414–427 Nucleic Acids Research, 2012, Vol. 40, No. 1 Published online 12 September 2011
doi:10.1093/nar/gkr719

Deep sequencing of virus-infected cells reveals HIV-encoded small RNAs

Nick C.T. Schopman[1], Marcel Willemsen[2], Ying Poi Liu[1], Ted Bradley[3], Antoine van Kampen[2,4], Frank Baas[3], Ben Berkhout[1,*] and Joost Haasnoot[1]

[1]Laboratory of Experimental Virology, Department of Medical Microbiology, [2]Bioinformatics Laboratory, Department of Clinical Epidemiology, Biostatistics and Bioinformatics, [3]Department of Genome Analysis, Academic Medical Center, University of Amsterdam, Meibergdreef 15, 1105 AZ and [4]Biosystems Data Analysis, Swammerdam Institute for Life Sciences, University of Amsterdam, Science Park 904, 1098 XH, Amsterdam, The Netherlands

Received June 24, 2011; Revised August 19, 2011; Accepted August 21, 2011

ABSTRACT

Small virus-derived interfering RNAs (viRNAs) play an important role in antiviral defence in plants, insects and nematodes by triggering the RNA interference (RNAi) pathway. The role of RNAi as an antiviral defence mechanism in mammalian cells has been obscure due to the lack of viRNA detection. Although viRNAs from different mammalian viruses have recently been identified, their functions and possible impact on viral replication remain unknown. To identify viRNAs derived from HIV-1, we used the extremely sensitive SOLiD™ 3 Plus System to analyse viRNA accumulation in HIV-1-infected T lymphocytes. We detected numerous small RNAs that correspond to the HIV-1 RNA genome. The majority of these sequences have a positive polarity (98.1%) and could be derived from miRNAs encoded by structured segments of the HIV-1 RNA genome (vmiRNAs). A small portion of the viRNAs is of negative polarity and most of them are encoded within the 3′-UTR, which may represent viral siRNAs (vsiRNAs). The identified vsiRNAs can potently repress HIV-1 production, whereas suppression of the vsiRNAs by antagomirs stimulate virus production. These results suggest that HIV-1 triggers the production of vsiRNAs and vmiRNAs to modulate cellular and/or viral gene expression.

INTRODUCTION

Virus infection of plants, insects, nematodes and fungi results in the production of virus-derived small interfering RNAs (viRNA), which are processed from double-stranded RNA (dsRNA) replication intermediates. These viRNAs trigger the RNA interference (RNAi) pathway and act as guides for the RNA-induced silencing complex (RISC) which catalyses the sequence-specific cleavage of the perfect complementary viral transcripts (1–5). These viruses usually express RNA silencing suppressors (RSSs) to counter this RNA-based antiviral response (6). Although the RNAi machinery is well conserved and functional in mammals, viRNAs could not easily be detected in virus-infected cells of mammalian origin (7). The large DNA viruses, mainly from the family of the herpesviridae, encode multiple viRNAs that are derived from structured single-stranded transcripts, and thus represent virus-encoded miRNAs (vmiRNA), which are believed to regulate the expression of specific viral and/or cellular mRNAs (8–10). However, research failed to detect virus-specific small RNAs in cells infected with mammalian viruses, which is possibly due to extremely low expression levels of the viRNAs.

Deep-sequencing technology can be used as a highly sensitive method to study all kinds of small RNA species in cells (10–13). The possibility to generate tens of millions of sequence reads from a single sample makes these methods effective tools for the discovery of unidentified low-abundant regulatory RNAs (14). Recent deep-sequencing studies showed that viRNAs do indeed accumulate in virus-infected mammalian cells. Parameswaran *et al.* (15) reported small virus-derived RNAs in cells infected with Dengue virus, vesicular stomatitis virus, polio virus, hepatitis C virus and West Nile virus. The 454-sequencing technology also revealed that transposon-specific siRNAs, so-called endo-siRNAs, are produced in mouse oocytes to block transposon activity (16). Because retroviruses such as the human immunodeficiency virus type 1 (HIV-1) are related to

*To whom correspondence should be addressed. Tel: +31 20 566 4822; Fax: +31 20 691 6531; Email: b.berkhout@amc.uva.nl

retrotransposons in terms of integration of the proviral DNA genome in the host cell DNA, it is likely that HIV-1-specific viRNAs can be produced in infected cells to regulate and possibly down-modulate virus replication. Indeed, Yeung *et al.* (17) reported the accumulation of small RNAs in HIV-1-infected cells. In that study, 454 sequencing of HIV-1-derived small RNAs yielded only about 100 reads that map to the HIV-1 genome. Another study used the SOLiD deep-sequencing technology, which was designed to be more sensitive than 454 sequencing (14). Cells were infected with a VSV-G pseudotyped HIV-1 vector and the authors yielded 0.7% sequences of viral origin, but half of these reads were vector-derived GFP sequences (18). In this study, we also used SOLiD sequencing to identify HIV-1-derived small RNAs.

