 its RIG-I-dependent immunostimulatory activity.

To validate these findings, we synthesized IVT RNY4 and RNY4ΔS3 RNAs using plasmids containing 3′ ribozyme sequences that generate discrete 3′ ends. We measured IFN-I response after transfection with these RNAs, confirming the difference observed between RNY4 and RNY4ΔS3 (Figures S2D and S2E). Finally, to confirm that endogenously transcribed Y RNAs can be immunostimulatory, we cloned the RNY4 sequence downstream of an RNA Pol3 promoter (U6) and used a Pol2 promoter (CMV) as control. Only endogenous transcription of RNY4 driven by RNA Pol3, but not driven by Pol2, elicited an IFN-I response dependent on the RIG-I/MAVS pathway (Figure 2C). To further understand the novel function of Y RNAs as a RIG-I ligands, we developed a secondary structure model of RNY4 and computed the

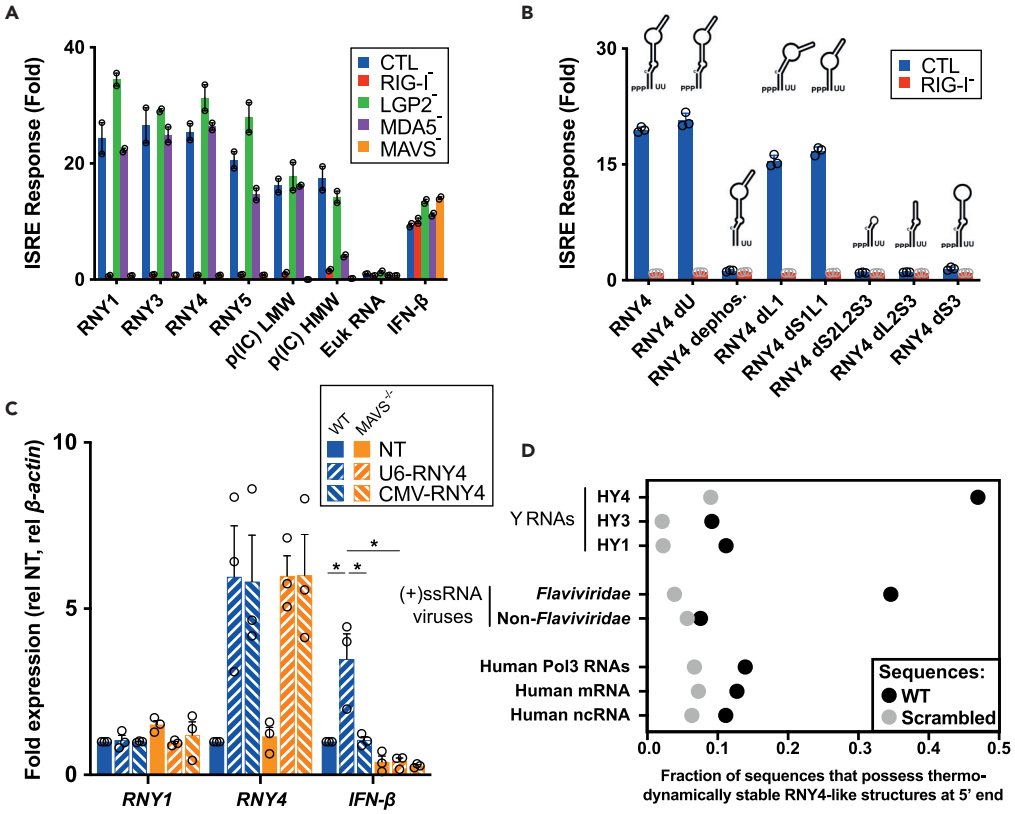


Figure 2. RNY4 RIG-I agonist activity is conferred by RNA 5′-PPP moieties and viral-mimicking specific secondary structure

(A) Promoter ISRE-luciferase reporter activity in HAP1 cells control (CTL) or knockout (ko) for each individual RLR or downstream adaptor MAVS, transfected with 30 ng/mL of IVT Y RNA (*RNY1*, *RNY3-5*), 10 ng/mL poly(I:C) low- or high-molecular-weight (p(IC) LMW/HMW), 30 ng/mL of total eukaryotic RNA, or treated with 100 U/mL recombinant IFN-β.

(B) Promoter ISRE-luciferase reporter activity in HAP1 cells CTL or RIG-I ko transfected with 30 ng/mL IVT RNY4 full length or lacking specific substructure (Figure S2C). RNY4 dephos: RNY4 was additionally pretreated with alkaline phosphatase to remove 5′- triphosphate extremity.

(C) *RNY1*, *RNY4*, and *IFN-β* RNA levels were measured by qPCR after transfection of WT or MAVS$^{-/-}$ 293T with plasmids coding for RNY4 sequence and supplemented with a plasmid coding for RIG-I. U6-RNY4: p2RZ plasmid encoding full-length RNY4 downstream of Pol3 U6 promoter with a ribozyme sequence placed directly in 3′. CMV-RNY4: same plasmid with Pol2 CMV promoter instead of U6 (Table S7). NT: empty plasmid.

(D) Probability of thermodynamically stable sequence folding along RNY4 secondary structure in the 5′ end of each transcript, for dataset of human Y RNA families, (+)ssRNA viruses genomes (*Flaviviridae* or non-*Flaviviridae*), or human noncoding RNA (ncRNA), mRNAs, and Pol3 transcripts, compared with average probability of the same sequences randomly scrambled.

(A–B) Data representative of n = 3 independent experiments. Bars show mean ± SEM of technical duplicates. (C) Bars show mean ± SEM of n = 3 independent experiments. Student's t test *p < 0.05.

probability of different sets of sequences of viral or human origins to fold along this model. Remarkably, RNY4-like structures were more often predicted in the 5′ end of positive-sense RNA virus genomic sequences from *Flaviviridae* virus family than in human RNA families (Figure 2D). Importantly, we observed similar findings when using RNY1 as model (Figure S2F). Altogether, these results suggest that Y RNAs display patterns of viral mimicry and can be mobilized as RIG-I ligands upon infection.

HIV-1-dependent downregulation of DUSP11 determines endogenous RNAs triphosphorylation in infected cells

As Y RNAs are readily expressed in various cell types at steady state (Kowalski and Krude, 2015), we questioned what triggers RIG-I binding upon viral infection. Our results indicate that a 5′-PPP end is required for RNY4 RIG-I agonist activity. We performed a differential enzymatic digestion assay (Figure S3A) to analyze

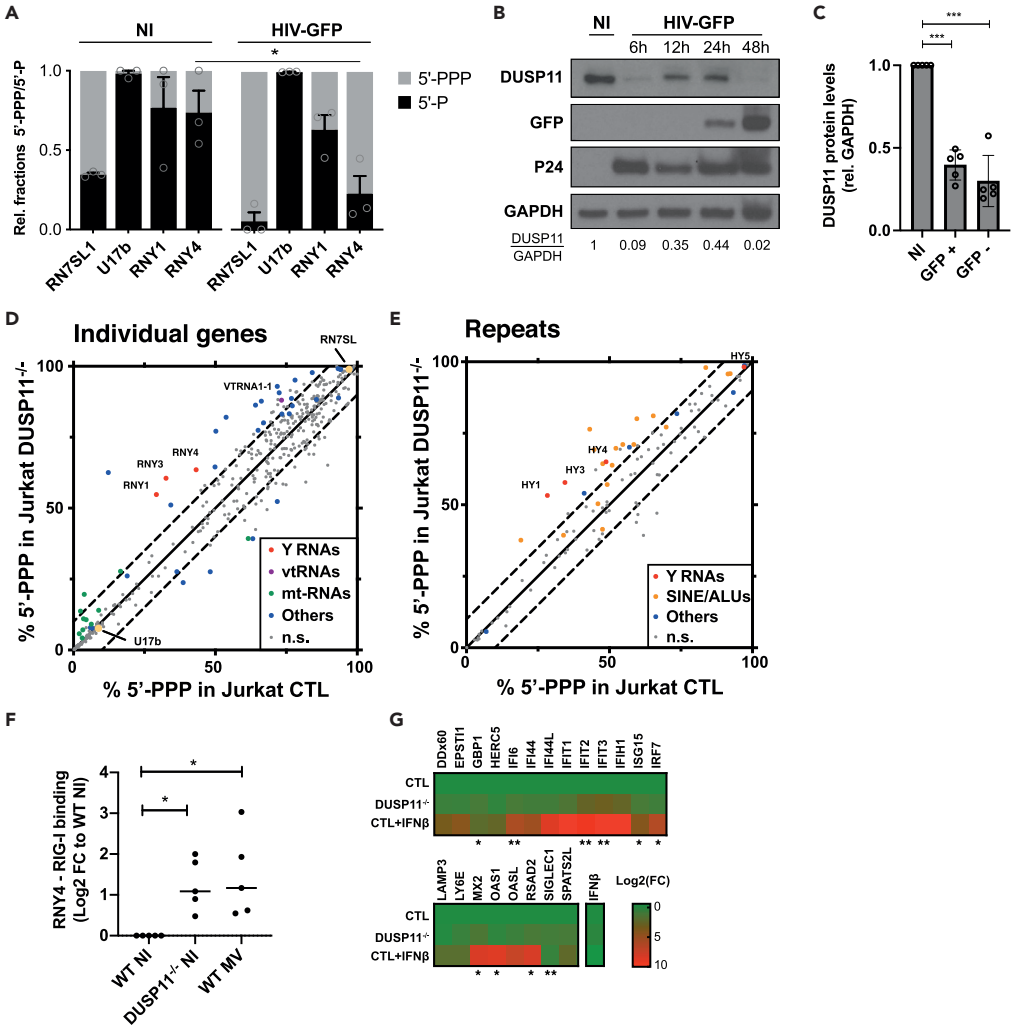


Figure 3. HIV-1-dependent downregulation of DUSP11 licenses endogenous 5′-PPP RNAs immunogenicity in infected cells

(A) Ratio of 5′-PPP and 5′-P-bearing RNY1 and RNY4 in Jurkat cells 48 h postinfection with HIV-GFP or in noninfected (NI) Jurkat cells. Relative 5′-PPP/5′-P RNA levels were determined through differential enzyme digestion followed by qPCR analysis relative to *β-actin* mRNA. *RN7SL1* and *U17b* are 5′-PPP and 5′-P RNA controls, respectively.

(B) DUSP11 protein levels measured at different times points after Jurkat T cell infection with HIV-GFP.

(C) DUSP11 protein levels measured in NI or HIV-GFP-infected productively (GFP+) or nonproductively (GFP−) CD4 primary cells from five different donors 48 h postinfection. CD4 T cells were beads-sorted from total PBMC and activated with phytohemagglutinin-L (PHA) for 72 h prior to infection with HIV-GFP. Forty-eight hours postinfection, productively infected cells were FACS sorted according to GFP expression. DUSP11 protein levels are quantified relative to GAPDH (see also Figure S3C).

(D and E) 5′-PPP RNA sequencing of Jurkat CTL or DUSP11$^{-/-}$ cells. Individual genes (D) and repeats (E) are plotted according to the average percentage of their 5′-PPP subsets in three Jurkat control clones (x axis) or DUSP11$^{-/-}$ clones (y axis). Positions of RN7SL1 and U17b genes are indicated, representing 5′-PPP and 5′-P RNA controls, respectively.

