JVI Journals.ASM.org

Large-Scale Nucleotide Optimization of Simian Immunodeficiency Virus Reduces Its Capacity To Stimulate Type I Interferon *In Vitro*

Nicolas Vabret,[a,d] Marc Bailly-Bechet,[b] Alice Lepelley,[c] Valérie Najburg,[a] Olivier Schwartz,[c] Bernard Verrier,[d] Frédéric Tangy[a]

Unité de Génomique Virale et Vaccination, CNRS UMR-3569, Institut Pasteur, Paris, France[a]; Laboratoire de Biométrie et Biologie Evolutive, CNRS UMR-5558, Université Claude Bernard Lyon 1, Villeurbanne, France[b]; Unité Virus et Immunité, Institut Pasteur, CNRS UMR-3569, Paris, France[c]; Laboratoire de Biologie Tissulaire et Ingénierie Thérapeutique, UMR-5305, University of Lyon 1, Lyon, France[d]

ABSTRACT

Lentiviral RNA genomes present a strong bias in their nucleotide composition with extremely high frequencies of A nucleotide in human immunodeficiency virus type 1 (HIV-1) and simian immunodeficiency virus (SIV). Based on the observation that human optimization of RNA virus gene fragments may abolish their ability to stimulate the type I interferon (IFN-I) response, we identified the most biased sequences along the SIV genome and showed that they are the most potent IFN-I stimulators. With the aim of designing an attenuated SIV genome based on a reduced capacity to activate the IFN-I response, we synthesized artificial SIV genomes whose biased sequences were optimized toward macaque average nucleotide composition without altering their regulatory elements or amino acid sequences. A synthetic SIV optimized with 169 synonymous mutations in *gag* and *pol* genes showed a 100-fold decrease in replicative capacity. Interestingly, a synthetic SIV optimized with 70 synonymous mutations in *pol* had a normal replicative capacity. Its ability to stimulate IFN-I was reduced when infected cells were cocultured with reporter cells. IFN regulatory factor 3 (IRF3) transcription factor was required for IFN-I stimulation, implicating cytosolic sensors in the detection of SIV-biased RNA in infected cells. No reversion of introduced mutations was observed for either of the optimized viruses after 10 serial passages. In conclusion, we have designed large-scale nucleotide-modified SIVs that may display attenuated pathogenic potential.

IMPORTANCE

In this study, we synthesized artificial SIV genomes in which the most hyperbiased sequences were optimized to bring them closer to the nucleotide composition of the macaque SIV host. Interestingly, we generated a stable synthetic SIV optimized with 70 synonymous mutations in *pol* gene, which had a normal replicative capacity but a reduced ability to stimulate type I IFN. This demonstrates the possibility to rationally change viral nucleotide composition to design replicative and genetically stable lentiviruses with attenuated pathogenic potentials.

RNA viruses carry their whole genetic information on a single or double-stranded genomic RNA molecule. They have relatively small genomes, probably due to their high mutation rate (1) and to capsid size constraints (2). Their genomic RNA encodes few proteins, sometimes with overlapping segments, and contains many noncoding sequences (3). These regulatory sequences play critical roles in controlling transcription, translation, subcellular trafficking, or packaging but also in virulence functions such as immune evasion (4–6). Recently, viral RNA properties relying on global features such as nucleotide composition or codon usage have also been revealed to be crucial in virus biology (7, 8).

It was proposed 30 years ago that each species was subjected to specific genomic pressures on nucleotide composition resulting in a distinctive bias in synonymous codon usage (9). This is true for viruses, which have species-specific nucleotide compositions, with most RNA viruses displaying A-rich and C-poor codons in their coding strand (10). For example, the genomes of human immunodeficiency virus type 1 (HIV-1) and other lentiviruses are strongly biased in their nucleotide composition compared to their primate hosts, with as much as 35% adenosines (11, 12). As a consequence, both their average amino acid composition and their synonymous codon usage are different from those of their hosts (13, 14). Lentivirus-biased nucleotide composition has been explained by deoxynucleoside triphosphate (dNTP) pool imbalance during reverse transcription (15) and by antiviral activity of the cellular Apobec 3G (A3G) cytidine deaminase, which mutates G to A in HIV provirus (16) and is counteracted by the viral protein Vif (15, 17). The specific nucleotide composition of lentiviruses may impact genome structure and stability (18, 19), nuclease sensitivity (20), and viral RNA recognition by innate immunity receptors (21, 22).

To study the impact of global nucleotide composition or codon usage on virus biology, it is necessary to rely on chemical synthesis of large genomic fragments and on reverse genetics to generate modified viruses. In the last years, several groups have synthesized large artificial viral sequences to proceed to genome-scale modifications. For example, the effect of shifting poliovirus codon usage on its replication has been investigated (8, 23). The poliovirus RNA genome was deoptimized without altering the amino acid

Received 31 October 2013 **Accepted** 20 January 2014
Published ahead of print 29 January 2014
Editor: G. Silvestri
Address correspondence to Frédéric Tangy, ftangy@pasteur.fr.
Supplemental material for this article may be found at http://dx.doi.org/10.1128/JVI.03223-13.

doi:10.1128/JVI.03223-13

sequence by changing synonymous codons from frequently to rarely used codons. This resulted in imbalanced synthesis of viral proteins and generated attenuated viruses. Attenuated polioviruses and influenza viruses were further developed by deoptimizing the codon pair bias (7, 24, 25). Both deoptimized viruses protected mice from infection after subsequent challenge with a wild-type virus, highlighting their potential as vaccine. In the case of HIV-1, it was observed that systematic replacement of wild-type codons by synonymous GC-rich codons in *gag* and *pol* genes led to a profound delay in replication kinetics with at least 5-fold loss of infectivity (26). Suppression of viral infectivity was caused by enhanced dimer stability of the viral RNA genome and subsequent reduction of viral cDNA synthesis (26). These examples demonstrate that genome-scale changes in viral sequence allow the design of attenuated vaccines that are genetically stable because of the large number of mutations involved.

We recently reported that HIV-1-biased nucleotide composition triggers overstimulation of the type I interferon (IFN-I) response after RNA transfection in human cells, indicating that RNA sequences are discriminated according to their nucleotide composition (27). Type I IFN is a major antiviral cytokine thought to contribute to chronic activation of the immune system and progression to AIDS during HIV infection (28, 29). We also proposed a putative link between pathogenicity and divergent nucleotide composition of HIV-1 compared to host (27). These results suggested a new determinant for the pathogenicity of lentiviral infections and raised the possibility of altering virus-host interactions by artificially changing the nucleotide frequency of the viral genome.

In the present work, we observed that codon optimization of viral genes, a technique commonly used to increase antigen expression in vaccine candidates (30–32), abolishes the capacity of viral RNA to induce IFN-I in human cells. Based on the demonstration that distinct regions of simian immunodeficiency virus (SIV) genomic RNA trigger different levels of IFN-I according to their nucleotide composition, we designed an attenuated artificial SIV genome by sequence optimization. Viral sequences were made closer to the average nucleotide composition of the SIV host macaque without altering the regulatory elements and the amino acid composition. The corresponding synthetic viruses were produced by reverse genetics and were analyzed for their ability to replicate and to stimulate IFN-I production *in vitro* in a human T-cell line and in human and macaque peripheral blood mononuclear cells (PBMCs).

MATERIALS AND METHODS

***In silico* design of SIVopt genomes.** To design new nucleotide-optimized SIV (SIVopt) genomes, we selected three modifiable regions within *gag*, *pol*, and *env* genes of the SIVmac239 genome in which nucleotide optimization was possible while preserving amino acid sequences and regulatory regions. For each region, the actual profile of chi-square divergence of nucleotide composition between the virus and the rhesus macaque (*Macaca mulatta*) host was computed with a sliding window of 101 nucleotides, as previously described (27). Then, a threshold of maximal acceptable chi-square divergence was selected, according to similar profiles computed with nonpathogenic lentivirus-host couples (SIVagm/African green monkey *Sabeus* and SIVsm/Sooty mangabey) or couples with reduced lentiviral pathogenicity (HIV-2/human and SIVcpz/chimpanzee). The selected thresholds were 0.04 in the *gag* region, 0.14 in the *pol* region, and 0.05 in the *env* region. To identify the nucleotide modifications that would reduce local divergence, each region was divided into peaks of divergence separated from nondivergent sections (see Fig. 2B). Inside each region, all genome positions with a chi-square value above the given threshold and distant less than 101 nucleotides from another were included in the same peak of divergence. The distance of 101 nucleotides corresponds to the window size for chi-square computation; thus, the local chi-square value at a given position depends only on the nucleotide composition of the 50 surrounding nucleotides, and two positions distant by 101 nucleotides are computed independently. This division of genomic regions into independent peaks of divergent and nondivergent sections allowed optimizing of the nucleotide content of each peak separately and quickened most computations. Inside each peak, all possible one-nucleotide changes synonymous for all coding sequences (CDS) for the position were ranked based on their impact on the local chi-square divergence. To minimize the number of mutations necessary to reduce the divergence to the predefined threshold, the single nucleotide modification causing the higher diminution of divergence was selected. After each modification, the chi-square profile of the peak was recomputed, and new modifications were made iteratively, until either no more genome position in the peak had a chi-square value above the threshold or all possible mutations in the peak had been realized, meaning that the selected threshold was unreachable. The next peak was then selected and solved iteratively in the same way until resolution of all peaks inside each region was achieved.

Cells and reagents. Human PBMCs were isolated from the blood of healthy donors by Ficoll centrifugation. The blood was provided by the EFS (Etablissement Français du Sang, the French Official Blood Bank). Macaque PBMCs were isolated from two Chinese rhesus monkeys (RMs) (*Macaca mulatta*) housed in single cages within the "Commissariat à l'Energie Atomique" (Fontenay-aux-Roses, France) facilities according to national guidelines. Whole blood was collected on sodium heparin. Human and macaque PBMCs and Cemx174 cells were grown in RPMI medium with 10% heat-inactivated fetal bovine serum (FBS). HeLa P4C5 (33) and HEK293T cells and derivatives were grown in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% FBS.