We performed small RNA cloning followed by ultra deep sequencing using SOLiD™ technology to analyze small RNA species of ~15–33 nt in HIV-1-infected cells. We identified 26 000 HIV-1-derived small RNAs, which is 1% of the total small RNA population. The majority of small RNAs are derived from cellular miRNAs, tRNAs and ribosomal RNAs. A small portion of sequences correspond to endogenous transposable elements. Although small RNAs can be found throughout the plus-strand HIV-1 RNA genome, several hot spots were identified. In addition, a restricted set of specific minus-strand viRNA were found. Collectively, these viRNAs represent two types of virus-derived small RNAs that originate from two distinct pathways (Figure 1). The vmiRNAs originate from hairpin structures in the plus-strand viral transcript that are processed by the RNAi machinery (Drosha, Dicer). The vsiRNAs are likely produced by Dicer from dsRNA formed between HIV-1 RNA and antisense transcripts, which are made from an antisense promoter located either within or downstream of the integrated provirus. Further analyses of the identified HIV-1-derived vmiRNAs revealed candidates that are functional and we describe vsiRNAs that can inhibit virus production.

MATERIAL AND METHODS

Virus infection

SupT1 suspension T cells were grown in Advanded Rosewell Park Memorial Institute medium (Invitrogen, Carlsbad, CA, USA) supplemented with L-glutamine, 1% fetal calf serum (FCS), penicillin (30 U/ml) and streptomycin (30 µg/ml), in a humidified chamber at 37°C and 5% CO_2. The full-length HIV-1 molecular clone pLAI was described (19) and a doxycycline (dox)-inducible HIV-1 variant, HIV-rtTA, was constructed previously (20). In HIV-rtTA, the Tat/TAR axis required for transactivation of transcription is replaced by the tet-operator/rtTA system for dox-inducible promoter activity and virus replication. HIV-1 and HIV-rtTA virus were produced by transfection of the respective molecular clones in 293T cells as described previously (21). Virus was harvested from the transfected cultures 2 days post infection. Virus production was measured by CA-p24 enzyme-linked immunosorbent assay (ELISA). SupT1 cells (20×10^6 cells) were infected with HIV-1 or the inducible HIV-rtTA variant for subsequent SOLiD sequencing. The latter culture received dox (1000 ng/ml) to drive viral gene expression. SupT1 cells infected with HIV-1 reached a 60% infection grade as determined by intracellular CA-p24 staining by FACS and HIV-rtTA reached a 2% infection level. For FACS analysis, 2×10^6 cells are pelleted (4 min, 4000 rpm) and fixated for 5 min at room temperature (RT) with 200 µl 37% formaldehyde (stabilized with 10% methanol) diluted in phosphate buffered saline. Next, cells are taken up in FACS buffer (2% BSA/FCS in phosphate buffered saline) and again pelleted. Supernatant was removed and cells were stored at 4°C. The cells were washed with BD Perm/Wash™ buffer (BD Pharmingen) and stained for minimally 30 min at 4°C with RD1-conjugated mouse monoclonal anti-CA-p24 antibody diluted (1:100) in BD Perm/Wash™ buffer. Excess antibody was removed by washing the cells with BD Perm/Wash™ buffer and the cells were resuspended in FACS buffer. Cells were analyzed on a BD FACSCanto II flow cytometer with