(F) *RNY4* RNA enrichment on RIG-I in DUSP11$^{-/-}$ cells. *RNY4* RNA level were measured by qPCR from total RNA, RIG-I-bound, and Cherry-bound fractions in 293 expressing ST-RIG-I or ST-Cherry and either WT, deficient for DUSP11 or after infection with MV. RIG-I binding is computed by measuring the abundance of RNY4 bound to RIG-I compared with protein control Cherry, after normalization to the total cellular RNA abundance in cells overexpressing RIG-I or Cherry, relative to *GAPDH* mRNA levels. Data are represented in Log2 FC compared with WT NI condition.

(G) Heatmap of qPCR values measuring expression level of a panel of IFN-I stimulated genes in Jurkat control, DUSP11$^{-/-}$, or control treated overnight with recombinant IFN-β. Expression levels are normalized to *β-actin* mRNA levels and to Jurkat control. Genes with significant enrichments in DUSP11$^{-/-}$ cells compared with control are indicated with stars.

Figure 3. *Continued*
(A) Bars show mean ± SEM of n = 3 independent experiments. Student's t test; *p < 0.05. (B) Western blot representative of n = 3 independent experiments. (C) Bars show mean ± SEM of n = 5 donors. (F) Bars show mean ± SEM of n = 5 independent experiments. (G) Heatmap shows mean of three control and three DUSP11$^{-/-}$ Jurkat clones. (H) Bars show mean ± SEM of n = 3 experimental replicates. (G) Student's paired t test; *p < 0.05, **p < 0.01.

the 5′ structure of RNAs in HIV-1-infected Jurkat T cells. Surprisingly, HIV-1 infection induced a hyper triphosphorylation of RNY4 compared with noninfected cells (Figure 3A). Pol3-transcribed RNA such as Y RNAs initially contains a 5′-PPP upon transcription that may be further edited by cellular enzymes. Among these, DUSP11 is a protein from the dual-specificity phosphatase family that displays 5′-triphophatase activity on several miRNA and other cellular noncoding RNAs (Burke et al., 2016; Choi et al., 2020; Deshpande et al., 1999). We hypothesized HIV-1 infection could modulate DUSP11 activity. Infection of Jurkat cells with an HIV-1 NL4.3 clone encoding GFP (HIV-GFP) led to profound DUSP11 protein downregulation in early (6 h) and late (48 h) time points (Figure 3B), whereas DUSP11 mRNA levels remained stable (Figure S3B). We further infected activated primary CD4^{+} T cells from five healthy donors with HIV-GFP, FACS-sorted the productively infected (GFP+), and nonproductively infected (GFP−) fractions and measured the levels of DUSP11 protein, which was downregulated in each fraction (Figures 3C, S3C, and S3D). We therefore hypothesized changes in Y RNAs 5′-triphosphorylation in HIV-1-infected cells were caused by DUSP11 downregulation.

To further explore the link between DUSP11 and triphosphorylation of cellular RNAs, we developed a novel sequencing approach that allows for the unbiased capture and quantification of 5′-PPP and 5′-P fractions of noncoding cellular RNAs (5′-PPPseq, Figure S3E). We applied this sequencing method to A549 WT or DUSP11$^{-/-}$ and observed a DUSP11-dependent increase of 5′-triphosphorylation of specific noncoding RNAs such as Y RNAs and vtRNAs (Figures S3F, S3G, and S3H), consistent with previous findings (Burke et al., 2016; Choi et al., 2020). We also detected a previously unreported increase in phosphorylation of mitochondrial RNAs and several ALU repeat families. We performed the same sequencing approach in Jurkat cells WT and DUSP11$^{-/-}$, observing similar specificity of DUSP11 RNA targets (Figures 3D, 3E and S3I), although with different magnitude. In addition, we confirmed the activity of DUSP11 on RNY1 and RNY4 by differential enzymatic digestion assay (Figure S3J). To establish the link between DUSP11-mediated changes in 5′-triphosphorylation and RIG-I binding, we generated DUSP11 knockout 293 cells stably expressing tagged RIG-I and measured the level of RNY4 binding. Deletion of DUSP11 led to a significant increase in RNY4 binding to RIG-I, similar to what is observed in MV-infected cells (Figure 3F), although we did not detect DUSP11 downregulation following MV infection (Figure S3K). Vault RNAs were previously found as the most enriched RNA in RIG-I-purified fraction during KSHV lytic reactivation (Zhao et al., 2018b). To compare vtRNAs and RNY4-specific binding to RIG-I, we performed an additional RIG-I-specific RNA purification from MV-infected WT and DUSP11 KO cells after UV crosslinking and measured similar RIG-I binding profiles of RNY4 and vtRNAs by RT-qPCR (Figure S3L). Finally, we found that DUSP11 deficiency was sufficient to trigger an innate immune response by measuring with qPCR the upregulation of genes from a panel of classic IFN-I ISGs (Figure 3G).

HIV-1 VPR induces DUSP11 downregulation and subsequent increase of endogenous 5′-PPP RNAs

The predominantly nuclear localization of DUSP11 (Burke and Sullivan, 2017), together with its rapid downregulation kinetics observed upon HIV-1 infection (Figure 3B), including in nonproductively infected cells (Figure 3C), points toward HIV-1 VPR as the possible mediator of DUSP11 downregulation. Indeed, *vpr* codes for a conserved accessory protein that incorporates into viral particles, contains a nuclear localization signal, and induces the proteasomal degradation of several host cell factors (Guenzel et al., 2014). We compared DUSP11 downregulation in Jurkat T cells infected by either WT HIV-1 or the same clone lacking *vpr* (HIVΔVPR). We found that VPR expression was required for HIV-1-induced DUSP11 protein downregulation (Figures 4A and S3B). Concordantly, expression of WT VPR after transduction by lentiviral vectors, but not of a VPR(Q65R) mutant unable to recruit the DCAF1/DDB/Cul4 ligase complex (Le Rouzic et al., 2007), was sufficient to induce DUSP11 downregulation in Jurkat (Figure 4B). To further confirm that VPR proteins incorporated in particles can readily downregulate DUSP11, we performed HIV-1-GFP infection in the presence of antiretroviral drugs that target different steps of HIV-1 life cycle. Only anti-fusion inhibitors, preventing entry of HIV particles after viral docking to cell membrane, but not drugs targeting the reverse-transcription or integration steps of HIV replication, could prevent DUSP11 downregulation (Figure 4C), including at early time point (Figure S4A). To confirm that VPR-dependent downregulation of

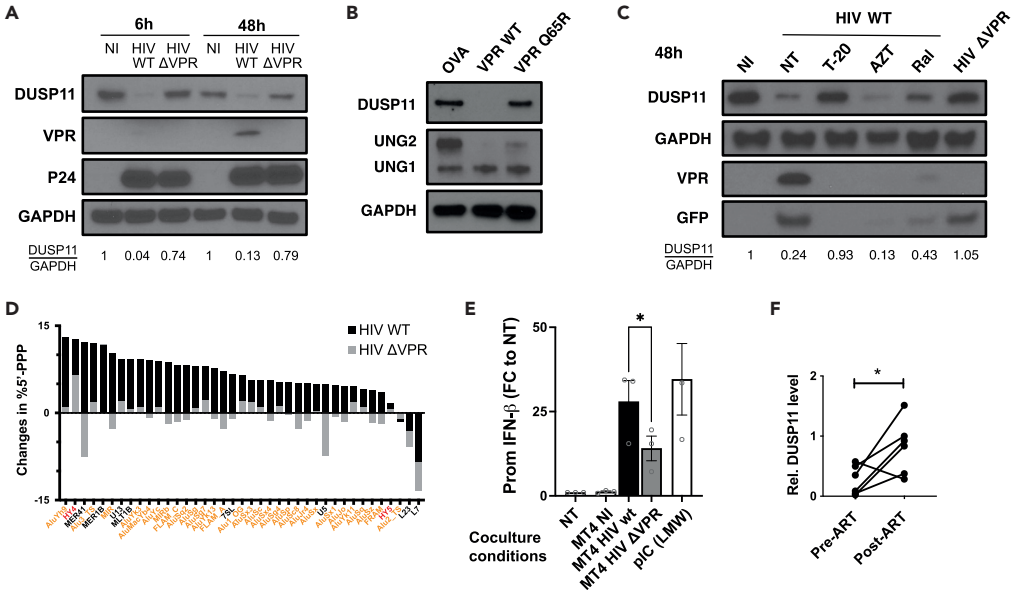


Figure 4. HIV-1 VPR induces DUSP11 downregulation and subsequent increase of endogenous 5′-PPP RNAs

(A) DUSP11 protein levels measured at 6 and 48 h after Jurkat T cell infection with WT NL4.3 HIV-1 or the same clone deleted for VPR protein.

(B) DUSP11 protein levels in FACS-sorted Jurkat cells 72 h following transduction with lentiviruses coding for ovalbumin (OVA) used as a control, HIV-1 VPR WT, or an HIV-1 VPR(Q65R) defective mutant. UNG2 serves as control of a VPR target downregulated by Q65R mutant (Langevin et al., 2009). UNG1 is a product detected by the same antibody that is not targeted by VPR.

(C) DUSP11 protein levels 48 h after HIV-1 infection in Jurkat cells treated with antiretroviral inhibitor or nontreated (NT). T-20: fusion inhibitor enfuvirtide (5 uM); AZT: reverse transcriptase inhibitor zidovudine (1 mM); Ral: integrase inhibitor raltegravir (10 uM).

(D) 5′-PPP RNA sequencing of FACS-sorted Jurkat 48 h after infection with HIV-GFP WT or HIV-GFP ΔVPR. Repeats are plotted according to their increase in 5′ triphosphorylation between infected and noninfected cells. Repeats with significant changes in 5′-PPP status upon HIV WT infection are plotted together with corresponding values upon HIV-ΔVPR infection (average of three infection replicates each).

(E) Promoter IFN-β-luciferase reporter activity in 293-4x4 after co-culture with HIV-GFPWT- or HIV-GFPdVPR-infected MT4C5. 293-4x4 without co-culture (NT) or co-cultured with noninfected (NI) MT4 serve as negative control, 293-4x4 transfected with 10 ng/ul of poly(I:C) low molecular weight (LMW) serve as positive control. The induction of IFN-β-luciferase is shown as fold change to NT control.

(F) Relative quantification of DUSP11 protein level in CD4 T cells from HIV + patients prior to and after antiretroviral treatment. Paired t test; *$p < 0.05$. See also Figure S4C.

(A and C) Numbers at the bottom indicate semi-quantification of relative DUSP11/GAPDH levels normalized to NI conditions. (A, B, and C) Western blot representative of n = 2 independent experiments. (E) Mean of three independent experiments.

DUSP11 is responsible for cellular changes in RNA 5′-triphosphorylation observed in infected cells, we performed 5′-PPPseq on FACS-sorted Jurkat cells infected with HIV-1-GFP or HIV-1-GFPΔVPR. The large majority of 5′-triphosphorylation increase observed for repeat RNA families during infection with WT HIV-1 were reduced or absent in cells infected by the mutant virus lacking VPR (Figure 4D). Finally, we measured the induction of IFN-I signaling in 293-4x4 upon co-culture with MT4 cells infected with either HIV-GFP WT or HIV-1-GFPΔVPR and found that despite similar infection levels, virus mutants lacking VPR had a reduced ability to trigger IFN-I signaling (Figures 4E and S4B).

We next investigated the significance of DUSP11 downregulation *in vivo*. In a cohort of HIV-1 positive (HIV+) patients, for which PBMCs were collected prior to and 6 months after antiretroviral treatment (ART) (Table S3) (Markowitz et al., 2014), we measured DUSP11 protein level on beads-isolated $CD4^+$ T cells. We found that DUSP11 protein levels were significantly increased in 5/6 patients after ART (Figures 4F and S4C), indicating HIV-1-mediated downregulation of DUSP11 before ART. Altogether, these results suggest that HIV-1 VPR actively alters the cellular pool of self-derived RLR ligands in

HIV-1 patients, through targeting of DUSP11, which functions as a mediator of endogenous 5′-PPP RNAs.