Virus production. Viruses were produced by transfection of infectious plasmids into HEK293T cells and cocultivation with Cemx174 cells. Mutated *gag*- and *pol*-optimized DNA sequences were chemically synthesized (GeneScript) and then cloned into plasmid p239SpSp5. A mutated *env*-optimized sequence was cloned in p239SpE3′ nef Open. Plasmids p239Sp5 and p239SpE3′ nef Open contain, respectively, the 5′ and 3′ halves of SIVmac239 infectious clone (34). Each plasmid (10 μg) containing the opt or the wild-type (wt) version of *gag*, *pol*, or *env* genes was digested with SphI (p239Sp5) or SphI + AatII (p239SpE3′ nef Open), purified on agarose gel, and ligated together with T4 DNA ligase for 48 h at 4°C. The total ligation product (40 μl) was transfected with 6 μl Lipofectamine (Invitrogen) in 8×10^5 HEK293T cells plated the day before on 6-well plates. Two days after transfection, DMEM was replaced by 3 ml RPMI–10% FBS containing 5×10^5/ml Cemx174 cells. Syncytium formation in the coculture indicated virus production. Two weeks after transfection, culture supernatant was harvested, centrifuged at low speed, and filtrated trough 0.45-μm pores. Viral stocks were kept at −80°C.

RNA preparation. Primers (see Table S1 in the supplemental material) were designed to amplify 40 overlapping fragments of approximately 500 bp by PCR (Enzyme Phusion; 35 cycles with amplification temperature [Ta] of 60°C and 30-s elongation). Primers were also designed to amplify sequences (1-min elongation) from wt or opt sequences of viral genes. A 5′ tail containing the T7 promoter sequence was included in every forward primer to allow the subsequent *in vitro* transcription reaction. PCR products were purified and used as the template for T7 RNA synthesis according to the manufacturer's instructions (T7 RiboMA Express; Promega). Resulting RNAs were purified using the RNeasy minikit (Qiagen), and concentrations were determined by Nanodrop measurement.

IFN-luciferase reporter assays. Expression of alpha/beta interferon (IFN-α/β) was determined by transient transfection of HEK293 cells with either reporter plasmid pISRE-Luc containing five interferon-stimulated

response elements (ISRE) upstream of the firefly luciferase gene (Stratagene) or reporter plasmid pIFNβ-Luc containing the firefly luciferase gene under the control of the IFN-β promoter (provided by R. Weil and J. Hiscott). For RNA transfections, HeLa P4C5 or HEK293T cells were plated in 24-well plates (2×10^5 per well). After 24 h, cells were transfected using 1 μl Lipofectamine 2000 (Invitrogen) with pISRE-Luc reporter plasmid (250 ng/well) (for HEK293T) or pIFNβ-Luc (for HeLa P4C5), a plasmid harboring a thymidine kinase (Tk) promoter upstream of the renilla luciferase gene (25 ng/well), and 12 ng of each RNA fragment (500 bp) or 20 ng of wt/opt viral genes. After 20 h, cells were lysed, and the firefly and renilla luciferase activities were measured in cell lysates using the Dual-luciferase Reporter Assay system (Promega) according to the manufacturer's instructions. Reporter activity was calculated as a triplicate of the ratio of firefly luciferase activity to the reference renilla luciferase activity.

For virus activation analysis during coculture, HeLa P4C5 cells were plated in 24-well plates (10^5 cells per well). One day later, cells were transfected using 1 μl Fugen (Roche) with pIFNβ-Luc reporter plasmid (200 ng/well) and a plasmid harboring a thymidine kinase promoter upstream of the renilla luciferase gene (20 ng/well). One day later, infected Cemx174 cells were added to the culture at a concentration of 10^6 cells per ml (Cemx174/HeLa P4C5 ratio, 1:1). The percentage of Cemx174 infection was assayed by SIV p27 staining and flow cytometry.

IFN-I detection. IFN-I secretion was quantified using the reporter cell line HL116, which carries the luciferase gene under the control of the IFN-inducible 6-16 promoter (35) (a kind gift from Sandra Pellegrini, Institut Pasteur, France). HL116 cells were grown in DMEM supplemented with 10% FBS and hypoxanthine-aminopterin-thymidine (HAT; H, 20 μg/ml; T, 20 μg/ml; A, 0.2 μg/ml). HL116 cells (2×10^4) plated 16 h prior to the assay in a 96-well plate were incubated for 7 h with the desired culture supernatants or standards containing a titration of human IFN-α2a (Immunotools). Cells were then lysed (Luciferase Cell Culture Lysis, 5× Reagent; Promega), and luciferase activity was measured using the Luciferase Assay reagent (Promega). Samples were analyzed using PerkinElmer Wallac 1420. IFN levels are expressed as the equivalent of IFN-α2a concentration in units/ml.

Flow cytometry staining. Cells were intracellularly stained with anti-SIV Gag p27 (Clone 55-2F12, NIH AIDS Research & Reference Reagent Program, from Niels Pedersen) and anti-human MxA (from O. Haller). Briefly, cells were fixed for 10 min with phosphate-buffered saline (PBS) 4% paraformaldehyde, washed, permeabilized, and stained for 45 min in PBS containing 1% bovine serum albumin (BSA) and 0.05% saponin. Isotype-matched monoclonal antibodies (MAbs) were used as negative controls. Samples were analyzed by flow cytometry using a FacsCalibur (Becton, Dickinson) or a FacsCanto II (Becton, Dickinson) instrument with FlowJo or FacsDIVA software.

Lentiviral transduction. HeLa P4C5 cells were transduced with a lentiviral vector (LV) expressing bovine viral diarrhea virus (BVDV)-Pro and previously described (36). The LV also expresses the puroR gene. Two days after transduction, HeLa P4C5 cells were selected in the presence of 1 μg/ml puromycin. Resistant populations grew in a few days and were used without further cloning.

WB analysis. One week after transduction and selection, 2×10^6 HeLa P4C5 cells were lysed in PBS–1% Triton X-100 (Sigma-Aldrich) supplemented with protease inhibitors (Roche). Cell lysates were analyzed by SDS-gel electrophoresis using 4 to 12% NuPAGE gels (Invitrogen). IFN regulatory factor 3 (IRF3) Western blot (WB) analysis was performed using rabbit anti-IRF3 (Clone FL-425; Santa Cruz). As a control, actin specific primary antibody was used.

Nucleotide sequence accession numbers. The optimized sequences determined in this study are available from GenBank under accession number KJ152770 for SIVopt1 and accession number KJ152771 for SIVopt2.

RESULTS

Codon optimization of viral RNAs abolishes their ability to induce IFN-I. Optimization of codon usage by introduction of host cell synonymous codons is a widely used means to improve recombinant protein expression for DNA-based (30, 31) or viral vector-based (32) vaccines produced in bacteria, yeasts, or plants (37). Genes with a codon usage matching the specific cellular tRNA abundance are the most highly expressed (38, 39). Interestingly, in mammalian cells, only a weak positive correlation between optimal codon usage and gene expression levels is found (40–43). Since we previously observed that the nucleotide composition of HIV-1 RNA can modulate IFN-I stimulation in human cells (27), we wondered whether codon optimization of various viral sequences would alter their potential immunostimulatory capacity.

We chose several genes of different sizes and functions from different viral species (*gag*, *pol*, and *env* from HIV-1, *gag* from SIVmac, *hemagglutinin* and *neuraminidase* from influenza viruses H5N1 and H1N1, *nucleoprotein* and *spike* from severe acute respiratory syndrome [SARS] coronavirus, and *core*, *E1*, and *E2* from hepatitis C virus [HCV]). These sequences were obtained from both wild-type viral cDNAs and commercial synthetic DNAs that were optimized for human codon usage (GeneScript). For each viral sequence, wild type (wt) and human codon-optimized (opt) versions were used as the template for *in vitro* transcription with T7 RNA polymerase into uncapped and unpolyadenylated RNA fragments. To ensure that no protein was expressed from these RNAs, the first nucleotide of the ATG codon was mutated in each construct. The capacity of these RNA fragments to stimulate IFN-I was determined by using a very sensitive method (27). RNA fragments were cotransfected into HEK293T cells together with a reporter plasmid expressing the luciferase gene under the control of five interferon-stimulated response elements (ISRE-luciferase) (44). The IFN-I response was determined by measuring luciferase activity 24 h after transfection (Fig. 1A). In this system, all wt genes (Fig. 1A, black bars) significantly induced IFN-I. However, differences in intensity were observed: RNA derived from HIV *pol* gene was the most potent stimulator of IFN-I, and RNAs from HCV were the least stimulatory. According to our previous results, human codon optimization of all viral genes reduced their capacity to stimulate IFN-I (Fig. 1A, gray bars), with the exception of HCV core RNA, whose optimization increased the luciferase signal. These data show that, regardless of their origin, RNAs derived from viral genes stimulate IFN-I, while codon optimization decreases this property (Wilcoxon paired unilateral test, $P = 10^{-3}$). We then measured IFN-I production by human PBMCs upon stimulation with the same RNAs complexed to DOTAP (N-[1-(2,3-dioleoyloxy)propyl]-N,N,N-trimethylammonium methylsulfate). IFN-I concentration in PBMC supernatants was measured 20 h after stimulation using a reporter cell line (35). Human codon optimization also reduced the ability of all genes tested to induce IFN-I in PBMC, albeit to a lesser extent than observed in HEK293T cells (see Fig. S1 in the supplemental material).