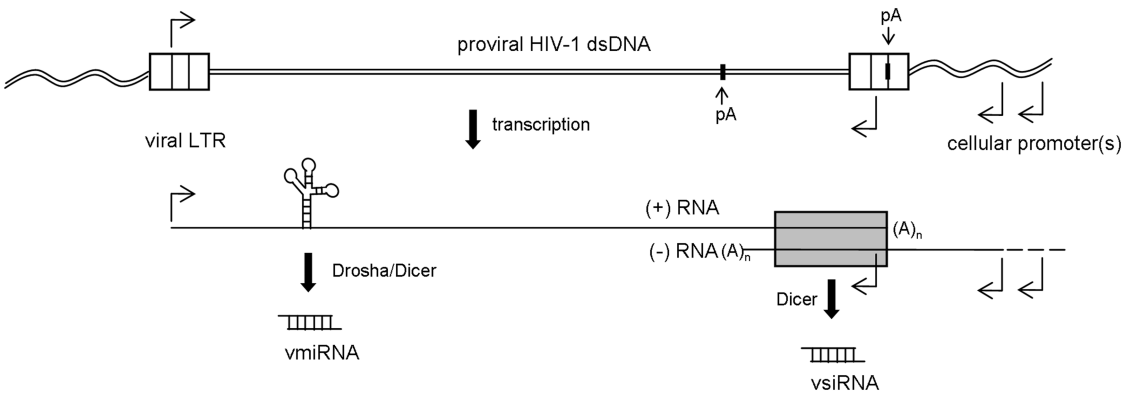

Figure 1. Origin of HIV-1 viRNAs. Two main classes of viRNAs can be produced by HIV-1 that we termed vmiRNAs and vsiRNAs. The vmiRNAs originate from structured regions in the viral RNA genome that trigger Drosha/Dicer processing. The vsiRNAs are produced from extended dsRNA formed by HIV-1 RNA and antisense RNA made by convergent transcription of the HIV-1 genome. These antisense transcripts are made from a promoter in the viral genome or the flanking cellular genes. Promoter elements are indicated by arrows. Putative polyadenylation signals (pA) are indicated (54). The dsRNA is marked by a gray box.

BD FACSDiva Software v6.1.2 (BD biosciences, San Jose, CA, USA). Uninfected SupT1 cells were used as negative control. Uninfected SupT1 cells treated with dox served as an additional control. To ensure single round infection, the fusion inhibitor T1249 was added (400 nM) to the media 24 h post infection. Total RNA was isolated 48 h post infection.

Small RNA library preparation and emulsion PCR for SOLiD sequencing

Total RNA was isolated from (HIV-1 or HIV-rtTA)-infected SupT1 cells and uninfected control cells using the mirVana miRNA isolation kit (Ambion). To purify the small RNA population, a denaturing PAGE gel (6%) was used for size fractionation. The ~15–40 nt RNA fragments were eluted and subsequently used to create a small RNA library that is compatible with the SOLiD-sequencing platform. We used the SOLiD™ Small RNA Expression Kit in which small RNAs are hybridized and ligated to adaptors. Next, the small RNA population with ligated adaptors is reverse transcribed into cDNA with a subsequent RNase H treatment to remove the RNA template. The cDNA library is amplified in 13–15 cycles of Polymerase Chain Reaction (PCR) with the primer set supplied in the kit. To concentrate samples and remove PCR by-products from the amplified cDNA library, PCR products of ~105–150 bp that correspond to the small RNA population were isolated from PAGE gel.

Samples are now ready for emulsion PCR (ePCR), which was performed according to the manufacturer's instructions (Applied biosystems, SOLiD System 3.0 user guide). The sequence templates were attached to beads and clonal bead populations were prepared in microreactors containing template, PCR reaction components, beads and primers. After ePCR, the templates are denatured and bead enrichment is performed for the templated beads, which was required to isolate templated beads from non-amplifying or poorly amplifying beads. Finally, the template on the selected beads was 3′ modified by Terminal Transferase to allow covalent bonding to a glass slide and subsequently sequenced in the SOLiD analyzer (Applied Biosystems).

Bioinformatics

Analysis of the SOLiD colorspace reads was performed with rna2map v0.5.0, also known as the 'small RNA pipeline' (SOLiD software development community). First, small RNA sequences were selected that mapped on the HIV-1 genome (HIV-1 wild-type isolate LAI or the HIV-rtTA construct was used as template). All HIV-1 genome mapping was done with a 15 nt seeding step length, allowing 0 mismatches, whereas after the seeding step mismatches and bulges were allowed up to 35 nt in length to determine the read length. The classification of the remaining small RNA sequences was based on sequence analysis using the provided filter that removes known rRNA, tRNA, snRNA, scRNA, repeats and human endogenous retrovirus elements (HERVs). In this step, the default filter provided by SOLiD was used, allowing up to two mismatches on alignments of 20 nt length. Next, mirBase was used to filter the known cellular miRNAs (MirBase version 15). The remaining small RNA sequences were matched against the human genome (UCSC hg18). The alignment against the human genome was performed with the same settings as used in the alignment against the HIV-1 genome. Matching files, generated by the small RNA pipeline, were converted to '.gff' files using matoGff 0.2.06 (SOLiD software development community), which were further examined with SOLiD Alignment Browser 2.1. Several R scripts (ShortReads package) were used to plot and connect all components.