DISCUSSION

Here we describe the contribution of endogenous RNA sensing by RIG-I during RNA virus infection. In principle, any RNA transcribed by RNA Pol3 may have the ability to trigger RIG-I-dependent immune responses, at least transiently, because they initially contain 5′-PPP terminal regions upon initiation of transcription. Indeed, a few of the Pol3-dependent RNAs, including RN7SL, RNA5SP141, and vtRNAs, have been shown thus far to trigger immune responses in different settings (Chiang et al., 2018; Choi et al., 2020; Haderk et al., 2017; Hung et al., 2015; Johnson et al., 2021; Leonova et al., 2013; Nabet et al., 2017; Zhao et al., 2018a, 2018b). In this work, we specifically identified the family of Y RNAs, particularly RNY4, as a model of endogenous RNAs whose unique structure confers a previously unknown function as RIG-I agonists. The reasons why RNY4 appears to be a preferential RIG-I ligand compared with other Y RNAs are unclear. Although we cannot exclude that our approach failed to capture other endogenous ligands, such as shorter transcripts, differences in cellular compartmentalization, accessibility to binding proteins, or discrete specificities in secondary structure could confer RNY4 a higher ability to act as RIG-I ligand.

Interestingly, Y RNAs, similarly to other Pol3-dependent RNAs that have been shown to have immune agonist activity, are readily expressed in homeostatic conditions, indicating the mechanisms that control their function as innate immune ligands do not simply rely on transcriptional induction. We observed a contribution of Y RNAs as ligands during infections by MV and DV-4, both RNA viruses replicating in the cytoplasm and producing RLR-specific viral RNA ligands (Chazal et al., 2018; Sanchez David et al., 2016). MV and DV-4 are acute viral infections where the speed of host response is critical to halt and ultimately clear viral replication. We speculate that the viral-mimicking structure of these endogenous RLRs ligands licenses them to act as innate immune guardians that prime immune responses at the onset of cell infection.

We observed the same contribution of Y RNAs during infection by HIV-1. However, in the case of HIV-1 we surprisingly failed to detect any ligand of viral origin to either RIG-I, or MDA5, or LGP2, suggesting that in this case self-RNAs, rather than nonself, represent the main RNA ligands contributing to the RLR pathway. Similar to previous studies that identify vtRNAs or RNA5SP141 as self-derived RIG-I ligand during DNA virus infection (Chiang et al., 2018; Zhao et al., 2018b), our results suggest that Y RNAs form another class of noncoding RNAs that act as RIG-I substrates in various RNA virus conditions. We further characterized the role of cellular triphosphatase DUSP11 as a key immune modulator that prevents unwarranted sensing of cellular RNAs in healthy cells. Importantly, we show that HIV-1 has evolved mechanisms to manipulate and subvert this process through a targeted, VPR-dependent, downregulation of DUSP11 that leads to a surge of cellular RNA triphosphorylation and subsequent binding of self-RNAs to RIG-I. Interestingly, the effect of DUSP11 downregulation on host RNA triphosphorylation in HIV-infected cells does not seem to fully recapitulate what is measured in DUSP11 knockout cells. Notably, although we detected an increase of RNY1 and RNY4 triphosphorylation in DUSP11$^{-/-}$ cells, we did not observe differences for RNY1 between uninfected and infected cells. As HIV-1 infection does not lead to a complete degradation of DUSP11, a combination of low DUSP11 levels and differences in DUSP11 affinity for RNY1 and RNY4 could explain this observation. Further, all the Y RNAs do not seem to be susceptible to DUSP11 activity, as we did not detect any difference in HY5 triphosphorylation status in WT or DUSP11 knockout cells (Figures 3E and S3G), suggesting that other cellular mechanisms modulate the immunogenicity of this RNA at steady state.

Another study identified a reduction of DUSP11 mRNA and protein levels during Kaposi's sarcoma-associated herpesvirus (KSHV) lytic reactivation (Zhao et al., 2018b), perhaps a consequence of the widespread host mRNA decay associated with KSHV infection (Covarrubias et al., 2009). Interestingly, our study found that HIV-1 VPR-dependent downregulation of DUSP11 was mediated at the protein and not mRNA level and could happen in the early events of cell infection, possibly via the VPR proteins packaged inside the viral particles (Cohen et al., 1990). Although it might appear surprising that an HIV-1 accessory gene developed a function that increases the pool of triphosphorylated RNAs through DUSP11 depletion, this observation is in line with previous reports showing direct or potentiating effects of VPR on IFN-I induction (Zahoor et al., 2014; Vermeire et al., 2016). Further work will be required to characterize the evolutionary

forces that selected this ability and how it impacts HIV-1 infection pathogenicity. We speculate that RLR activation by Y RNAs acts as a rapid response mechanism for the host to detect when RNA dephosphorylation is disrupted in an infected cell. In the case of HIV-1, importantly, during the chronic phase infection, higher levels of IFN-I signaling correlate with sustained levels of inflammation, immune exhaustion, CD4 T cell depletion, and disease progression (Moir et al., 2011). In fact, chronic IFN-I signaling is considered by many as central to HIV-1 pathogenesis, to the point where the use of IFN-I blockade treatment is discussed as a supplement during ART (Deeks et al., 2017). In this context, it will be important to determine whether, and to what extent, immune activation subsequent to DUSP11 downregulation participates to drive chronic IFN-I signaling and disruption of immune homeostasis observed in HIV-1-infected patients. Finally, our results emphasize the role of PRRs in sensing not only microbial ligands but also self-derived ligands. In the context of host-pathogen interactions, these endogenous ligands, possibly owing to their molecular mimicry of specific pathogen-associated features, constitute a new class of immunomodulatory molecules that provide the host the unique advantage to control both their potency and accessibility to innate sensors.

Limitations of the study

Although canonicals Y RNA (RNY1, RNY3-5) are known to be transcribed by RNA Pol3, it is unclear which polymerase governs the transcription of Y RNA pseudogenes. As most of these pseudogenes lack RNA Pol3 promoters, they may be primarily transcribed as part of larger Pol2 transcripts.

Although we used multiple experimental systems to demonstrate the immunogenicity of Y RNAs (co-purification with RIG-I, transfection of IVT RNAs, or of plasmids inducing RNY4 transcription from two different promoters), characterizing the immunogenicity of an RNA sequence using *in vitro* transcribed molecule may be biased by inherent properties of the T7 polymerase, such as abortive transcription and intramolecular priming. Further, using IVT RNA does not interrogate other parameters that can modulate the immunogenicity of endogenous RNA molecules, such as epitranscriptomic modifications, shielding with endogenous proteins, and the intracellular stability of RNA molecules after transfection.

MAVS deficiency does not completely abrogate the IFN signaling after co-culture of infected cells with 293-4x4. It is possible that VPR is also responsible for the residual IFN signaling, as its role has previously been also shown for the cGAS pathway in a different experimental system (Vermeire et al., 2016).

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Cells lines
- METHOD DETAILS
  - Generation of CRISPR-edited cell lines
  - Affinity chromatography of RLR-RNP complexes and RNA extraction
  - Isolation of primary cells
  - *In vitro* transcription
  - Luciferase-based reporter assay
  - Infection with virus/transduction with vector
  - Differential 5′-PPP RNA digestion
  - RNA purification
  - Quantitative PCR
  - Western blotting
  - Antibodies
  - Purification of RLRs and RNA extraction with crosslinking
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - RNA-seq analysis of total and RLR-bound RNA

- 5′-PPP RNA sequencing
- RNA seq analysis (repetitive elements)
- Polymerase-III transcript annotation
- Modeling RNY4 and RNY1-like structure in transcripts

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2022.104599.

ACKNOWLEDGMENTS

NV would like to thank Alice Lepelley, Sarah Moyon and Miriam Merad for insightful discussions and suggestions. The authors would like to thank members of the Tangy, van der Werf, Bhardwaj, and Greenbaum labs for support and valuable discussions. We are grateful to Lisset Hermida and Gerardo Enrique Guillen Nieto (CIGB, Havana, Cuba) for providing DV-4 isolate, Christopher Sullivan (UT Austin) for sharing a DUSP11 overexpressing lentiviral clone and Namita Sajita and Benjamin Chen (Icahn school of Medicine at Mount Sinai, NYC, USA) for providing the HIV-GFP and HIV-mCherry clones and valuable technical help. VN and AK would like to thank Florence Guivel-Benhassine (Virus and Immunity Laboratory of Institut Pasteur), Caroline Demeret (Molecular Genetics of RNA Viruses laboratory if Institut Pasteur), and Pierre-Henri Commere (Plate-Forme de Cytométrie, Institut Pasteur) for technical support.

This work was supported by Agence Nationale de Recherches sur le SIDA et les hépatites virales 2012-1 and 2012-2 HIV signature on RIG-I-like receptors to AVK and FT; Agence Nationale pour la Recherche ANR-16-CE15-0025-01 ViroStorm to NJ and GB; Fondation pour la Recherche Medicale FDT20140931129 to RYSD; National Institutes of Health R01CA201189 and R01CA180913 to NB; R01AI081848 to NV, NB, and BDG; U01CA228963 to NV, NB, AS and BDG, R01CA240924 and P30CA008748 to AS and BDG. BDG was supported by a Stand Up To Cancer—National Science Foundation—Lustgarten Foundation Convergence Dream Team Grant sponsored by Stand Up to Cancer, the Lustgarten Foundation, the V Foundation, and the National Science Foundation (NSF 1545935).BDG BDG is The Pershing Square Sohn Prize-Mark Foundation Fellow supported by funding from The Mark Foundation for Cancer Research.

Icahn School of medicine has a patent related to this work (WO 2016/131048 Al) on which BDG and NB are inventors. BDG has received honoraria for speaking engagements from Merck, Bristol–Meyers Squibb, and Chugai Pharmaceuticals; has received research funding from Bristol-Meyers Squibb; and has been a compensated consultant for PMV Pharma, Darwin Health, Shennon Biotechnologies, and Rome Therapeutics of which he is a cofounder. NB has received honoraria for speaking engagements from Pfizer and research funds from the Cancer Research Institute, Merck, Regeneron, Ludwig Institute, Melanoma Research Alliance, Leukemia & Lymphoma Society, Antidote Health Foundation, and reports consultancy and/or advisory roles for Novartis, Apricity, Rome Therapeutics, CureVac, Genotwin, BioNTech, Gilead, Tempest Therapeutics, Genentech, and Rubis Therapeutics.

AUTHOR CONTRIBUTIONS

Experimental methodology and analysis: N.V., A.K. Experimental procedure: N.V., V.N., R.G., C.M., S.B., Y.R., M.C., O.Sis., R.Y.S.D., A.K. Computational methodology and analysis: A.S., P.S., G.B., R.L., H.V., N.V., and B.D.G. Resources: M.M. and O.Sch. Supervision: L.C., N.J., O.Sch., S.v.d.W., F.T., B.D.G, N.B., and A.K. Writing—original draft: N.V. Writing—review and editing: N.V., A.K., B.D.G., and N.B. A.K., B.D.G., and N.B. are credited with senior authorship.

DECLARATION OF INTERESTS

The authors declare no conflicting interests.