Codon optimization modifies both codon usage and global nucleotide composition. To evaluate the influence of nucleotide composition in our observations, we performed a principal-component factorial correspondence analysis on the nucleotide composition of all wt and opt RNA fragments (Fig. 1B). This analysis highlights that wt viral genes are either A or U rich, with the ex-

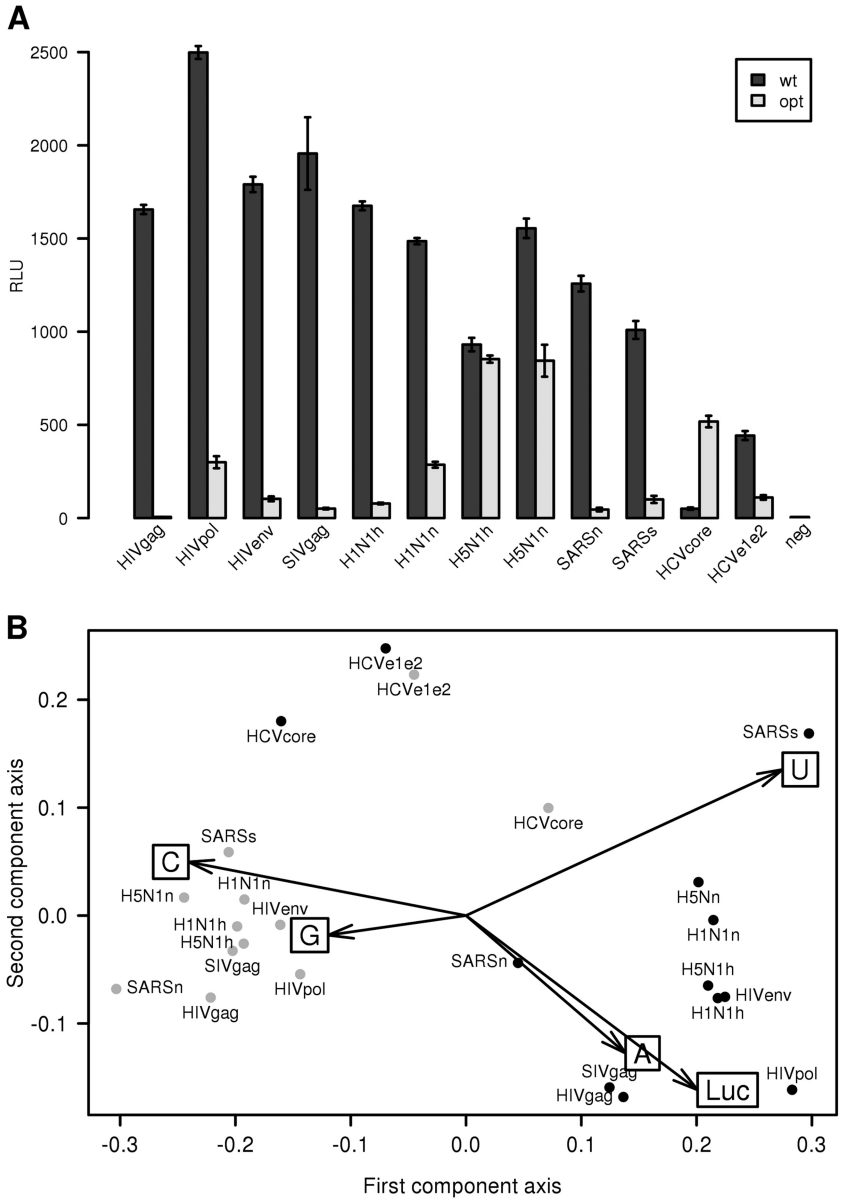

FIG 1 Codon optimization of viral RNAs modulates their ability to stimulate IFN-I. (A) Measure of luciferase expression reflecting IFN-I stimulation after transfection of RNA fragments transcribed *in vitro* from wild type (wt) or optimized (opt) viral sequences from different origins. RLU, relative luciferase units. Virus names are written in uppercase letters and viral gene names in lowercase letters. (B) Plot of the two principal components (96.2% of the total variance) of the correspondence analysis of the nucleotide composition of viral sequences. The positions of viral sequences depend on their composition (black dots correspond to wt sequences and gray dots to opt sequences). Projection of the A, C, G, and U components indicates which nucleotide is enriched in each direction on the graph. Projection of the luciferase expression (as in panel A) is also shown (Luc) and is highly correlated with A nucleotide frequency in the strains (Pearson correlation test, $R = 0.85$, $P = 1.4 \times 10^{-7}$), as can visually be assessed by the small angle between the A and Luc arrows.

ception of HCV genes, while most optimized versions are G/C rich. Comparison of nucleotide composition with luciferase activity showed that the A richness of RNAs correlated with IFN-I stimulation (Fig. 1B, $R = 0.856$, $P = 1.4 \times 10^{-7}$).

The most biased regions of the SIV genome are the most potent IFN-I stimulators. IFN-I is the principal mediator of antiviral innate immunity, and its sustained expression is a major difference observed in host responses between pathogenic and nonpathogenic lentiviral infections (28, 29). Viral RNAs with high content of A/U nucleotides are strong stimulators of IFN-I, and lentiviral genomes have a particularly A-rich composition. With the aim of designing an attenuated SIV genome based on a reduced capacity to activate IFN-I response, we investigated whether local A-rich regions of SIV genes were stronger stimulators of IFN-I. To analyze the repartition of IFN-stimulating sequences along the SIVmac239 genome, we measured the ability of a set of 40 overlapping RNA fragments of approximately 500 bp, covering the entire genome of SIVmac239, to induce IFN-I *in vitro* in HEK293T cells (Fig. 2A). Activating fragments were found mostly clustered in the *pol* region, the 5′ region of *Gag*, and the 3′ region of *env*. We then looked for the repartition of local nucleotide bias on the whole genome. We computed the chi-square dis-

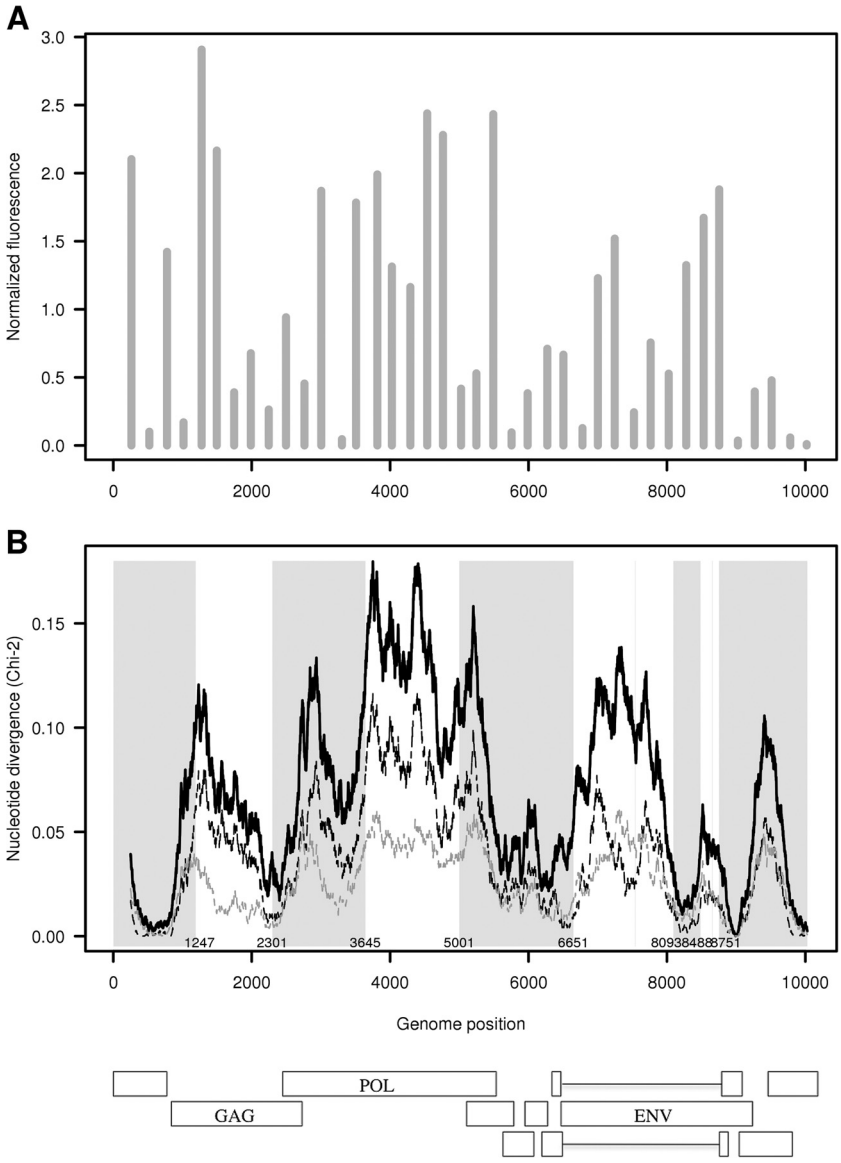

FIG 2 The nucleotide divergence of SIVmac239 RNA with that of the rhesus macaque correlates with the ability to stimulate IFN-I. (A) Luciferase reporter activity determined in HEK-293T cells cotransfected with SIVmac239 RNA fragments and pISRE-Luc reporter plasmid. (B) Repartition of nucleotide divergence along the SIVmac239 genome. The black line shows the nucleotide divergence of the SIVmac239 genome compared to the macaque rhesus genome, as measured by a chi-square distance in a 500-bp sliding window along the SIVmac239 sequence. Individual contributions of A and C nucleotides to this divergence are shown in black and gray dotted lines, respectively. Nucleotide divergence correlates with luciferase activity ($R = 0.36$, $P < 0.02$). White areas indicate regions modified by our optimization algorithm; the limits of the region are reported along the *x* axis. An SIVmac schematic genome map is presented at the bottom of the figure.

tance between the A/C/G/U frequencies of a sliding window 500 nucleotides (nt) wide along the SIVmac239 genome and the corresponding frequencies of the entire coding sequences of the macaque genome (Fig. 2B). This analysis shows that the most divergent regions locate within the three large viral genes (*gag*, *pol*, and *env*), while overlapping coding regions and *cis*-active regulatory sequences are not biased. Richness in A contributes to the majority of the observed divergence, with the exception of a short central portion of *env* where C paucity is the major contributor. The ability of each fragment to induce IFN-I (Fig. 2A) and the divergence to host in nucleotide composition (Fig. 2B) correlated significantly ($R = 0.36$, $P = 0.02$), as expected from our previous work on the HIV-1 genome (27).