Plasmid and siRNAs construction

To generate vmiRNA expression constructs, the candidate vmiRNA hairpin domains with ~100-bp flanking sequences were amplified from LAI DNA. The primers encoded the BamHI and XhoI restriction sites for subsequent cloning purposes. PCR products were excised from gel, purified with the QIAquick gel extraction kit (Qiagen), BamHI/XhoI digested and cloned in the pcDNA6.2-GW/EmGFP-miR vector (Invitrogen). These plasmids were transformed into bacterial TOP10 cells. Positive clones were selected by colony PCR and sequenced using the BigDye Terminator v1.1 Cycle Sequencing Kit (Perkin Elmer Applied Biosystem).

Synthetic vsiRNAs were designed from the HIV-1 antisense reads obtained by SOLiD sequencing. The most abundant HIV-1 derived antisense reads were selected and a total of eight vsiRNAs were constructed with the Silencer siRNA Construction Kit (Ambion) according to the manufacturer's instructions. An siRNA targeting the HIV-1 Nef gene (siNef) was used as positive control, and an irrelevant siRNA targeting firefly luciferase (siLuc) was used as a negative control. To inhibit the vsiRNAs, we designed locked nucleic acids (LNA) molecules that are perfectly complementary to the vsiRNA (Eurogentec).

Transfection experiments

Human embryonic kidney 293T adherent cells were grown as monolayer in Dulbecco's modified Eagle's medium (Invitrogen, Carlsbad, CA, USA) supplemented with 10% FCS, penicillin (100 U/ml) and streptomycin (100 µg/ml) in a humidified chamber at 37°C and 5% CO_2. Co-transfections of pLAI and the siRNA and/or LNA were performed in the 96-well format. Per well, 2×10^4 293T cells were seeded in 100 µl DMEM with 10% FCS without antibiotics. The next day, 100 ng pLAI (or 25 ng pGL-3), 50 nM siRNA with or without 50 nM LNA, and 0.5 ng pRL (Renilla luciferase) were transfected with 0.5 µl Lipofectamine 2000 in 50 µl according to the manufacturer's instructions (Invitrogen).

Following transfection of vmiRNA expression constructs, 100 ng of pLAI (or 25 ng of pGL-3) together with 0, 25 or 100 ng of pcDNA6.2-GW/EmGFP-miR and 0.5 ng of pRL (Renilla luciferase) were transfected with 0.5 µl Lipofectamine 2000 in 50 µl according to the manufacturer's instructions (Invitrogen). Two days after transfection, the supernatant was harvested, virus was inactivated and CA-p24 ELISA was performed to

quantify virus production. The cells were lysed for Renilla luciferase activity measurements with the Renilla Luciferase Assay System (Promega). To correct for transfection variation, the CA-p24 values were divided by the Renilla values. We set the condition that for a valid experiment the ratio between the highest and the lowest Renilla values should differ by less than a factor of 2. Transfection experiments were corrected for between session variations as described previously (22).

Detection of vsiRNA-mediated cleavage by 5′-RACE PCR

Human embryonic kidney 293T cells (2×10^6 cells) were co-transfected with 10 ug of pLAI together with 50 nM vsiRNAs using Lipofectamine 2000 reagent. Total RNA was extracted 2 days post transfection with the mirVana miRNA isolation kit (Ambion) according to the manufacturer's protocol. The RNA concentration was measured using the Nanodrop 1000 (Thermo Fisher Scientific). Genomic DNA was removed by DNase treatment using the TURBO DNA-free kit (Ambion). Of the total RNA, 5 μg was used as template in the 5′-Rapid Amplification of cDNA Ends (RACE) system (Invitrogen) and the reaction was carried out according to manufacturer's protocol.