Received: October 25, 2021
Revised: March 1, 2022
Accepted: June 7, 2022
Published: July 15, 2022

REFERENCES

Barski, A., Chepelev, I., Liko, D., Cuddapah, S., Fleming, A.B., Birch, J., Cui, K., White, R.J., and Zhao, K. (2010). Pol II and its associated epigenetic marks are present at Pol III-transcribed noncoding RNA genes. Nat. Struct. Mol. Biol. *17*, 629–634. https://doi.org/10.1038/nsmb.1806.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. Roy. Stat. Soc. B *57*, 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

Brister, J.R., Ako-Adjei, D., Bao, Y., and Blinkova, O. (2015). NCBI viral genomes resource. Nucleic Acids Res. *43*, D571–D577. https://doi.org/10.1093/nar/gku1207.

Burke, J.M., Kincaid, R.P., Nottingham, R.M., Lambowitz, A.M., and Sullivan, C.S. (2016). DUSP11 activity on triphosphorylated transcripts promotes Argonaute association with noncanonical viral microRNAs and regulates steady-state levels of cellular noncoding RNAs. Genes Dev. *30*, 2076–2092. https://doi.org/10.1101/gad.282616.116.

Burke, J.M., and Sullivan, C.S. (2017). DUSP11 - an RNA phosphatase that regulates host and viral non-coding RNAs in mammalian cells. RNA Biol. *14*, 1457–1465. https://doi.org/10.1080/15476286.2017.1306169.

Canella, D., Praz, V., Reina, J.H., Cousin, P., and Hernandez, N. (2010). Defining the RNA polymerase III transcriptome: genome-wide localization of the RNA polymerase III transcription machinery in human cells. Genome Res. *20*, 710–721. https://doi.org/10.1101/gr.101337.109.

Chazal, M., Beauclair, G., Gracias, S., Najburg, V., Simon-Lorière, E., Tangy, F., Komarova, A.V., and Jouvenet, N. (2018). RIG-I recognizes the 5′ region of dengue and zika virus genomes. Cell Rep. *24*, 320–328. https://doi.org/10.1016/j.celrep.2018.06.047.

Chiang, J.J., Sparrer, K.M.J., van Gent, M., Lässig, C., Huang, T., Osterrieder, N., Hopfner, K.P., and Gack, M.U. (2018). Viral unmasking of cellular 5S rRNA pseudogene transcripts induces RIG-I-mediated immunity. Nat. Immunol. *19*, 53–62. https://doi.org/10.1038/s41590-017-0005-y.

Chiappinelli, K.B., Strissel, P.L., Desrichard, A., Li, H., Henke, C., Akman, B., Hein, A., Rote, N.S., Cope, L.M., Snyder, A., Makarov, V., Budhu, S., Slamon, D., Wolchok, J., Pardoll, D., Beckmann, M., Zahnow, C., Merghoub, T., Chan, T., Baylin, S., and Strick, R. (2016). Inhibiting DNA methylation causes an interferon response in cancer via dsRNA including endogenous retroviruses. Cell *164*, 1073. https://doi.org/10.1016/j.cell.2015.10.020.

Choi, J.H., Burke, J.M., Szymanik, K.H., Nepal, U., Battenhouse, A., Lau, J.T., Stark, A., Lam, V., and Sullivan, C.S. (2020). DUSP11-mediated control of 5′-triphosphate RNA regulates RIG-I sensitivity. Genes Dev. *34*, 1697–1712. https://doi.org/10.1101/gad.340604.120.

Chow, K.T., Gale, M., Jr., and Loo, Y.M. (2018). RIG-I and other RNA sensors in antiviral immunity. Annu. Rev. Immunol. *36*, 667–694. https://doi.org/10.1146/annurev-immunol-042617-053309.

Cohen, E.A., Dehni, G., Sodroski, J.G., and Haseltine, W.A. (1990). Human immunodeficiency virus vpr product is a virion-associated regulatory protein. J. Virol. *64*, 3097–3099. https://doi.org/10.1128/jvi.64.6.3097-3099.1990.

Combredet, C., Labrousse, V., Mollet, L., Lorin, C., Delebecque, F., Hurtrel, B., McClure, H., Feinberg, M.B., Brahic, M., and Tangy, F. (2003). A molecularly cloned Schwarz strain of measles virus vaccine induces strong immune responses in macaques and transgenic mice. J. Virol. *77*, 11546–11554. https://doi.org/10.1128/jvi.77.21.11546-11554.2003.

Covarrubias, S., Richner, J.M., Clyde, K., Lee, Y.J., and Glaunsinger, B.A. (2009). Host shutoff is a conserved phenotype of gammaherpesvirus infection and is orchestrated exclusively from the cytoplasm. J. Virol. *83*, 9554–9566. https://doi.org/10.1128/jvi.01051-09.

Deeks, S.G., Odorizzi, P.M., and Sekaly, R.P. (2016). The interferon paradox: can inhibiting an antiviral mechanism advance an HIV cure? J. Clin. Invest. *127*, 103–105. https://doi.org/10.1172/jci91916.

Deshpande, T., Takagi, T., Hao, L., Buratowski, S., and Charbonneau, H. (1999). Human PIR1 of the protein-tyrosine phosphatase superfamily has RNA 5′-triphosphatase and diphosphatase activities. J. Biol. Chem. *274*, 16590–16594. https://doi.org/10.1074/jbc.274.23.16590.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.

Guenzel, C.A., HÃ©rate, C., and Benichou, S. (2014). HIV-1 Vpr-a still "enigmatic multitasker". Front. Microbiol. *5*, 127. https://doi.org/10.3389/fmicb.2014.00127.

Haderk, F., Schulz, R., Iskar, M., Cid, L.L., Worst, T., Willmund, K.V., Schulz, A., Warnken, U., Seiler, J., Benner, A., et al. (2017). Tumor-derived exosomes modulate PD-L1 expression in monocytes. Sci Immunol. *2*, eaah5509. https://doi.org/10.1126/sciimmunol.aah5509.

Heidegger, S., Wintges, A., Stritzke, F., Bek, S., Steiger, K., Koenig, P.A., Göttert, S., Engleitner, T., Ollinger, R., Nedelko, T., et al. (2019). RIG-I activation is critical for responsiveness to checkpoint blockade. Sci Immunol. *4*, eaau8943. https://doi.org/10.1126/sciimmunol.aau8943.

Hubley, R., Finn, R.D., Clements, J., Eddy, S.R., Jones, T.A., Bao, W., Smit, A.F., and Wheeler, T.J. (2016). The Dfam database of repetitive DNA families. Nucleic Acids Res. *44*, D81–D89. https://doi.org/10.1093/nar/gkv1272.

Hubner, W., McNerney, G.P., Chen, P., Dale, B.M., Gordon, R.E., Chuang, F.Y.S., Li, X.D., Asmuth, D.M., Huser, T., and Chen, B.K. (2009). Quantitative 3D video microscopy of HIV transfer across T cell virological synapses. Science *323*, 1743–1747. https://doi.org/10.1126/science.1167525.

Hung, T., Pratt, G.A., Sundararaman, B., Townsend, M.J., Chaivorapol, C., Bhangale, T., Graham, R.R., Ortmann, W., Criswell, L.A., Yeo, G.W., and Behrens, T.W. (2015). The Ro60 autoantigen binds endogenous retroelements and regulates inflammatory gene expression. Science *350*, 455–459. https://doi.org/10.1126/science.aac7442.

Ishizuka, J.J., Manguso, R.T., Cheruiyot, C.K., Bi, K., Panda, A., Iracheta-Vellve, A., Miller, B.C., Du, P.P., Yates, K.B., Dubrot, J., et al. (2019). Loss of ADAR1 in tumours overcomes resistance to immune checkpoint blockade. Nature *565*, 43–48. https://doi.org/10.1038/s41586-018-0768-9.

Johnson, L.R., Lee, D.Y., Eacret, J.S., Ye, D., June, C.H., and Minn, A.J. (2021). The immunostimulatory RNA RN7SL1 enables CAR-T cells to enhance autonomous and endogenous immune function. Cell *184*, 4981–4995.e14. https://doi.org/10.1016/j.cell.2021.08.004.

Kowalski, M.P., and Krude, T. (2015). Functional roles of non-coding Y RNAs. Int. J. Biochem. Cell Biol. *66*, 20–29. https://doi.org/10.1016/j.biocel.2015.07.003.

Laforge, M., Limou, S., Harper, F., Casartelli, N., Rodrigues, V., Silvestre, R., Haloui, H., Zagury, J.F., Senik, A., and Estaquier, J. (2013). DRAM triggers lysosomal membrane permeabilization and cell death in CD4(+) T cells infected with HIV. PLoS Pathog. *9*, e1003328. https://doi.org/10.1371/journal.ppat.1003328.

Langevin, C., Maidou-Peindara, P., Aas, P.A., Jacquot, G., Otterlei, M., Slupphaug, G., and Benichou, S. (2009). Human immunodeficiency virus type 1 Vpr modulates cellular expression of UNG2 via a negative transcriptional effect. J. Virol. *83*, 10256–10263. https://doi.org/10.1128/jvi.02654-08.

Law, K.M., Komarova, N.L., Yewdall, A.W., Lee, R.K., Herrera, O.L., Wodarz, D., and Chen, B.K. (2016). In vivo HIV-1 cell-to-cell transmission promotes multicopy micro-compartmentalized infection. Cell Rep. *15*, 2771–2783. https://doi.org/10.1016/j.celrep.2016.05.059.

Le Rouzic, E., Belaïdouni, N., Estrabaud, E., Morel, M., Rain, J.C., Transy, C., and Margottin-Goguet, F. (2007). HIV1 Vpr arrests the cell cycle by recruiting DCAF1/VprBP, a receptor of the Cul4-DDB1 ubiquitin ligase. Cell Cycle *6*, 182–188. https://doi.org/10.4161/cc.6.2.3732.

Lehmann, S.M., Krüger, C., Park, B., Derkow, K., Rosenberger, K., Baumgart, J., Trimbuch, T., Eom, G., Hinz, M., Kaul, D., et al. (2012). An unconventional role for miRNA: let-7 activates Toll-like receptor 7 and causes neurodegeneration. Nat. Neurosci. *15*, 827–835. https://doi.org/10.1038/nn.3113.

Leonova, K.I., Brodsky, L., Lipchick, B., Pal, M., Novototskaya, L., Chenchik, A.A., Sen, G.C., Komarova, E.A., and Gudkov, A.V. (2013). p53 cooperates with DNA methylation and a suicidal interferon response to maintain epigenetic silencing of repeats and noncoding RNAs. Proc. Natl. Acad. Sci. USA *110*, E89–E98. https://doi.org/10.1073/pnas.1216922110.

Lepelley, A., Louis, S., Sourisseau, M., Law, H.K.W., Pothlichet, J., Schilte, C., Chaperot, L.,

Plumas, J., Randall, R.E., Si-Tahar, M., et al. (2011). Innate sensing of HIV-infected cells. PLoS Pathog. *7*, e1001284. https://doi.org/10.1371/journal.ppat.1001284.

Liao, Y., Smyth, G.K., and Shi, W. (2013). The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. Nucleic Acids Res. *41*, e108. https://doi.org/10.1093/nar/gkt214.

Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930. https://doi.org/10.1093/bioinformatics/btt656.

Lin, R., Génin, P., Mamane, Y., and Hiscott, J. (2000). Selective DNA binding and association with the CREB binding protein coactivator contribute to differential activation of alpha/beta interferon genes by interferon regulatory factors 3 and 7. Mol. Cell. Biol. *20*, 6342–6353.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550. https://doi.org/10.1186/s13059-014-0550-8.

Lucas-Hourani, M., Dauzonne, D., Jorda, P., Cousin, G., Lupan, A., Helynck, O., Caignard, G., Janvier, G., André-Leroux, G., Khiar, S., et al. (2013). Inhibition of pyrimidine biosynthesis pathway suppresses viral growth through innate immunity. PLoS Pathog. *9*, e1003678. https://doi.org/10.1371/journal.ppat.1003678.