Design of SIVmac239-optimized sequences. Based on this observation, we hypothesized that a synthetic SIV with a nucleotide composition optimized to be closer to its host would have a reduced capacity to stimulate IFN-I and might thus represent a model of attenuated macaque lentiviral infection. SIVmac239 is highly pathogenic in rhesus macaques (45). To attenuate its virulence, we optimized the SIVmac239 genomic sequence by changing its nucleotide content toward the macaque average composition. We used the original SIVMM239 sequence available at GenBank (accession number M33262). Our strategy was to reduce as much as possible the major peaks of nucleotide divergence, which are also the most potent stimulators of IFN-I, while preserving the amino acid sequence. Based on previous work showing

TABLE 1 Genomic regulatory sequences excluded from algorithm optimization in SIVmac239

Sequence(s) excluded	Function during viral replication	Reference(s)
Long terminal repeats	Regulation of transcription, translation, and genome packaging	66
First 300 nucleotides of Gag	Genome packaging in feline immunodeficiency virus- or SIV-derived vector	67, 68
Gag-Pol frameshift	Translation of Pol gene	33
5′ A-rich sequence of Pol	Important for cDNA synthesis during HIV-1 reverse transcription	26
Central polypurine tract and central terminal sequence	Regulation of reverse transcription	68
Rev responsive element	RNA subcellular trafficking	69, 70
Sequences containing donor and acceptor splicing sites	RNA maturation	71–73

that codon optimization of the HIV-1 genome may lead to defective virus (26), we excluded all known or putative regulatory sequences from optimized regions. From the literature available on the SIVmac239 genome and by homology with known regulatory sequences of the HIV-1 genome, we excluded from the optimized regions the sequences corresponding to the genomic features described in Table 1. Overlapping open reading frames (ORFs) that were impossible to optimize while maintaining amino acid sequences were unmodified. This led us to select three separate regions on the SIVmac239 genome to be optimized to the macaque average nucleotide composition. These regions are located within *gag*, *pol*, and *env* genes (Fig. 2A, white areas). Each region was divided into short domains of nucleotide divergence, and synonymous mutations were locally computed to reduce divergence. This method was applied iteratively on all divergent peaks until the divergence in each region was below a preselected threshold. We tested different thresholds inside each region and selected for synthesis and further experiments those leading to a total of 348 changes divided into 99 synonymous mutations in the *gag* region, 70 in the *pol* region, and 189 in the *env* region. The optimized sequences are available from GenBank as indicated above in Materials and Methods, and the statistics of mutations are presented in Table S1 in the supplemental material. Mutated fragments were chemically synthesized (GeneScript) and subcloned in replacement of wt counterparts into p239Sp5 and p239SpE3′ nef Open plasmids that contain, respectively, the 5′ and 3′ halves of the SIVmac239 infectious clone (34).

Optimization of *pol* or *gag/pol* regions leads to replicating viruses. Wild-type SIV (SIVwt) and optimized SIV strains were produced after transfection of reconstituted proviral DNA into HEK293 cells and cocultivation with Cemx174 cells. All combinations of wt and optimized *gag*, *pol*, and *env* fragments were tested. Replicating viruses could be obtained with genomes containing an optimized *pol* (SIVopt1) or optimized *gag* and *pol* (SIVopt2) (Fig. 3A). No replicating virus was obtained with optimized *env* or optimized *gag* alone. Viral stocks of SIVopt1, SIVopt2, and SIVwt were produced, and their replication kinetics were compared by infecting Cemx174 cells and macaque PBMCs with different viral doses. Viral replication was assessed by p27 immunostaining and flow cytometry (Fig. 3B and C). SIVwt and SIVopt1 displayed the same replicative kinetics in both Cemx174 and PBMCs, while SIVopt2 was greatly reduced in Cemx174 and did not replicate in PBMCs. In Cemx174 cells, a 100-fold-higher viral input was needed for SIVopt2 to reach viral peak on day 8 after infection than for SIVwt and SIVopt1. Because SIVopt2 had a reduced infectivity, it was not further evaluated. Conversely, the 70 synonymous mutations introduced in *pol* gene did not alter the replicative capacity of SIVopt1.

In the perspective of vaccine design, the stability of attenuated synthetic viruses is of uttermost importance to prevent reversion. We therefore evaluated the stability of the optimized virus genomes. SIVopt1, SIVopt2, and SIVwt were passaged 10 times on Cemx174 cells (cultured for 10 weeks, which represents approximately 30 replication cycles). The genomes of the resulting viruses were fully sequenced, and their consensus sequences were compared to their original sequences before passaging. A few mutations occurred in the three passaged viruses (listed in Table S2 in the supplemental material). One mutation common to the three passaged viruses was found in the 5′-long terminal repeat (LTR). Other mutations were found in different regions of the genome (see Table S2 in the supplemental material). However, and interestingly, no reversion was observed in the optimized regions of optimized viruses, indicating that these artificially introduced mutations are not less stable than the wild-type SIVmac239 sequence. The lack of reversion remains to be determined in primary cells such as macaque PBMCs.

SIVopt1 has a reduced capacity to stimulate IFN-I response in human PBMCs. To compare the capacity of SIVopt1 and SIVwt to stimulate IFN-I response in human PBMCs, we measured the induction of the MxA gene, an interferon-stimulated gene (46). As expected from the previous experiment, SIVwt and SIVopt1 grew at similar rates in human PBMCs. Percentages of SIV Gag p27-positive-infected cells and cellular MxA induction were assessed by immunostaining and flow cytometry analysis at different time points. To standardize with viral growth, MxA induction was calculated relative to the percentage of Gag p27-positive cells for both viruses (Fig. 4A). When comparing samples from 7 human PBMC donors, we found that SIVopt1 had a slightly reduced capacity to stimulate MxA expression compared to SIVwt (Fig. 4B, paired Wilcoxon test, $P = 3 \times 10^{-5}$). This effect is not due to a difference in replication between the two viruses (paired Wilcoxon test, $P = 0.07$, calculated for 7 donor samples at each time point on more than 150 pairs), and these results are not affected by the postreplication peak cytotoxic effect of viruses ($P = 4 \times 10^{-4}$ for the reduction of MxA expression, and $P = 0.12$ for the difference in replicativity when removing all data points beyond 12 days for the same tests). Viral replication and MxA induction profiles for individual human and macaque PBMCs are available in Fig. S2 and S3 in the supplemental material). This effect indicates a reduced ability of SIVopt1 to induce IFN-I and suggests that a molecular sensor involved in innate sensing and the IFN-I response is able to detect the difference between SIVwt and SIVopt1, which consists of 70 synonymous mutations within the *pol* region in genomic RNA.

SIVopt1-infected cells induce a reduced IRF3-dependent IFN-β response. To investigate the role of cellular sensors in differential detection of SIVwt and SIVopt1, we used a coculture system consisting of infected Cemx174 lymphocytes and HeLa cell line-derived reporter cells (P4C5 cells expressing SIV receptors

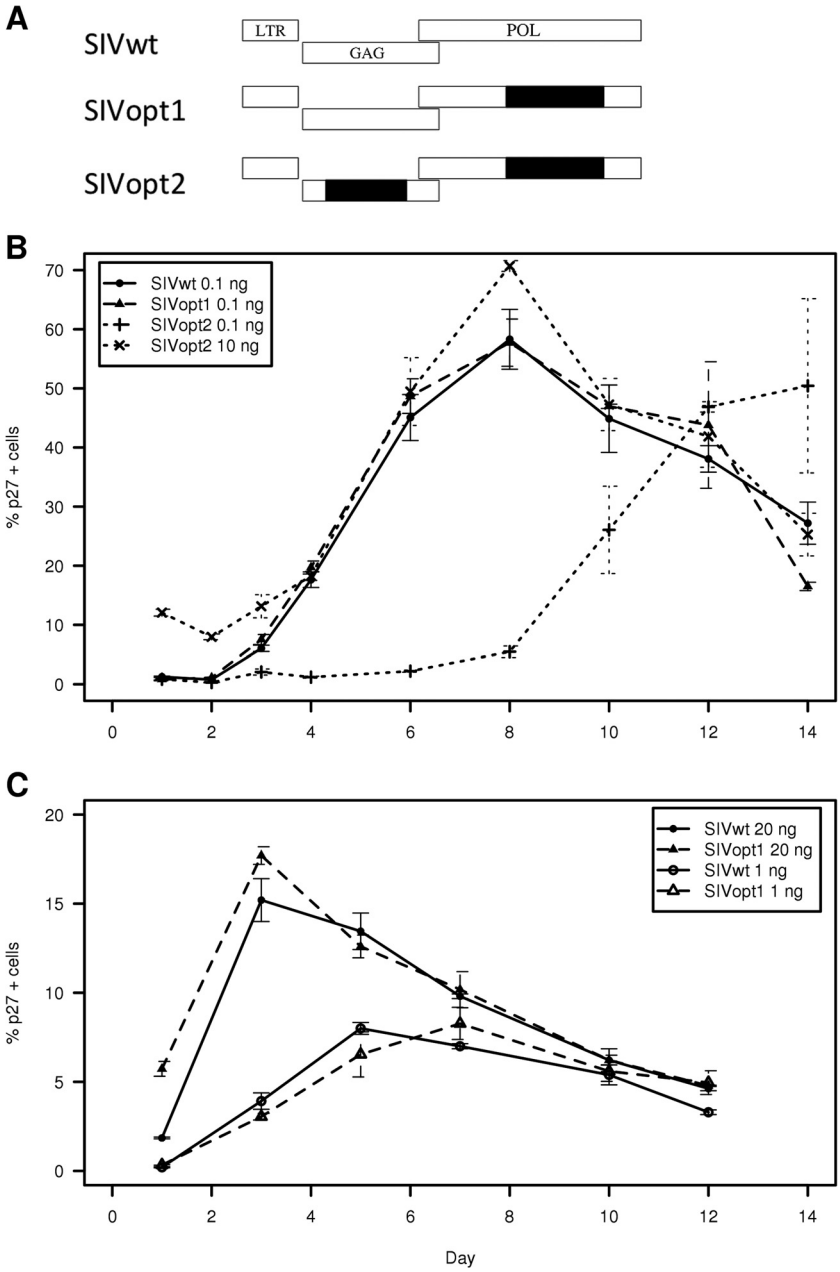

FIG 3 Replicative capacity of SIVopt1 and SIVopt2. (A) Schematic representation of the SIVmac239-based constructs used in this study. The codon-modified regions are indicated in black. (B) Replicative kinetics of SIVopt1 and SIVopt2 compared to SIVwt in Cemx174 cells. Nucleotide modification does not impair the viral replication kinetics of SIVopt1. However, SIVopt2 displays a 2-log decrease in viral infectivity. (C) Replication kinetics of SIVwt and SIVopt1 in macaque PBMCs. The amount of viral input used to infect cells is indicated (in ng of p27).

CD4 and CCR5 [47]) transfected with a luciferase gene under the control of IFN-β promoter (IFN-β-luc). This system allows investigation of the cytoplasmic sensing of viral RNA after cell-to-cell contact, which is a major mode of virus spreading in lentiviral infections (48). SIVwt and SIVopt1 differ only by their genomic RNA. Viral RNAs are detected by endosomal Toll-like receptor 3 (TLR3), TLR7, and TLR8 and cytosolic receptors RIG-I and MDA5 (49). The TLR7 pathway is not functional in HeLa cells, as confirmed by the absence of IFN-β stimulation by CL097, a TLR7 ligand (not shown). Therefore, this experimental system allows addressing whether TLR7-independent pathways (50) can detect SIV-infected lymphocytes and if this detection is sensitive to the nucleotide composition of the viral genome.

We first controlled whether our optimization algorithm decreased the capacity of SIV RNA to induce IFN-I in this cell line, as did commercial human codon optimization of most viral RNAs in HEK293 cells (Fig. 1A). We produced *in vitro*-transcribed RNA from the *Pol* region of SIVwt and SIVopt1 and evaluated their ability to trigger an IFN-I response after transfection in HeLa P4C5 target cells (Fig. 5A, P4C5ctrl). As expected, the optimized version of *pol* RNA induced a reduced IFN-β response compared to wt RNA.

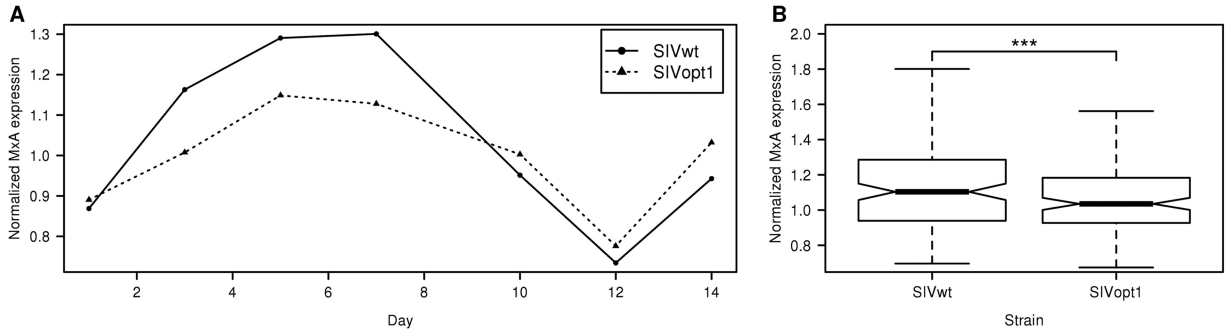

FIG 4 Differential sensing of SIV replication by PBMCs. (A) Example for one experiment. Human PBMCs were infected by SIVwt or SIVopt1. IFN-I induction was measured with the profile of intracellular MxA expression, an interferon-stimulated gene. SIVopt1 shows a reduced ability to induce MxA expression compared to SIVwt. Rebound at day 14 is likely caused by cell mortality. (B) Box plot of SIVwt and SIVopt1 for every measured MxA expression point from 7 different human PBMC donors infected by SIVopt1 and SIVwt. Each PBMC donor sample was divided into several cultures that were infected with different viral input concentrations (always paired between SIVwt and SIVopt1). MxA expression was measured at regular time points following infection (see Fig. S3 in the supplemental material). SIVopt1 globally stimulated less MxA synthesis than SIVwt; ***, paired Wilcoxon test, $P = 3.\ 10^{-5}$ on more than 150 pairs.

We then evaluated the induction of IFN-β-luciferase when Cemx174 cells were cocultured with P4C5 reporter cells. Cemx174 cells infected with SIVwt or SIVopt1 activated the IFN-β promoter. The intensity of luciferase induction was directly related to the percentage of SIV-infected cells in the coculture (Fig. 5B; SIVwt line, $R^2 = 0.87$; SIVopt1 line, $R^2 = 0.86$). Interestingly, SIVopt1 displayed a lower rate of IFN-β stimulation than SIVwt, as illustrated by the difference in the slopes of the regression lines (Fig. 5B, *t* test, $P = 0.001$). Thus, SIV triggers IFN-β when transmitted from cell to cell, and SIVopt1 is attenuated in this capacity.

During infection, binding of certain viral RNAs to cytosolic receptor RIG-I or MDA5 leads to conformational changes that expose their caspase activation and recruitment domain (CARD)-like domains to mitochondrial antiviral signaling proteins (MAVS), inducing downstream signaling for IFN-β transcription through IRF3 (51–58). To characterize the pathway that mediates recognition of SIV-infected cells, we stably expressed in HeLa P4C5 reporter cells BVDV-Pro, a molecule that degrades IRF3 (36). Western blot analysis confirmed that IRF3 levels were reduced in HeLa P4C5 cells expressing BVDV-Pro protease (Fig. 5A). As expected (56), the absence of IRF3 abrogated the activation of IFN-β upon RNA transfection (Fig. 5A, P4C5 BVDV-Pro). IRF3 was also involved in cellular signaling induced by SIV-infected Cemx174 cells (Fig. 5C, reduction of approximately 50% for coculture with SIVwt-infected cells in IRF3-silenced cells). This observation indicates that cytosolic sensors play a role in the detection of SIV-infected cells.

DISCUSSION

The redundancy of the genetic code leads to a specific codon usage bias for each species that refers to differences in occurrence frequency of synonymous codons. The origin of nucleotide and codon bias has been extensively debated and can be explained by mutational and selective mechanisms (41). Consequently, RNA viruses, which have coevolved with their hosts, can still express their proteins in host cells despite a very biased RNA nucleotide composition imposed by their genomic constraints. However, such biased nucleotide composition is likely detected as nonself in infected host cells. Innate antiviral immune response is initiated by the recognition of nonself pathogen-associated molecular patterns (PAMPs) by pattern recognition receptors (PRRs). Viral nucleotide sequences are recognized by TLRs and RIG-I-like cytosolic receptors (RLRs) (59), which upon activation, trigger signaling cascades that converge to induce IFN-α/β secretion and amplification of the antiviral response (59). In this work, we show that nucleotide optimization of viral RNA sequences affects their recognition by PRRs. Indeed, once optimized to human codon usage, viral RNAs virtually lose their capacity to induce IFN-α/β expression in human cells. The A/U richness of viral RNAs correlated with their capacity to stimulate IFN-α/β, whereas optimized G/C-rich sequences were attenuated in this capacity. As a practical consequence of this observation, the use of optimized viral genes in genetic vaccines strategies could be reevaluated by considering the potential immunostimulatory effect of viral sequences.

We further identified local A-rich regions of SIV genes as strong stimulators of IFN-I, as we previously showed for the HIV-1 genome (27). With the aim of designing an attenuated SIV, we generated 70 synonymous mutations that unbalanced the *pol* gene toward the macaque average nucleotide composition. The resulting SIVopt1 was growing as efficiently as SIVwt in a T-cell line and PBMCs, showing that it is possible to proceed to large-scale modification of a lentivirus genome without interfering with its *in vitro* replicative capacity. This also suggests the absence of any unknown regulatory sequences in the optimized *pol* region. Conversely, the delay in replication kinetics observed with SIVopt2, which contains 99 additional mutations in *gag*, and the failure to obtain a replicative virus with 189 mutations in *env* may be due to the existence of regulatory sequences or secondary structures necessary for viral replication. This could be experimentally investigated by designing less optimized versions of these genomic regions. More intriguingly, we were unable to obtain a *gag*-only optimized virus, whereas a virus optimized both in *gag* and *pol* was replicative. We cannot rule out the possibility that a *gag*-optimized replicative virus may be obtained after additional attempts. However, since nucleotide optimization alters the flexibility of RNA molecules (26), it is likely that high local structural constraints induced by optimizing the *gag* region would inhibit certain viral replication functions, such as packaging or reverse transcription. Changing additionally the nucleotide composition of the *pol* region might have induced a compensatory effect in the global flexibility of the genomic RNA molecule, which would re-