Mackow, E., Makino, Y., Zhao, B.T., Zhang, Y.M., Markoff, L., Buckler-White, A., Guiler, M., Chanock, R., and Lai, C.J. (1987). The nucleotide sequence of dengue type 4 virus: analysis of genes coding for nonstructural proteins. Virology *159*, 217–228. https://doi.org/10.1016/0042-6822(87)90458-2.

Markowitz, M., Evering, T.H., Garmon, D., Caskey, M., La Mar, M., Rodriguez, K., Sahi, V., Palmer, S., Prada, N., and Mohri, H. (2014). A randomized open-label study of 3- versus 5-drug combination antiretroviral therapy in newly HIV-1-infected individuals. J. Acquir. Immune Defic. Syndr. *66*, 140–147. https://doi.org/10.1097/qai.0000000000000111.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal *17*, 10–12.

Mathews, D.H., Sabina, J., Zuker, M., and Turner, D.H. (1999). Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J. Mol. Biol. *288*, 911–940. https://doi.org/10.1006/jmbi.1999.2700.

Moir, S., Chun, T.W., and Fauci, A.S. (2011). Pathogenic mechanisms of HIV disease. Annu. Rev. Pathol. *6*, 223–248. https://doi.org/10.1146/annurev-pathol-011110-130254.

Moqtaderi, Z., Wang, J., Raha, D., White, R.J., Snyder, M., Weng, Z., and Struhl, K. (2010). Genomic binding profiles of functionally distinct RNA polymerase III transcription complexes in human cells. Nat. Struct. Mol. Biol. *17*, 635–640. https://doi.org/10.1038/nsmb.1794.

Nabet, B.Y., Qiu, Y., Shabason, J.E., Wu, T.J., Yoon, T., Kim, B.C., Benci, J.L., DeMichele, A.M., Tchou, J., Marcotrigiano, J., and Minn, A.J. (2017). Exosome RNA unshielding couples stromal activation to pattern recognition receptor signaling in cancer. Cell *170*, 352–366.e13. https://doi.org/10.1016/j.cell.2017.06.031.

Oler, A.J., Alla, R.K., Roberts, D.N., Wong, A., Hollenhorst, P.C., Chandler, K.J., Cassiday, P.A., Nelson, C.A., Hagedorn, C.H., Graves, B.J., and Cairns, B.R. (2010). Human RNA polymerase III transcriptomes and relationships to Pol II promoter chromatin and enhancer-binding factors. Nat. Struct. Mol. Biol. *17*, 620–628. https://doi.org/10.1038/nsmb.1801.

Perreault, J., Noel, J.F., Briere, F., Cousineau, B., Lucier, J.F., Perreault, J.P., and Boire, G. (2005). Retropseudogenes derived from the human Ro/SS-A autoantigen-associated hY RNAs. Nucleic Acids Res. *33*, 2032–2041. https://doi.org/10.1093/nar/gki504.

Perreault, J., Perreault, J.-P., and Boire, G. (2007). Ro-associated Y RNAs in metazoans: evolution and diversification. Mol. Biol. Evol. *24*, 1678–1689.

Raha, D., Wang, Z., Moqtaderi, Z., Wu, L., Zhong, G., Gerstein, M., Struhl, K., and Snyder, M. (2010). Close association of RNA polymerase II and many transcription factors with Pol III genes. Proc. Natl. Acad. Sci. USA *107*, 3639–3644. https://doi.org/10.1073/pnas.0911315106.

Ranoa, D.R.E., Parekh, A.D., Pitroda, S.P., Huang, X., Darga, T., Wong, A.C., Huang, L., Andrade, J., Staley, J.P., Satoh, T., et al. (2016). Cancer therapies activate RIG-I-like receptor pathway through endogenous non-coding RNAs. Oncotarget. *7*, 26496–26515. https://doi.org/10.18632/oncotarget.8420.

Roulois, D., Loo Yau, H., Singhania, R., Wang, Y., Danesh, A., Shen, S.Y., Han, H., Liang, G., Jones, P.A., Pugh, T.J., et al. (2015). DNA-demethylating agents target colorectal cancer cells by inducing viral mimicry by endogenous transcripts. Cell *162*, 961–973. https://doi.org/10.1016/j.cell.2015.07.056.

Sanchez David, R.Y., Combredet, C., Sismeiro, O., Dillies, M.A., Jagla, B., Coppée, J.Y., Mura, M., Guerbois Galla, M., Despres, P., Tangy, F., and Komarova, A.V. (2016). Comparative analysis of viral RNA signatures on different RIG-I-like receptors. Elife *5*, e11275. https://doi.org/10.7554/elife.11275.

Solovyov, A., Vabret, N., Arora, K.S., Snyder, A., Funt, S.A., Bajorin, D.F., Rosenberg, J.E., Bhardwaj, N., Ting, D.T., and Greenbaum, B.D. (2018). Global cancer transcriptome quantifies repeat element polarization between immunotherapy responsive and T cell suppressive classes. Cell Rep. *23*, 512–521. https://doi.org/10.1016/j.celrep.2018.03.042.

Tanne, A., Muniz, L.R., Puzio-Kuter, A., Leonova, K.I., Gudkov, A.V., Ting, D.T., Monasson, R., Cocco, S., Levine, A.J., Bhardwaj, N., and Greenbaum, B.D. (2015). Distinguishing the immunostimulatory properties of noncoding RNAs expressed in cancer cells. Proc. Natl. Acad. Sci. USA *112*, 15154–15159. https://doi.org/10.1073/pnas.1517584112.

Turowski, T.W., and Tollervey, D. (2016). Transcription by RNA polymerase III: insights into mechanism and regulation. Biochem. Soc. Trans. *44*, 1367–1375. https://doi.org/10.1042/bst20160062.

Vabret, N., Bhardwaj, N., and Greenbaum, B.D. (2017). Sequence-specific sensing of nucleic acids. Trends Immunol. *38*, 53–65. https://doi.org/10.1016/j.it.2016.10.006.

Vermeire, J., Roesch, F., Sauter, D., Rua, R., Hotter, D., Van Nuffel, A., Vanderstraeten, H., Naessens, E., Iannucci, V., Landi, A., et al. (2016). HIV triggers a cGAS-dependent, vpu- and vpr-regulated type I interferon response in CD4(+) T cells. Cell Rep. *17*, 413–424. https://doi.org/10.1016/j.celrep.2016.09.023.

White, R.J. (2011). Transcription by RNA polymerase III: more complex than we thought. Nat. Rev. Genet. *12*, 459–463. https://doi.org/10.1038/nrg3001.

Wolin, S.L., and Steitz, J.A. (1983). Genes for two small cytoplasmic Ro RNAs are adjacent and appear to be single-copy in the human genome. Cell *32*, 735–744. https://doi.org/10.1016/0092-8674(83)90059-4.

Zahoor, M.A., Xue, G., Sato, H., Murakami, T., Takeshima, S.N., and Aida, Y. (2014). HIV-1 Vpr induces interferon-stimulated genes in human monocyte-derived macrophages. PLoS One *9*, e106418. https://doi.org/10.1371/journal.pone.0106418.

Zhao, K., Du, J., Peng, Y., Li, P., Wang, S., Wang, Y., Hou, J., Kang, J., Zheng, W., Hua, S., and Yu, X.F. (2018a). LINE1 contributes to autoimmunity through both RIG-I- and MDA5-mediated RNA sensing pathways. J. Autoimmun. *90*, 105–115. https://doi.org/10.1016/j.jaut.2018.02.007.

Zhao, Y., Ye, X., Dunker, W., Song, Y., and Karijolich, J. (2018b). RIG-I like receptor sensing of host RNAs facilitates the cell-intrinsic immune response to KSHV infection. Nat. Commun. *9*, 4841. https://doi.org/10.1038/s41467-018-07314-7.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse monoclonal HIV-1 p24 (clone 39/5.4A)	Abcam	Cat: #ab9071
Rabbit polyclonal anti-MAVS	Cell Signaling Technology	Cat: #3993
Rabbit polyclonal anti-DUSP11	Proteintech	Cat: #10204-2-AP
Mouse monoclonal anti-Tubulin (clone 1E4C11)	Cell Signaling Technology	Cat: #2118
Rabbit monoclonal GFP (clone D5.1)	Cell Signaling Technology	Cat: #2956
Rabbit monoclonal GAPDH (clone 14C10)	Cell Signaling Technology	Cat: #2118
Mouse monoclonal UNG (clone OTI2C12)	Origene	Cat# : TA503563
Rabbit polyclonal VPR	Proteintech	Cat: # 51143-1-AP
Mouse monoclonal STrEP-Tag	Qiagen	Cat: #34850
Peroxidase-conjugated secondary antibodies against rabbit IgG	Cell Signaling Technology	Cat: #7074
Peroxidase-conjugated secondary antibodies against mouse IgG	Cell Signaling Technology	Cat: #7076
Mouse monoclonal anti-p24-FITC	Beckman and Coulter	Cat: #KC-57
Mouse monoclonal anti-CD3-Pacific Blue	Biolegend	Cat: #300539
Mouse monoclonal anti-CD19-APC	Biolegend	Cat: #302212
Mouse monoclonal anti-CD14-APC	Biolegend	Cat: #325608
Mouse monoclonal anti-CD56-APC	Biolegend	Cat: #318310
Mouse monoclonal anti-CD8-PerCP/Cy5.5	Bdbiosciences	Cat: #341051
Bacterial and virus strains		
HIVNL4.3WT	Gift from Olivier Schwartz laboratory	N/A
HIVNL4.3ΔVPR	Gift from Olivier Schwartz laborator	N/A
HIVNL4.3-GFP	Gift from Benjamin Chen laboratory Law et al. (2016)	N/A
HIVNL4.3-mcherry	from Benjamin Chen laboratory Law et al. (2016)	N/A
HIVNL4.3-GFP ΔVPR	This paper	N/A
MV Schwarz	Combredet et al. (2003)	N/A
DV-4 Dominica	Chazal et al. (2018)	N/A
Biological samples		
PBMC from healthy donors	New York blood center	N/A
HIV-1 cohort patients	Markowitz et al. (2014)	N/A
Chemicals, peptides, and recombinant proteins		
Ficoll	GE Healthcare Life Sciences	Cat: #GE17-1440-02
Trizol	Invitrogen	Cat: #15596026
Antarctic phosphatase	New England BioLabs	Cat: #M0289S
Lipofectamine 2000	Invitrogen	Cat: #11668030
Passive Lysis 5X Buffer	Promega	Cat: #E1941
low molecular weight poly(I:C)	Invivogen	Cat: #tlrl-picw
high molecular weight poly(I:C)	Invivogen	Cat: #tlrl-pic