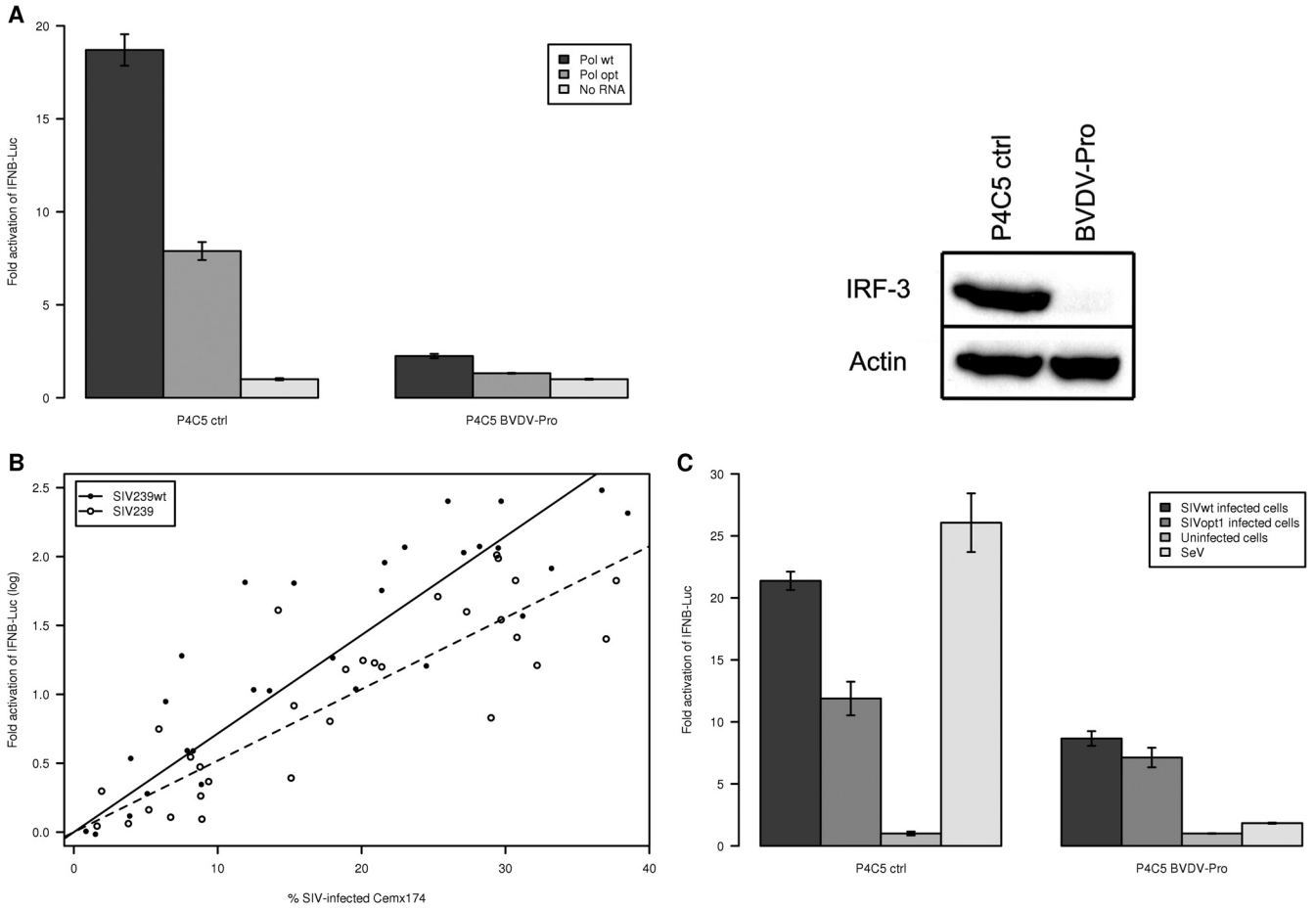

FIG 5 Differential sensing of SIV-infected lymphocytes by HeLa cell line-derived epithelial cells. (A) SIV Pol-derived RNA transfection. HeLa P4C5 cells were transduced with a lentiviral vector (LV) expressing BVDV-Pro or an irrelevant (ctrl) small interfering RNA (siRNA). IFN-I stimulation was measured in cells cotransfected with IFN-β-luciferase reporter plasmid and SIVwt or SIVopt1 Pol RNA sequence. The fold induction of luciferase activity compared to a negative control with no RNA transfection is shown. Wild-type Pol RNA was able to stimulate a higher IFN-I response than the nucleotide-optimized version of the RNA. IFN-I induction was lost in cells expressing BVDV-Pro (which cleaves IRF3). (Right panel) IRF3 and actin levels, assessed by Western blotting, in control (ctrl) and silenced HeLa P4C5 cells. (B) Sensing of SIV-infected lymphocytes. Induction of IFN-I in HeLa P4C5 cells transfected with IFN-β-luciferase reporter plasmid and cocultured for 16 h with CEMx174 cells infected by SIVwt or SIVopt1 at various levels of infection (*x* axis). The fold induction of luciferase activity compared to that in nonstimulated cells is shown (*y* axis). SIVwt-infected cells show a higher IFN-β promoter induction in HeLa P4C5 than in SIVopt1-infected cells ($P = 0.001$). (C) Role of IRF3. Control HeLa P4C5 cells or cells expressing BVDV-Pro were cocultured with approximately 25% SIVwt- or SIVopt1-infected Cemx174 cells and assayed for IFN-β promoter activity. The paramyxovirus Sendai virus (SeV) was used as a positive control.

store viral replicative function, although with a reduced efficacy. In a previous work (26), codon optimization of A-rich sequences within HIV-1 *pol* led to a 100-fold decrease of infectivity, which was explained by enhanced dimer stability of the viral RNA genome and reduction of viral cDNA synthesis. A report published while we were preparing the manuscript (60) describes the use of the synthetic attenuated virus engineering (SAVE) approach (24) to generate optimized and deoptimized HIV-1. Although very few synonymous mutations (<40) were introduced, deoptimized viruses had significantly lower viral replication capacity and reverted to wild-type virulence after serial passages. In contrast, optimized viruses remained stable but were not attenuated (60). Due to the low number of mutations introduced, the global A richness of these genomes was probably minimally affected.

We previously highlighted a correlation between the nucleotide composition of different HIV-1 subtypes and their pathogenicity (27). We proposed that biased RNA would drive a high IFN-I level in persistently infected cells and might be involved in the overactivation of the immune system during chronic pathogenic lentiviral infection. Here, we studied the consequences of optimizing the nucleotide composition of the SIVopt1 genome on IFN-I stimulation. We observed that both transfection of optimized RNA and coculture with SIVopt1-infected cells triggered a lower level of cellular IFN-I response than did SIVwt. In nonimmune cells, this suggests the existence of a cellular sensor able to recognize RNAs according to their nucleotide composition. IRF3 is the transcription factor leading to IFN-I upregulation after triggering of cytosolic viral sensors (61). We demonstrated that IRF3 is involved in IFN-I induction upon detection and differentiation of opt and wt SIV RNAs after transfection or coculture with infected cells. Upstream of IRF3, the cytosolic RLR (RIG-I-like receptor) helicases RIG-I and MDA5 directly recognize multiple and distinct forms of intracellular double-stranded RNA (dsRNA). For instance, putative RIG-I ligands include short 5′-triphosphate RNAs with double-stranded structures (51, 53, 54, 62, 63). The RNAs used in our experiment were *in vitro* tran-

scribed by T7 RNA polymerase and, in consequence, were not capped or polyadenylated. Thus, RIG-I and MDA5 could be involved in their recognition. Addressing precisely their role will require further investigation.

Differences observed between the recognition of SIVopt1 and SIVwt also raise the question of the nature of the viral ligand involved. Our results strongly suggest a role for nucleotide bias recognition in our system. However, other motifs can change with nucleotide modification. HIV-1 genomic RNA is highly structured (19), and this is likely the case for the SIV genome. Changes in RNA nucleotide composition may be detected either directly by RLRs or through the modification of viral RNA structure. Codon optimization of HIV-1 *pol* in a previous study (26) led to destabilization of RNA dimerization and impairment of reverse transcriptase processing. We cannot exclude that nucleotide modifications in SIVopt1 affected recognition by innate sensors without impacting viral replication. This could be the case, by increased production of defective particles or accumulation of nucleic acid replication intermediates.

In vivo studies will be required to analyze the fitness of SIVopt1 and SIVopt2 in macaques and the influence of their nucleotide composition on AIDS induction. Live attenuated SIV vaccines are highly protective in the macaque AIDS model (64), but their transposition to human application is not considered feasible because of safety concerns. Indeed, attenuated SIV persists, and reversion to virulence eventually occurs (65). In this context, original vaccine strategies are desired to reproduce the protection efficacy of attenuated viruses without retaining their associated risks. The mutations artificially introduced in SIVopt1 (70 mutations) and SIVopt2 (169 mutations) did not revert after 10 weeks of culture on Cemx174 cells. Although genetic stability has to be further documented in primary cells and *in vivo*, this first promising observation indicates that our optimization strategy did not alter viral fitness, at least in Cemx174 cells. Using our optimization system, it is thus possible to increase the number of synonymous mutations with the aim to finely tuning viral replication and/or the IFN-I response while preventing reversion of attenuated virus. SIVopt1 has a reduced ability to induce IFN-I but shows wt replicative kinetics *in vitro*. Although a type I IFN response is obviously necessary for a vaccine to shape adaptive immune responses and memory, an excessive response might be deleterious to the reduced pathogenicity and the stability of a live attenuated SIV/HIV vaccine. A considerable amount of work will be required to analyze if these properties are conserved *in vivo* and whether they impact the pathogenicity of infection. In conclusion, our work demonstrates the possibility of rationally changing viral nucleotide composition to design replicative and genetically stable lentiviruses with attenuated pathogenic potentials.

ACKNOWLEDGMENTS

This work was supported by the Centre National de la Recherche Scientifique (CNRS-PEPS), the Agence Nationale de Recherche contre le Sida (ANRS-2010), and the Institut Pasteur. N.V. was a fellow from the Ecole Normale Supérieure (ENS-Lyon, France).

The following reagents were obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH: p239SpE3′ nef Open and p239SpSp5′ from Ronald Desrosiers and SIVmac p27 MAb (55-2F12) from Niels Pedersen.

REFERENCES

1. **Holmes EC.** 2011. What does virus evolution tell us about virus origins? J. Virol. **85:**5247–5251. http://dx.doi.org/10.1128/JVI.02203-10.
2. **Nurmemmedov E, Castelnovo M, Catalano CE, Evilevitch A.** 2007. Biophysics of viral infectivity: matching genome length with capsid size. Q. Rev. Biophys. **40:**327–356. http://dx.doi.org/10.1017/S0033583508004666.
3. **Liu Y, Wimmer E, Paul AV.** 2009. Cis-acting RNA elements in human and animal plus-strand RNA viruses. Biochim. Biophys. Acta **1789:**495–517. http://dx.doi.org/10.1016/j.bbagrm.2009.09.007.
4. **Baum A, Garcia-Sastre A.** 2011. Differential recognition of viral RNA by RIG-I. Virulence **2:**166–169. http://dx.doi.org/10.4161/viru.2.2.15481.
5. **Gale M, Jr, Katze MG.** 1998. Molecular mechanisms of interferon resistance mediated by viral-directed inhibition of PKR, the interferon-induced protein kinase. Pharmacol. Ther. **78:**29–46. http://dx.doi.org/10.1016/S0163-7258(97)00165-4.
6. **Hu C, Saenz DT, Fadel HJ, Walker W, Peretz M, Poeschla EM.** 2010. The HIV-1 central polypurine tract functions as a second line of defense against APOBEC3G/F. J. Virol. **84:**11981–11993. http://dx.doi.org/10.1128/JVI.00723-10.
7. **Mueller S, Coleman JR, Papamichail D, Ward CB, Nimnual A, Futcher B, Skiena S, Wimmer E.** 2010. Live attenuated influenza virus vaccines by computer-aided rational design. Nat. Biotechnol. **28:**723–726. http://dx.doi.org/10.1038/nbt.1636.
8. **Mueller S, Papamichail D, Coleman JR, Skiena S, Wimmer E.** 2006. Reduction of the rate of poliovirus protein synthesis through large-scale codon deoptimization causes attenuation of viral virulence by lowering specific infectivity. J. Virol. **80:**9687–9696. http://dx.doi.org/10.1128/JVI.00738-06.
9. **Grantham R, Gautier C, Gouy M, Mercier R, Pave A.** 1980. Codon catalog usage and the genome hypothesis. Nucleic Acids Res. **8:**r49–r62.
10. **Auewarakul P.** 2005. Composition bias and genome polarity of RNA viruses. Virus Res. **109:**33–37. http://dx.doi.org/10.1016/j.virusres.2004.10.004.
11. **van der Kuyl AC, Berkhout B.** 2012. The biased nucleotide composition of the HIV genome: a constant factor in a highly variable virus. Retrovirology **9:**92. http://dx.doi.org/10.1186/1742-4690-9-92.
12. **van Hemert FJ, van der Kuyl AC, Berkhout B.** 2013. The A-nucleotide preference of HIV-1 in the context of its structured RNA genome. RNA Biol. **10:**211–215. http://dx.doi.org/10.4161/rna.22896.
13. **Berkhout B, Grigoriev A, Bakker M, Lukashov VV.** 2002. Codon and amino acid usage in retroviral genomes is consistent with virus-specific nucleotide pressure. AIDS Res. Hum. Retroviruses **18:**133–141. http://dx.doi.org/10.1089/08892220252779674.
14. **Berkhout B, van Hemert FJ.** 1994. The unusual nucleotide content of the HIV RNA genome results in a biased amino acid composition of HIV proteins. Nucleic Acids Res. **22:**1705–1711. http://dx.doi.org/10.1093/nar/22.9.1705.
15. **Deforche K, Camacho R, Laethem KV, Shapiro B, Moreau Y, Rambaut A, Vandamme AM, Lemey P.** 2007. Estimating the relative contribution of dNTP pool imbalance and APOBEC3G/3F editing to HIV evolution in vivo. J. Comput. Biol. **14:**1105–1114. http://dx.doi.org/10.1089/cmb.2007.0073.
16. **Bishop KN, Holmes RK, Sheehy AM, Malim MH.** 2004. APOBEC-mediated editing of viral RNA. Science **305:**645. http://dx.doi.org/10.1126/science.1100658.
17. **Sheehy AM, Gaddis NC, Choi JD, Malim MH.** 2002. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. Nature **418:**646–650. http://dx.doi.org/10.1038/nature00939.
18. **Lesnik EA, Freier SM.** 1995. Relative thermodynamic stability of DNA, RNA, and DNA:RNA hybrid duplexes: relationship with base composition and structure. Biochemistry **34:**10807–10815. http://dx.doi.org/10.1021/bi00034a013.
19. **Watts JM, Dang KK, Gorelick RJ, Leonard CW, Bess JW, Jr, Swanstrom R, Burch CL, Weeks KM.** 2009. Architecture and secondary structure of an entire HIV-1 RNA genome. Nature **460:**711–716. http://dx.doi.org/10.1038/nature08237.
20. **Malathi K, Dong B, Gale M, Jr, Silverman RH.** 2007. Small self-RNA generated by RNase L amplifies antiviral innate immunity. Nature **448:**816–819. http://dx.doi.org/10.1038/nature06042.
21. **Li M, Kao E, Gao X, Sandig H, Limmer K, Pavon-Eternod M, Jones TE, Landry S, Pan T, Weitzman MD, David M.** 2012. Codon-usage-based