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
DMEM, low glucose, GlutaMAX™ Supplement, pyruvate	ThermoFisher Scientific (GIBCO)	Cat: #10567014
heat-inactivated fetal calf serum	ThermoFisher Scientific (GIBCO)	Cat: #16000044
Penicillin-Streptomycin	ThermoFisher Scientific (GIBCO)	Cat: # 15140122
Gibco™ RPMI 1640 Medium	ThermoFisher Scientific (GIBCO)	Cat: # 11875101
human serum	Gemini Bio	Cat: #100-110
HEPES buffer	ThermoFisher Scientific (GIBCO)	Cat: #15630106
L-glutamine	ThermoFisher Scientific (GIBCO)	Cat: #A2916801
G-418 Solution	Sigmaaldrich	Cat: #4727878001
Iscove's Modified Dulbecco's Medium (IMDM)	ThermoFisher Scientific (GIBCO)	Cat: #31980030
Puromycin	Invivogen	Cat: #ant-pr-1
Polyjet	Signagen	Cat: #SL100688
Polybrene	EMD Millipore	Cat: #TR-1003-G
Phytohemagglutinin-L (PHA-L)	Sigmaaldrich	Cat: #11249738001
RNA 5′ polyphosphatase	Lucigen	Cat: #RP8092H
Terminator™ 5′-Phosphate-Dependent Exonuclease	Lucigen	Cat: #TER51020
4X Laemmi sample buffer	BIO-RAD	Cat: #1610747
TGX gels	BIO-RAD	Cat: #5671101
Tris/Glycine/SDS buffer	BIO-RAD	Cat: #161-0732EDU
Tris/Glycine buffer	BIO-RAD	Cat: #1610734EDU
Tris-buffered saline	BIO-RAD	Cat: #1706435EDU
0.1% Tween 20	ThermoFisher Scientific	Cat: #28320
ECL Plus Western Blotting Substrate	ThermoFisher Scientific (Pierce)	Cat: #34580
CL-Xposure Film	ThermoFisher Scientific	Cat: #34089
poly-L-Lysine-hydrobromide	Sigma	Cat: # P2636-25MG
RNasin	Promega	Cat: #N2511
Streptactin Sepharose beads	GE Healthcare	Cat: #28-9355-99
10X elution buffer	IBA, Biotin Elution Buffer 10X	Cat: # 2-1000-025
Proteinase K	Roche	Cat: # 3115887001
TRI Reagent LS	Sigma	Cat: # T3934
5′-phosphate-dependent XRN-1	New England BioLabs	Cat: #M0338S
Critical commercial assays		
EasySep Human $CD4^+$ T Cell Isolation Kit	Stemcell Technologies	Cat: #19052
EasySep Dead Cell Removal Annexin V Kit	Stemcell technology	Cat: #17899
T7 RiboMAX express large-scale RNA production system	Promega	Cat: #P1320
RNeasy kit	Qiagen	Cat: #NC9307831
Dual Luciferase Reporter Assay system	Promega	Cat: #E1910
Bright-Glo Luciferase Assay System	Promega	Cat: #E2610
Zymo RNA Clean and Concentrator	Zymo Research	Cat: #R1013
TURBO DNA-free Kit	Invitrogen	Cat: #10792877
TaqMan Fast Advanced Master Mix	Thermo Fisher Scientific	Cat: #4444556
Universal SYBR Green Supermix	BIO-RAD	Cat: #1725270EDU
Bioanalyser total RNA nano/pico kit	Agilent	Cat: #5067-1511/#5067-1513
TruSeq stranded total RNA library prep kit	Illumina	Cat: #20020596
riboPOOL Kit for human	siTOOLs Biotech	Cat: #054
ERCC RNA Spike-In Mix 1	Thermo Fisher Scientific	Cat: #4456740

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
cDNA EcoDry Premix Double Primed	Clontech	Cat: #639549
Deposited data		
Raw NGS data (total and RLR-bound RNA)	GEO	GEO: GSE134861
Raw NGS data (5′-PPP sequencing)	GEO	GEO: GSE203128
Experimental models: Cell lines		
HEK293	ATCC	Cat: #CRL-1573
HEK293T	ATCC	Cat: #CRL-3216
Jurkat T	Gift from Brian Brown laboratory	N/A
MT4C5	Gift from Olivier Schwartz laboratory	N/A
HEK293 ST-RIG-I	Sanchez-David et al. (2016)	N/A
HEK293 ST-MDA5	Sanchez-David et al. (2016)	N/A
HEK293 ST-LGP2	Sanchez-David et al. (2016)	N/A
HEK293 ST-CH	Sanchez-David et al. (2016)	N/A
HEK293 ST-RIG-I-4x4	This paper	N/A
HEK293 ST-MDA5-4x4	This paper	N/A
HEK293 ST-LGP2-4x4	This paper	N/A
HEK293 ST-CH-4x4	This paper	N/A
HEK293 STING-37	Lucas-Hourani et al. (2013)	N/A
HAP1 RIG-I ko	Horizon Discovery	Cat: #HZGHC001441c001
HAP1 LGP2 ko	Horizon Discovery	Cat: #HZGHC002927c011
HAP1 MDA5 ko	Horizon Discovery	Cat: #HZGHC001448c012
HAP1 MAVS ko	Horizon Discovery	Cat: #HZGHC001456c011
HAP parental cell line	Horizon Discovery	Cat: #C631
HEK293 ST-RIG-I DUSP11KO	This paper	N/A
HEK293 ST-CH DUSP11KO	This paper	N/A
Jurkat DUSP11KO	This paper	N/A
A549	ATCC	Cat: # CCL-185
A549 DUSP11KO	This paper	N/A
HEK293 MAVSKO	This paper	N/A
HEK293 negative KO	This paper	N/A
HEK293 ST-CH-4x4 MAVSKO	This paper	N/A
HEK293 ST-CH-4x4 negative KO	This paper	N/A
Experimental models: Organisms/strains		
N/A	N/A	N/A
Oligonucleotides		
Primers for DNA template used for *in vitro* transcription	This paper	Table S4
qPCR primers and probes used in this stud	This paper	Table S5
Recombinant DNA		
CRISPR–Cas9-expressing knockout plasmid MAVS	Santa Cruz	Cat: #sc-400769-ko-2
CRISPR–Cas9-expressing knockout plasmid RIG-I	Santa Cruz	Cat: #sc-400812
Homology Directed Repair plasmids MAVS	Santa Cruz	Cat: #sc-400769-HDR-2
Homology Directed Repair plasmids RIG-I	Santa Cruz	Cat: #sc-400812-HDR
Control plasmid CRISPR/Cas9	Santa Cruz	Cat: #sc-418922
CRISPR–Cas9-expressing knockout plasmid DUSP11	Santa Cruz	Cat: #sc-408162

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
p2RZ plasmid	Addgene	Cat: #27664
p2RZ RNY4 plasmid	This study	N/A
p2RZ RNY4ds3 plasmid	This study	N/A
pISRE-Luc	Agilent (Stratagene)	Cat: #219089
pIFNβ-Fluc (IFN-b-pGL3)	Lin et al. (2000)	N/A
pTK-Rluc	Promega	Cat: #E2231
pCIneo	Promega	Cat: #E1841
pCMV-RNY4	This study	N/A
pU6-RNY4	This study	N/A
Software and algorithms		
Attune NxT	Thermo Fisher Scientific	version 2.6
FlowJo	Tree Star	version 10.0.8
Prism	GraphPad Software	version 9.2
FACS aria	BD biosciences	N/A
R		version 3.5.1
Bioconductor package DESeq2	Love et al. (2014)	version 1.20.0
Raw Illumina reads trimmed using trim_galore	Babraham bioinformatics	N/A
NGS reads bioinformatic analysis with cutadapt	Martin (2011)	version 1.18
NGS reads bioinformatic analysis with STAR	Dobin et al. (2013)	versions 1.2.2 and 2.6.1c
featureCounts	Liao et al. (2013)	version 1.4.6-p3
RNAMotif tool	Perreault et al. (2007)	N/A
Other		
HiSeq2500 sequencer	Illumina	N/A
CFX384 Touch Real-Time PCR Detection System	BIO-RAD	N/A

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Nicolas Vabret (nicolas.vabret@mssm.edu).

Materials availability

The plasmids and cell lines generated in this study are available upon request.

Data and code availability

Raw sequencing data and expression count matrices have been deposited at GEO: GSE134861 and GSE203128, Accession numbers are also listed in the key resources table. This paper does not report original code.

Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cells lines

HEK293 (293, ATCC CRL-1573) and HEK293T (293T, ATCC CRL-3216) cells were maintained in DMEM-Glutamax supplemented with 10% heat-inactivated fetal calf serum (FCS, ThermoFisher Scientific) and Penicillin-Streptomycin (PS, Life Technologies). Jurkat T cells (Gift from Brown laboratory, Mount Sinai), MT4C5 cells (a derivate MT4 cells expressing CCR5) were used for co-culture with 293-4x4 and were cultured as described in (Lepelley et al., 2011). Primary T cells were maintained in RPMI supplemented with 5% pooled human serum (Gimini

Bio-products) and HEPES buffer, non-essential amino acids, PS and L-glutamine (all Life Technologies). One-STrEP-tagged RLRs (ST-RLR: ST-RIG-I, ST-MDA5, ST-LGP2), CHERRY (ST-CH) cells (described in (Sanchez David et al., 2016)) and STING-37 cell line corresponded to HEK293 cells stably transfected with an ISRE-luciferase reporter gene (described in (Lucas-Hourani et al., 2013)) were maintained in DMEM-Glutamax supplemented with 10% heat-inactivated FCS and 100 U/mL/100 mg/mL of PS and G418 (SIGMA) at 400 µg/mL. HAP1 RIG-I ko., LGP2 ko., MDA5 ko., MAVS ko. and control cell lines were purchased from Horizon Discovery (cat# HZGHC001441c001, HZGHC002927c011, HZGHC001448c012, HZGHC001456c011 and C631, respectively) and maintained in Iscove's Modified Dulbecco's Medium (ThermoFisher Scientific) with 10% FCS and PS. In order to generate ST-RLR cells susceptible to HIV-1 infection, they were transduced with lentiviral vectors encoding HIV-1 receptor (CD4) and co-receptor (CXCR4). After transduction, cells were sorted for the high level of expression of both CD4 and CXCR4 receptors. These cell lines were assigned ST-RLR-4X4. Stable cell line (assigned ST-CH-4X4) expressing the Cherry protein instead of tagged RLRs was generated and used as a negative control to allow subtraction of non-specific RNA binding.

METHOD DETAILS

Generation of CRISPR-edited cell lines

293T knock out cell lines were generated by cotransfection (lipofectamine 2000, Invitrogen) of CRISPR–Cas9-expressing knockout plasmids (MAVS: sc-400769-ko-2, RIG-I: sc-400812, both Santa Cruz) and Homology Directed Repair plasmids containing puromycin resistance gene (MAVS: sc-400769-HDR-2, RIG-I: sc-400812-HDR, both Santa Cruz). The knockout plasmids are a mixture of three plasmids, each carrying a different guide RNA specific for the target gene, as well as the Cas- and GFP-coding regions. 72 h after transfection cells were treated with puromycin (Invivogen, 1 µg/mL) for 1 week 293 and 2934x4 knock-out clones were generated by transfection (lipofectamine 2000, Invitrogen) of CRISPR–Cas9-expressing knockout plasmids (MAVS, sc-400769-ko-2 and control, sc-418922, both Santa Cruz). ST-RIG-I and ST-CH DUSP11 knock-out cells were generated by transfection (lipofectamine 2000) of CRISPR–Cas9-expressing knockout plasmids (DUSP11, sc-408162; control, sc-418922, both Santa Cruz). 72 h after transfection cells were treated with puromycin (Invivogen, 1 µg/mL) for 1 week. Jurkat knock-out clones were generated by electroporation (Neon Transfection System, Thermo Fisher Scientific) of CRISPR-Cas9-expressing knock-out plasmid (DUSP11, sc-408162; control, sc-418922, both Santa Cruz). 48 h after transfection (293 & 2934x4) or electroporation (Jurkat), GFP^+ cells were selected by cell sorting, and single clones were isolated in 96-well plates then cultured for 2 weeks. Depletion of target proteins was verified by western blotting.

Affinity chromatography of RLR-RNP complexes and RNA extraction

ST-RLR cells were infected with MV (MOI = 1) or DV-4 (MOI = 0.5) for 24 h or left uninfected (NI). In the case of HIV-1, ST-RLR4X4 (tagged) cells were co-cultivated with HIV-1-infected cells MT4C5 cells as described in (Lepelley et al., 2011). Briefly, 5×10^7 MT4C5 (donor cells) were exposed 150 ng (equivalent p24) of HIV-1 NL4.3 for 2 h at 37 °C. After washing the virus, the cells were grown for 48 h. The infection was monitored by flow cytometry analysis by intracellular Gag staining. Infection was then performed via co-culture of ST-RLR4X4 cells and MT4C5 cells at a donor:target cell ratio of 1:1. 24 h post infection, cells were lysed and affinity purification of ST-tagged proteins and RNA extraction was performed as described in (Chazal et al., 2018; Sanchez David et al., 2016).

Isolation of primary cells

Isolation of T cells from healthy donors

PBMCs were prepared by Ficoll (GE Healthcare) gradient centrifugation from buffy coats received from New York blood center (Long island city, NY, USA). Buffy coats were diluted in 1:2 ratios (v/v) with PBS and 30 mL of the diluted buffy coats were loaded on 15 mL Ficoll in 50 mL falcon tubes. The tubes were centrifuged for 25 min at 2000 rpm at low acceleration and break. Mononuclear cells were collected and pooled from the tubes and washed twice by centrifuging. $CD4^+$ T cell isolation was performed through beads-mediated negative selection (EasySep Human $CD4^+$ T Cell Isolation Kit, Stemcell Technologies) and $CD4^+$ cell purity was assessed by flow cytometry.

Isolation of T cells from HIV-1 cohort patients

Frozen PBMC from a cohort of intravenous drug using HIV-1+ patients (described in (Markowitz et al., 2014)) were thawed at 37°C and dead cells were removed through Annexin V beads-mediated selection

(EasySep Dead Cell Removal Annexin V Kit, Stemcell technology). $CD4^+$ T cells were further isolated through beads-mediated negative selection (EasySep Human $CD4^+$ T cell Isolation Kit, Stemcell technology) and resuspended in Trizol (Invitrogen) (a small fraction was resuspended in PBS to check viability and purity by flow cytometry). After addition of chloroform and phase separation, the top aqueous phase was used to subsequently isolate cellular RNA and the bottom organic phase was used to purify cellular proteins. Metadata of the cohort patients is listed in Table S3.

In vitro transcription

In vitro transcribed (IVT) RNAs were generated using T7 RiboMAX express large-scale RNA production system (Promega), using oligoDNA containing the sequence of interest behind a T7 promoter. For full-length Y RNAs, dsDNAs covering the entire sequence were used as templates. For RNY4 substructures fragments, a single oligo corresponding to the specific cDNA sequence of interest was annealed to another sense oligo containing the T7 promoter sequence, generating a DNA molecule as template where only the T7 promoter sequence was dsDNA. Sequences of DNA strands are listed in Table S4 below, with T7 sequence in bold. T7 reaction mixed where then treated with DNAse to remove DNA template and purified using RNeasy kit (Qiagen). When specified, IVT RNAs were additionally treated with Antarctic phosphatase (NEB) to remove their 5′-PPP moieties then repurified. When indicated, RNAs were also generated from *in vitro* transcription of a modified p2RZ plasmid (Addgene #27664) where RNY4 or RNY4ds3 sequences were cloned in 5′ of a T7 promoter, with the HDV Ribozyme sequence placed directly in 3′ of RNY4/RNY4ds3.

Luciferase-based reporter assay

ISRE & IFN-β promoter reporter assays

293T cells, 293, 293-4x4 or HAP1 cells were seeded in 24-well plates. 24 h later, reporter plasmids *p*-ISRE-Fluc (containing five ISRE promoter sequences upstream of the Firefly luciferase gene) or pIFNβ-Fluc (containing the IFN-β promoter upstream of the Firefly luciferase gene), and pTK-Rluc (containing a thymidine kinase promoter upstream of the Renilla luciferase gene) were transfected at a concentration of 100 ng/mL and 10 ng/mL, respectively. For experiments measuring responses to *in vitro* transcribed RNAs, plasmids were transfected together with 30 ng/mL of RNA of interest using Lipofectamine 2000 (Invitrogen). For experiments measuring responses to virus infection, cells were infected 24 h later with MV (MOI 1) or DV-4 (MOI 0.5), or co-cultured with MT4C5 infected with HIV-1 at a donor:target cell ratio of 1:1. 24 h later cells were lysed (Passive Lysis buffer, Promega) and Firefly luciferase and Renilla Luciferase activities were measured using Dual Luciferase Reporter Assay system (Promega). Renilla values were used as transfection normalization control. Short 5′-PPP bearing RNA molecules (obtained from pCIneo linearized with XbaI before transcription), low molecular weight poly(I:C), and high molecular weight poly(I:C) (both from Invivogen) were used as positive control of activation at a concentration of 10, 10 and 30 ng/mL, respectively.

STING-37 assay

STING-37 cells, corresponding to HEK293 cells stably transfected with the ISRE-luciferase reporter gene (described in (Lucas-Hourani et al., 2013)) were plated in 24-well plates. 24 h later, cells were transfected with 5-20 ng/well of *in vitro* transcribed RNA using lipofectamine 2000 (Invitrogen). 24 h after transfection, cells were lysed (Passive Lysis buffer, Promega) and Firefly luciferase activity was measured using the Bright-Glo Luciferase Assay System (Promega).

Infection with virus/transduction with vector

$HIV_{NL4.3}$WT and $HIV_{NL4.3}\Delta$VPR (described in (Hubner et al., 2009; Laforge et al., 2013)) were from Olivier Schwartz lab. $HIV_{NL4.3}$-GFP and $HIV_{NL4.3}$-mcherry encode sfGFP or mCherry, respectively, in place of nef, and nef expression is rescued with insertion of an internal ribosome entry site (Law et al., 2016). Both were from Benjamin Chen lab. $HIV_{NL4.3}$-GFP ΔVPR (this article) was produced by cloning the BamH1 and XhoI fragment from $HIV_{NL4.3}$-GFP into the corresponding backbone of $HIV_{NL4.3}\Delta$VPR. Positive clones were confirmed by enzymatic restriction and Sanger sequencing. A single clone was picked and amplified for subsequent experiments. Each virus was produced through transfection (Polyjet, Signagen) of 293T cells with plasmids coding for full-length viral genomes.

48 h after transfection, supernatants were collected, spun down and filtered to removed cellular debris and used to infect (HIV) or transduce (lentiviral vectors) T cell (Jurkat or Primary cells) by spinoculation in 96-wells

plate (1,200 x g, 90 min at 16°C) with polybrene (4 μg/mL, EMD Millipore). 6 h after infection/transduction, supernatants were replaced with fresh medium. In the case of primary cells, CD4 T cells were activated using PHA-L (2 μg/mL) for 72 h prior to infection.

When specified, HIV-GFP productively infected cells were sorted based on the GFP expression on a FACS aria (BD biosciences) with biosafety cabinet facility. The cells were stained for CD3 and CD4 expression to identify CD4 T cell populations and CD8, CD14, CD19 and CD56 to identify other contaminating immune cell populations. Viable cells were discriminated with a viability dye (blue fluorescent dye, Thermofischer). $CD3^+CD4^{+/dim}$ GFP^+ cells were sorted as infected and the $CD3^+CD4^+GFP^-$ cells were sorted as non-infected to the collection tube and used for RNA and protein isolation. The MV-Schwarz vaccine strain (GenBank accession no. AF266291.1) has been previously described (Combredet et al., 2003). DV-4 strain Dominica (AF326825) (Mackow et al., 1987) was obtained from the Centro de Ingeniería Genética y Biotecnología (CIGB), Cuba.

Differential 5′-PPP RNA digestion

1ug of total cellular RNA was treated with RNA 5′ polyphosphatase (enzyme that converts 5′-triphosphorylated RNA into 5′-monophosphorylated RNA, Lucigen) for 30 min at 37°C or mock-treated. RNAs were then purified and treated with Terminator 5′-Phosphate-Dependent Exonuclease (processive 5′ to 3′ riboexonuclease that specifically digests RNA with 5′-monophosphate ends, Lucigen) for 90 min at 30°C or mock-treated. Resulting RNAs were then purified and processed for qPCR analysis.

RNA purification

Total cellular RNA was extracted using Trizol followed by RNA purification from the aqueous phase using a modified version of the Zymo RNA Clean and Concentrator (Zymo Research), through the addition of 2x volumes of ethanol to increase the retention of small RNA species. RNA was then subjected to DNAse digestion (TURBO DNA-free Kit, Invitrogen) then purified again using Zymo RNA Clean and Concentrator and finally resuspended in DPEC-treated water.

Quantitative PCR

1ug of total RNA was converted to cDNA using RNA to cDNA EcoDry Premix Double Primed (Clontech) and resulting cDNA was diluted 10X in water. For differential enzymatic digestion analysis, qPCR reactions were carried out in 10 μL reaction volumes with 5 μL of TaqMan Fast Advanced Master Mix (Thermo Fisher Scientific), 2 μL of Primers/Probe mix and 3 μL of each cDNA sample. For all other analysis, qPCR reactions were carried out in 10 μL reaction volume with 5 μL iTaq Universal SYBR Green Supermix (Bio-rad), 2 μL Primer mix and 3 μL of each cDNA sample. The qPCR reactions were run using a CFX384 Touch Real-Time PCR Detection System (Bio-Rad) in clear wells plates. Targets amplification were quantified using the $\Delta\Delta C_t$ method relative to β-actin or to GAPDH. The list of the primers used in this study is provided in Table S5.

Western blotting

Whole-cell lysates were resuspended in Laemmi sample buffer (Bio-rad) completed with 10% β-mercaptoethanol and heated for 5 min at 95°C. 10-15ul of lysates were loaded onto 10% or 4-12% mini-protean TGX gels (Bio-rad) and the gel was run in Tris/Glycine/SDS buffer (Bio-rad). Proteins were transferred to 0.45 um PVDF membranes (Immobilon) in Tris/Glycine buffer (Bio-rad) supplemented with 20% methanol. Membranes were blocked in Tris-buffered saline (Bio-rad) plus 0.1% Tween 20 (Fisher) (TBS-T) containing 5% non-fat dry milk for 30 min at room temperature followed by overnight incubation with primary antibody at 4 °C. Membranes were then washed with TBS-T and incubated with HRP-conjugated secondary antibodies for 3 h at room temperature. Membranes were then washed and HRP was activated with ECL Plus Western Blotting Substrate (Pierce) for 5 min before being exposed to CL-Xposure Film (Thermo scientific). Relative HRP signals were quantified using image Lab software (Bio-rad), relative to GAPDH or Tubulin controls.

Antibodies

Western Blot and protein purification

The following antibodies were used: HIV-1 p24 (mouse monoclonal, Abcam, clone 39/5.4A, #ab9071); MAVS (rabbit polyclonal, Cell Signaling Technology - CST, #3993); DUSP11 (rabbit polyclonal, Proteintech,

#10204-2-AP); α-tubulin (mouse monoclonal, Proteintech, clone 1E4C11, #66031-1-Ig); GFP (rabbit monoclonal, CST, clone D5.1, #2956); GAPDH (rabbit monoclonal, CST, clone 14C10, #2118); VPR (rabbit polyclonal, Proteintech, # 51143-1-AP), StrEP-Tag (mouse monoclonal, Qiagen, #34850), UNG (mouse monoclonal, OTI2C12, Origene). Peroxidase-conjugated secondary antibodies against rabbit IgG (#7074) and mouse IgG (#7076) were purchased from CST.