inhibition of HIV protein synthesis by human schlafen 11. Nature **491:** 125–128. http://dx.doi.org/10.1038/nature11433.
22. **Saito T, Owen DM, Jiang F, Marcotrigiano J, Gale M, Jr.** 2008. Innate immunity induced by composition-dependent RIG-I recognition of hepatitis C virus RNA. Nature **454:**523–527. http://dx.doi.org/10.1038/nature07106.
23. **Burns CC, Shaw J, Campagnoli R, Jorba J, Vincent A, Quay J, Kew O.** 2006. Modulation of poliovirus replicative fitness in HeLa cells by deoptimization of synonymous codon usage in the capsid region. J. Virol. **80:** 3259–3272. http://dx.doi.org/10.1128/JVI.80.7.3259-3272.2006.
24. **Coleman JR, Papamichail D, Skiena S, Futcher B, Wimmer E, Mueller S.** 2008. Virus attenuation by genome-scale changes in codon pair bias. Science **320:**1784–1787. http://dx.doi.org/10.1126/science.1155761.
25. **Gutman GA, Hatfield GW.** 1989. Nonrandom utilization of codon pairs in Escherichia coli. Proc. Natl. Acad. Sci. U. S. A. **86:**3699–3703. http://dx.doi.org/10.1073/pnas.86.10.3699.
26. **Keating CP, Hill MK, Hawkes DJ, Smyth RP, Isel C, Le SY, Palmenberg AC, Marshall JA, Marquet R, Nabel GJ, Mak J.** 2009. The A-rich RNA sequences of HIV-1 pol are important for the synthesis of viral cDNA. Nucleic Acids Res. **37:**945–956. http://dx.doi.org/10.1093/nar/gkn1015.
27. **Vabret N, Bailly-Bechet M, Najburg V, Muller-Trutwin M, Verrier B, Tangy F.** 2012. The biased nucleotide composition of HIV-1 triggers type I interferon response and correlates with subtype D increased pathogenicity. PLoS One **7:**e33502. http://dx.doi.org/10.1371/journal.pone.0033502.
28. **Harris LD, Tabb B, Sodora DL, Paiardini M, Klatt NR, Douek DC, Silvestri G, Muller-Trutwin M, Vasile-Pandrea I, Apetrei C, Hirsch V, Lifson J, Brenchley JM, Estes JD.** 2010. Downregulation of robust acute type I interferon responses distinguishes nonpathogenic simian immunodeficiency virus (SIV) infection of natural hosts from pathogenic SIV infection of rhesus macaques. J. Virol. **84:**7886–7891. http://dx.doi.org/10.1128/JVI.02612-09.
29. **Jacquelin B, Mayau V, Targat B, Liovat AS, Kunkel D, Petitjean G, Dillies MA, Roques P, Butor C, Silvestri G, Giavedoni LD, Lebon P, Barre-Sinoussi F, Benecke A, Muller-Trutwin MC.** 2009. Nonpathogenic SIV infection of African green monkeys induces a strong but rapidly controlled type I IFN response. J. Clin. Invest. **119:**3544–3555. http://dx.doi.org/10.1172/JCI40093.
30. **Abdulhaqq SA, Weiner DB.** 2008. DNA vaccines: developing new strategies to enhance immune responses. Immunol. Res. **42:**219–232. http://dx.doi.org/10.1007/s12026-008-8076-3.
31. **Barouch DH.** 2006. Rational design of gene-based vaccines. J. Pathol. **208:**283–289. http://dx.doi.org/10.1002/path.1874.
32. **Li S, Locke E, Bruder J, Clarke D, Doolan DL, Havenga MJ, Hill AV, Liljestrom P, Monath TP, Naim HY, Ockenhouse C, Tang DC, Van Kampen KR, Viret JF, Zavala F, Dubovsky F.** 2007. Viral vectors for malaria vaccine development. Vaccine **25:**2567–2574. http://dx.doi.org/10.1016/j.vaccine.2006.07.035.
33. **Marcheschi RJ, Staple DW, Butcher SE.** 2007. Programmed ribosomal frameshifting in SIV is induced by a highly structured RNA stem-loop. J. Mol. Biol. **373:**652–663. http://dx.doi.org/10.1016/j.jmb.2007.08.033.
34. **Regier DA, Desrosiers RC.** 1990. The complete nucleotide sequence of a pathogenic molecular clone of simian immunodeficiency virus. AIDS Res. Hum. Retroviruses **6:**1221–1231.
35. **Uze G, Di Marco S, Mouchel-Vielh E, Monneron D, Bandu MT, Horisberger MA, Dorques A, Lutfalla G, Mogensen KE.** 1994. Domains of interaction between alpha interferon and its receptor components. J. Mol. Biol. **243:**245–257. http://dx.doi.org/10.1006/jmbi.1994.1651.
36. **Hilton L, Moganeradj K, Zhang G, Chen YH, Randall RE, McCauley JW, Goodbourn S.** 2006. The NPro product of bovine viral diarrhea virus inhibits DNA binding by interferon regulatory factor 3 and targets it for proteasomal degradation. J. Virol. **80:**11723–11732. http://dx.doi.org/10.1128/JVI.01145-06.
37. **Gustafsson C, Govindarajan S, Minshull J.** 2004. Codon bias and heterologous protein expression. Trends Biotechnol. **22:**346–353. http://dx.doi.org/10.1016/j.tibtech.2004.04.006.
38. **Akashi H, Eyre-Walker A.** 1998. Translational selection and molecular evolution. Curr. Opin. Genet. Dev. **8:**688–693. http://dx.doi.org/10.1016/S0959-437X(98)80038-5.
39. **Ikemura T.** 1985. Codon usage and tRNA content in unicellular and multicellular organisms. Mol. Biol. Evol. **2:**13–34.
40. **Comeron JM.** 2004. Selective and mutational patterns associated with gene expression in humans: influences on synonymous composition and intron presence. Genetics **167:**1293–1304. http://dx.doi.org/10.1534/genetics.104.026351.
41. **Plotkin JB, Kudla G.** 2011. Synonymous but not the same: the causes and consequences of codon bias. Nat. Rev. Genet. **12:**32–42. http://dx.doi.org/10.1038/nrg2899.
42. **Plotkin JB, Robins H, Levine AJ.** 2004. Tissue-specific codon usage and the expression of human genes. Proc. Natl. Acad. Sci. U. S. A. **101:**12588–12591. http://dx.doi.org/10.1073/pnas.0404957101.
43. **Urrutia AO, Hurst LD.** 2003. The signature of selection mediated by expression on human genes. Genome Res. **13:**2260–2264. http://dx.doi.org/10.1101/gr.641103.
44. **Caignard G, Komarova AV, Bourai M, Mourez T, Jacob Y, Jones LM, Rozenberg F, Vabret A, Freymuth F, Tangy F, Vidalain PO.** 2009. Differential regulation of type I interferon and epidermal growth factor pathways by a human Respirovirus virulence factor. PLoS Pathog. **5:**e1000587. http://dx.doi.org/10.1371/journal.ppat.1000587.
45. **Silvestri G.** 2008. AIDS pathogenesis: a tale of two monkeys. J. Med. Primatol. **37**(Suppl 2)**:**6–12. http://dx.doi.org/10.1111/j.1600-0684.2008.00328.x.
46. **Haller O, Stertz S, Kochs G.** 2007. The Mx GTPase family of interferon-induced antiviral proteins. Microbes Infect. **9:**1636–1643. http://dx.doi.org/10.1016/j.micinf.2007.09.010.
47. **Amara A, Gall SL, Schwartz O, Salamero J, Montes M, Loetscher P, Baggiolini M, Virelizier JL, Arenzana-Seisdedos F.** 1997. HIV coreceptor downregulation as antiviral principle: SDF-1alpha-dependent internalization of the chemokine receptor CXCR4 contributes to inhibition of HIV replication. J. Exp. Med. **186:**139–146. http://dx.doi.org/10.1084/jem.186.1.139.
48. **Sattentau Q.** 2008. Avoiding the void: cell-to-cell spread of human viruses. Nat. Rev. Microbiol. **6:**815–826. http://dx.doi.org/10.1038/nrmicro1972.
49. **Brennan K, Bowie AG.** 2010. Activation of host pattern recognition receptors by viruses. Curr. Opin. Microbiol. **13:**503–507. http://dx.doi.org/10.1016/j.mib.2010.05.007.
50. **Lepelley A, Louis S, Sourisseau M, Law HK, Pothlichet J, Schilte C, Chaperot L, Plumas J, Randall RE, Si-Tahar M, Mammano F, Albert ML, Schwartz O.** 2011. Innate sensing of HIV-infected cells. PLoS Pathog. **7:**e1001284. http://dx.doi.org/10.1371/journal.ppat.1001284.
51. **Jiang F, Ramanathan A, Miller MT, Tang GQ, Gale M, Jr, Patel SS, Marcotrigiano J.** 2011. Structural basis of RNA recognition and activation by innate immune receptor RIG-I. Nature **479:**423–427. http://dx.doi.org/10.1038/nature10537.
52. **Kawai T, Takahashi K, Sato S, Coban C, Kumar H, Kato H, Ishii KJ, Takeuchi O, Akira S.** 2005. IPS-1, an adaptor triggering RIG-I- and Mda5-mediated type I interferon induction. Nat. Immunol. **6:**981–988. http://dx.doi.org/10.1038/ni1243.
53. **Kowalinski E, Lunardi T, McCarthy AA, Louber J, Brunel J, Grigorov B, Gerlier D, Cusack S.** 2011. Structural basis for the activation of innate immune pattern-recognition receptor RIG-I by viral RNA. Cell **147:**423–435. http://dx.doi.org/10.1016/j.cell.2011.09.039.
54. **Luo D, Ding SC, Vela A, Kohlway A, Lindenbach BD, Pyle AM.** 2011. Structural insights into RNA recognition by RIG-I. Cell **147:**409–422. http://dx.doi.org/10.1016/j.cell.2011.09.023.
55. **Meylan E, Curran J, Hofmann K, Moradpour D, Binder M, Bartenschlager R, Tschopp J.** 2005. Cardif is an adaptor protein in the RIG-I antiviral pathway and is targeted by hepatitis C virus. Nature **437:**1167–1172. http://dx.doi.org/10.1038/nature04193.
56. **Seth RB, Sun L, Ea CK, Chen ZJ.** 2005. Identification and characterization of MAVS, a mitochondrial antiviral signaling protein that activates NF-kappaB and IRF 3. Cell **122:**669–682. http://dx.doi.org/10.1016/j.cell.2005.08.012.
57. **Wu B, Peisley A, Richards C, Yao H, Zeng X, Lin C, Chu F, Walz T, Hur S.** 2013. Structural basis for dsRNA recognition, filament formation, and antiviral signal activation by MDA5. Cell **152:**276–289. http://dx.doi.org/10.1016/j.cell.2012.11.048.
58. **Xu LG, Wang YY, Han KJ, Li LY, Zhai Z, Shu HB.** 2005. VISA is an adapter protein required for virus-triggered IFN-beta signaling. Mol. Cell **19:**727–740. http://dx.doi.org/10.1016/j.molcel.2005.08.014.
59. **Iwasaki A.** 2012. A virological view of innate immune recognition. Annu. Rev. Microbiol. **66:**177–196. http://dx.doi.org/10.1146/annurev-micro-092611-150203.
60. **Martrus G, Nevot M, Andres C, Clotet B, Martinez MA.** 2013. Changes in codon-pair bias of human immunodeficiency virus type 1 have pro-

found effects on virus replication in cell culture. Retrovirology **10:**78. http://dx.doi.org/10.1186/1742-4690-10-78.
61. **Dixit E, Kagan JC.** 2013. Intracellular pathogen detection by RIG-I-like receptors. Adv. Immunol. **117:**99–125. http://dx.doi.org/10.1016/B978-0-12-410524-9.00004-9.
62. **Rehwinkel J, Reis e Sousa C.** 2010. RIGorous detection: exposing virus through RNA sensing. Science **327:**284–286. http://dx.doi.org/10.1126/science.1185068.
63. **Schlee M, Hartmann E, Coch C, Wimmenauer V, Janke M, Barchet W, Hartmann G.** 2009. Approaching the RNA ligand for RIG-I? Immunol. Rev. **227:**66–74. http://dx.doi.org/10.1111/j.1600-065X.2008.00724.x.
64. **Koff WC, Johnson PR, Watkins DI, Burton DR, Lifson JD, Hasenkrug KJ, McDermott AB, Schultz A, Zamb TJ, Boyle R, Desrosiers RC.** 2006. HIV vaccine design: insights from live attenuated SIV vaccines. Nat. Immunol. **7:**19–23. http://dx.doi.org/10.1038/ni1296.
65. **Whitney JB, Wainberg MA.** 2007. Recovery of fitness of a live attenuated simian immunodeficiency virus through compensation in both the coding and non-coding regions of the viral genome. Retrovirology **4:**44. http://dx.doi.org/10.1186/1742-4690-4-44.
66. **Arya SK, Gallo RC.** 1988. Human immunodeficiency virus type 2 long terminal repeat: analysis of regulatory elements. Proc. Natl. Acad. Sci. U. S. A. **85:**9753–9757. http://dx.doi.org/10.1073/pnas.85.24.9753.
67. **Browning MT, Mustafa F, Schmidt RD, Lew KA, Rizvi TA.** 2003. Sequences within the gag gene of feline immunodeficiency virus (FIV) are important for efficient RNA encapsidation. Virus Res. **93:**199–209. http://dx.doi.org/10.1016/S0168-1702(03)00098-4.
68. **Mangeot PE, Negre D, Dubois B, Winter AJ, Leissner P, Mehtali M, Kaiserlian D, Cosset FL, Darlix JL.** 2000. Development of minimal lentivirus vectors derived from simian immunodeficiency virus (SIVmac251) and their use for gene transfer into human dendritic cells. J. Virol. **74:**8307–8315. http://dx.doi.org/10.1128/JVI.74.18.8307-8315.2000.
69. **Le SY, Malim MH, Cullen BR, Maizel JV.** 1990. A highly conserved RNA folding region coincident with the Rev response element of primate immunodeficiency viruses. Nucleic Acids Res. **18:**1613–1623. http://dx.doi.org/10.1093/nar/18.6.1613.
70. **Olsen HS, Beidas S, Dillon P, Rosen CA, Cochrane AW.** 1991. Mutational analysis of the HIV-1 Rev. protein and its target sequence, the Rev responsive element. J. Acquir. Immune Defic.Syndr. **4:**558–567.
71. **Reinhart TA, Rogan MJ, Haase AT.** 1996. RNA splice site utilization by simian immunodeficiency viruses derived from sooty mangabey monkeys. Virology **224:**338–344. http://dx.doi.org/10.1006/viro.1996.0539.
72. **Unger RE, Stout MW, Luciw PA.** 1991. Simian immunodeficiency virus (SIVmac) exhibits complex splicing for tat, rev, and env mRNA. Virology **182:**177–185. http://dx.doi.org/10.1016/0042-6822(91)90661-T.
73. **Victoria JG, Robinson WE, Jr.** 2005. Disruption of the putative splice acceptor site for SIV(mac239)Vif reveals tight control of SIV splicing and impaired replication in Vif non-permissive cells. Virology **338:**281–291. http://dx.doi.org/10.1016/j.virol.2005.05.007.