Flow cytometry

The following antibodies were used: anti-p24-FITC (mouse monoclonal, Beckman and Coulter, #KC-57); anti-CD3-Pacific Blue (mouse monoclonal, Biolegend #300330); anti-CD4-PE (mouse monoclonal, Biolegend, #300539); anti-CD19-APC (mouse monoclonal, Biolegend #302212); anti-CD14-APC (mouse monoclonal, Biolegend, #325608); anti-CD56-APC (mouse monoclonal, Biolegend, #318310); anti-CD8-PerCP/Cy5.5 (mouse monoclonal, Bdbiosciences, #341051).

Purification of RLRs and RNA extraction with crosslinking

Three 15-cm2 tissue culture dishes per cell line were pretreated with 0.1 mg/mL poly-L-Lysine-hydrobromide (Sigma), rinsed with distilled water and PBS before plating the cells. Cells (20x106) were plated per dish in 20 mL of DMEM medium for 24 h before infection with MV at MOI1. 24 h post-infection plates were rinsed twicewith ice-cold PBS, crosslinked at 400 mJ/cm^2 in 10 mL of ice-cold PBS/plate and cells were then scraped, pelleted and resuspended in 2 mL of MOPS lysis buffer (20mMMOPS-KOH pH 7.4, 120mMKCl, 0.5% Igepal, 2mMbeta-mercaptoethanol, supplemented with protease inhibitors mixture and RNasin at 0.2 U/mL and protease inhibitors mixture (Roche). Cell lysates were incubated on ice for 20 min with gentle mixing every 5 min and then clarified by centrifugation at 16000xg for 15 min at 4 °C. Streptactin Sepharose beads (GE Healthcare, 100 μL/dish) were washed in MOPS washing buffer (20 mM MOPS-KOH pH 7.4, 120 mM KCl, 2 mM beta-mercaptoethanol, supplemented with RNasin 0.2 U/mL and protease inhibitors mixture) and finally resuspended in 1 mL of MOPS lysis buffer per initial culture dish. Clarified cell lysate was incubated with Streptactin beads for 2 h at 4 °C. The beads were washed three times with MOPS washing buffer and centrifuged at 1600xg, 5 min at 4 °C. Strep-tagged proteins were then eluted twice for 15 min at 4 °C in 250 μL/dish of 1X elution buffer (IBA, Biotin Elution Buffer 10X). Each sample was treated with proteinase K (Roche) in v/v of 2X proteinase K buffer (200 mM Tris pH 8, 100 mM NaCl, 20 mM EDTA, 4M urea) for 20 min at 4 °C that has been preincubated 20 min at 37 °C to remove RNase contamination. RNA purification was performed using TRI Reagent LS (Sigma). RNA was dissolved in 50 μL of DNase-free and RNase-free ultrapure water. Extracted RNAs were analyzed using Nanovue (GE Healthcare) and Bioanalyser total RNA nano/pico kit (Agilent).

QUANTIFICATION AND STATISTICAL ANALYSIS

RNA-seq analysis of total and RLR-bound RNA

Protocols for NGS library preparation and NGS of total and RLR-bound RNA from MV, DV-4-infected cells have been described in (Chazal et al., 2018; Sanchez David et al., 2016). Before RNA-seq analysis of total and RLR-bound RNA from HIV-1- and mock-infected cells, depletion of ribosomal RNA was done using the riboZero reagents included in the TruSeq stranded total RNA library prep kit (#20020596, Illumina). NGS libraries were generated following the manufacturer's protocol. The indexed samples were multiplexed per 4 or 6 and sequenced on a HiSeq2500 sequencer (Illumina) to produce single-ends 65 bases reads, bearing strand specificity.

Reads were cleaned of adapter sequences and low-quality sequences using cutadapt (Martin, 2011) version 1.11. Only sequences at least 25 nt in length were considered for further analysis. STAR version 2.5.0a (Dobin et al., 2013) with default parameters, was used for alignment on the reference genome (Human genome hg19 from UCSC). Genes were counted using featureCounts version 1.4.6-p3(Liao et al., 2014) from Subreads package (parameters: -t gene -g ID -s 1). For statistical analysis of NGS data, count data were analyzed using R version 3.5.1 and the Bioconductor package DESeq2 version 1.20.0 (Love et al., 2014). The normalization and dispersion estimation were performed with DESeq2 using the default parameters and statistical tests for differential expression were performed applying the independent filtering algorithm. For each virus, a generalized linear model including the replicate, beads and protein factors as well as the beads x protein interaction was set in order to test for the differential expression between the biological conditions. For each pairwise comparison, raw p-values were adjusted for multiple testing according to the Benjamini and Hochberg (BH) procedure (Benjamini and Hochberg, 1995) and genes

with an adjusted p value lower than 0.05 were considered differentially expressed. Bioinformatics analysis of NGS reads for viral reads was performed as described in (Chazal et al., 2018). The MV-Schwarz vaccine strain (AF266291.1), DV-4 strain Dominica (AF326825) and HIV-1 NL4.3 (AF324493.2) were used as references. The data have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Super-Series accession number GSE134861.

5′-PPP RNA sequencing

After RNA purification, ribosomal RNA molecules were depleted using the riboPOOL Kit for human (siTOOLs Biotech). To control for enrichment bias during library preparation, an aliquot of ERCC RNA Spike-In Mix 1 (ThermoFisher) was treated with 5′ RNA polyphosphatase (Epicentre) and mixed with untreated spike-in mix in approximately equal amounts, to form a 5′-PPP/5′-P 1:1 ratio. 1 μL of a 1:10 dilution was added to each total RNA preparation. RNA samples were first poly(A)-tailed using poly(A) polymerase. Then, the 5′ Illumina TruSeq sequencing adapters, which carry sequence tags ATTACTCG and TCCGGAGA were ligated to the 5′ mono-phosphate groups (5′-P) of sampled transcripts. Remaining fragments with unligated 5′-P ends are removed with 5′-phosphate-dependent XRN-1 (NEB). The samples were then treated with 5′ RNA polyphosphatase (Epicentre) to convert 5′ triphosphate structures into 5′ monophosphate ends. To the newly formed 5′-P groups the 5′ Illumina TruSeq sequencing adapters were ligated, which carry sequence tag CGCTCATT and GAGATTCC. First-strand cDNA synthesis was performed using an oligo(dT)-adapter primer and the M-MLV reverse transcriptase. After sequencing, reads corresponding to RNA molecules initially harboring 5′-PPP and 5′-P moieties were inferred from sequence tags. Library preparations and sequencing were performed at Vertis Biotechnology. After library preparation, sequencing and reads alignments, RNA totaling less than an average of 100 RNA reads per sample were discarded. Individual RNA ratios of 5′-PPP/5′-P forms were calculated in each sample. Additional intra-sample normalization was performed using the median ratio of ERCC RNA spike-in mix. Percentage of 5′-PPP fractions were then calculated from normalized ratio and fold changes of the average of 5′-PPP percentage triplicates were computed using DESeq2 package (Love et al., 2014). The data have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Super-Series accession number GSE203128.

RNA seq analysis (repetitive elements)

Raw Illumina reads were trimmed using trim_galore (Babraham bioinformatics) and cutadapt (Martin, 2011) version 1.18 with default settings. Reads were then mapped to the human genome (gencode annotation, build 38) and to repbase elements (release 20) using STAR aligner (Dobin et al., 2013) version 2.6.1c. Aligned reads were assigned to genes using the featureCounts function of subread package using the annotation (Liao et al., 2013). This produced the raw read counts for each gene. Mapping and counting of the reads were done in two stages. First, reads were mapped to the human genome, and the counts were determined using the gencode annotation and the annotation derived from the repeatmasker output. Second, the reads which were not assigned to any feature in either gencode or repeatmasker annotation were re-aligned to the repeat consensus sequence (repbase). Counts obtained from repeatmasker and repbase corresponding to the same family were added together. Differential expression analysis was performed using DESeq2 package (Love et al., 2014).

Polymerase-III transcript annotation

Identification of regions of genome transcribed by polymerase-III (Pol3) remains a topic of active research (Turowski and Tollervey, 2016; White, 2011). In order to check if a given transcript is transcribed by Pol3 for our analysis, we created a curated list of tentative ncRNA transcripts that are likely transcribed by Pol3. In the list, we included all ncRNA that are known to be transcribed by Pol3 (White, 2011) as well as ncRNA from hg19 genome assembly which overlap or are in proximity of an annotated Pol3 binding site or known Pol3 transcript, as based on curation of published datasets from Pol3 transcription studies from the following references (Barski et al., 2010; Canella et al., 2010; Moqtaderi et al., 2010; Oler et al., 2010; Raha et al., 2010).

Modeling RNY4 and RNY1-like structure in transcripts

We follow the computational approach originally conceived in the work (Perreault et al., 2007) to find Y RNA homologs in genomes and used the RNAMotif tool to identify motifs that fold into a Y RNA like structure. The RNAMotif software searches given RNA sequences for regions that are able to fold into a specified secondary structure. It identifies all regions in the RNA sequence capable of adopting the specified structure

and calculates the free-energy contribution of the RNA region folded into the structure using the nearest-neighbor model for RNA (Mathews et al., 1999). For our study, we use the following constraints for the "RNY4" structure search with RNAMotif. The motif is required to consist of (annotation illustrated as in Figure S2B): stem (S1) of length ranging from 5 to 13 base pairs, a loop (L1) (with 0 to 2 unpaired nucleotides on 5′ end side and 1 to 3 nucleotides on 3′ end side), followed by a stem (S2) of length from 7 up to 11 base pairs, with a loop (L2) of size 6 to 17 unpaired nucleotides on the 5′ end side and 6 to 17 nucleotides on the 3′-end side, and a stem (S3) of length 7 to 17 base pairs, with a terminal hairpin loop of length ranging from 3 to 10 bases. For RNY1 structure: stem (S1) ranging from 5 to 13 base pairs, a loop (L1) (with 1 to 6 unpaired nucleotides on the 5′ side and 1 to 7 nucleotides on the 3′ end side), followed by a stem (S2) of length from 7 up to 11 base pairs, then a terminal loop (L2) containing, from 5′ end to 3′ end: a single-stranded stretch of 4 to 10 nucleotides, then a hairpin with a stem of 7 to 12 nucleotides and a terminal loop of 3 to 6 nucleotides, a single-stranded stretch of 5 to 11 nucleotides, a second hairpin with a stem of 4 to 6 nucleotides and a terminal loop of 3 to 9 nucleotides and finally a single-stranded stretch of 12 to 23 nucleotides. Specifically, the motifs tested do not include the presence of a nearly invariant bulged helix that is the binding site of the Ro60 protein and an ending with three U that is the binding site of the La protein.

In each stem, we allow up to 1 mismatch, and the stem base-pairs can contain both Watson-Crick and wobble base pairs. We only search for presence of the motif at the 5′ start of the transcripts, so for the evaluated sequences, we only consider the motif to be present if a possible RNY4-like structure (with assigned free energy by Turner model (Mathews et al., 1999) smaller than 0) is detected by RNAMotif within 6 bases from the 5′-end of the transcript.

The sequence datasets evaluated for RNY4 motif presence were the following: human cDNA and non-coding RNA sequences (from hg38 reference genome assembly), complete genomes of positive-sense viruses with human host (Table S6, obtained from NCBI viral genome database (Brister et al., 2015)), and inserts in genome that were annotated as belonging to Y RNA family in the repetitive DNA element database (Hubley et al., 2016). For each sequence, we also constructed a scrambled sequence, which was obtained by randomly permuting all nucleotides in the respective sequence, so that the frequency distribution of respective nucleotides remains the same, but their order is random. The set of scrambled sequences was also used to search for RNY4-like motifs.