siRNA detection by northern blotting

Northern blot experiments were performed as previously described (23). Briefly, human embryonic kidney 293T cells (5×10^6 cells) were transfected with 40 μg of shRNA constructs or pLAI using Lipofectamine 2000 reagent. Small RNA was extracted 2 days post transfection with the mirVana miRNA isolation kit (Ambion, Austin, TX, USA) according to the manufacturer's protocol. The RNA concentration was measured using the Nanodrop 1000 (Thermo Fisher Scientific). For northern blot analysis, 2.4 μg small RNA was electrophoresed in an urea denaturing 15% polyacrylamide gel (precast Novex TBU gel, Invitrogen, Carlsbad, CA, USA). RNA molecular weight markers (ABI) were prepared as suggested by the manufacturer's protocol and run alongside. After electrophoresis, the gels were stained with 2 μg/ml ethidium bromide for 20 min and destained with milliQ water for 10 min. The tRNA signal was visualised under UV light to check for equal sample loading. The RNA in the gel was electrotransferred to a positively charged nylon membrane (Boehringer Mannheim, GmbH, Mannheim, Germany). The RNA was crosslinked to the membrane using UV light at a wavelength of 254 nm (1200 μJ × 100). Hybridizations were performed at 42°C with radiolabeled locked nucleic acid (LNA) oligonucleotides in 10 ml ULTRAhyb hybridization buffer (Ambion, Austin, TX, USA) according to the manufacturer's instructions. LNA oligonucleotide probes were 5′-end labeled with the kinaseMax kit (Ambion) in the presence of 1 μl [γ-^{32}P] ATP (0.37 MBq/μl Perkin Elmer). To remove unincorporated nucleotides, the probes were purified on Sephadex G-25 spin columns (Amersham Biosciences) according to the manufacturer's protocol. We used the following vmiRNA-specific probes (LNA positions underlined): 5′-AAGCAGTGGGTTCCCTAGTTAG-3′ (vmiRNA-43/9175) and 5′-TGTTTCCCATGTTTCCTTT-3′ (vmiRNA-3270). The signal was detected by autoradiography and quantified using a phosphorimager (Amersham Biosciences).

RESULTS

Sequencing of small RNAs from HIV-1-infected cells

To study small RNAs in virus-infected mammalian cells, we infected SupT1 T cells with the primary CXCR4-using HIV-1 LAI isolate. We performed an acute infection with relatively high virus input and harvested the cells (60% infected) at 48 h post infection. Uninfected SupT1 cells served as a negative control. To study the previously reported vmiRNA encoded by the viral TAR hairpin motif, SupT1 cells were also infected with the HIV-rtTA variant. HIV-rtTA is modified such that Tat-mediated transcription transactivation is inactivated by mutations in the TAR hairpin that affect binding of the Tat protein and the tetO and rtTA components of the Tet-On system were inserted (24,25). This virus is fully dependent on doxycycline (dox) for gene expression and replication and we analyzed parallel infected cultures with and without dox. Two days post infection total RNA was isolated from these four cultures, size fractionated (15–40 nt), ligated to adaptors, amplified by RT–PCR and sequenced on the SOLiD sequencer. From a single run, a total of 16 127 962 raw sequences were analyzed, resulting in 5 210 754 identified small RNAs. Sequences were categorized as cellular miRNAs, ribosomal RNA, small nuclear RNA (snRNA), small cytoplasmic RNA (scRNA), tRNA, LINEs and SINEs (repeats), retrotransposon RNA, other RNAs originating from the human genome and HIV-1-specific RNAs. The results obtained for the HIV-rtTA control infection without dox largely overlap with the uninfected SupT1 control cells (results not shown). The distribution among the small RNA classes for uninfected, HIV-1 and HIV-rtTA-infected cells is summarized in Figure 2A.

Infected versus non-infected cells: the tRNA^{Lys} case

Virus-specific small RNAs were detected in cells infected with HIV-1 and HIV-rtTA. In the HIV-1-infected cells, 25 981 HIV-1-specific reads (1% of total sequences) were scored. The HIV-rtTA-infected cells yielded 2474 (0.1%) virus-specific reads. HIV-derived small RNA sequences were up to 33 nt long, but most reads were 15–22 nt in length with a peak at around 18 nt (Figure 2B, middle panel). HIV-rtTA yielded a similar pattern (Figure 2B, lower panel). The analyses of the small RNAs from two different cultures infected with HIV-1 yielded a similar small RNA class and size distribution, indicating that the sequencing method is reproducible (results not shown). Surprisingly, a HIV-specific signal with a sharp peak at 18 nt was also detected in the uninfected control cells (Figure 2B, upper panel). In fact, this 'HIV-1' signal corresponds to a single antisense RNA read that is fully complementary to the tRNA primer binding site (PBS) of the HIV-1 RNA genome (Figure 3, lower panel). This cellular signal was also present in the virus-infected

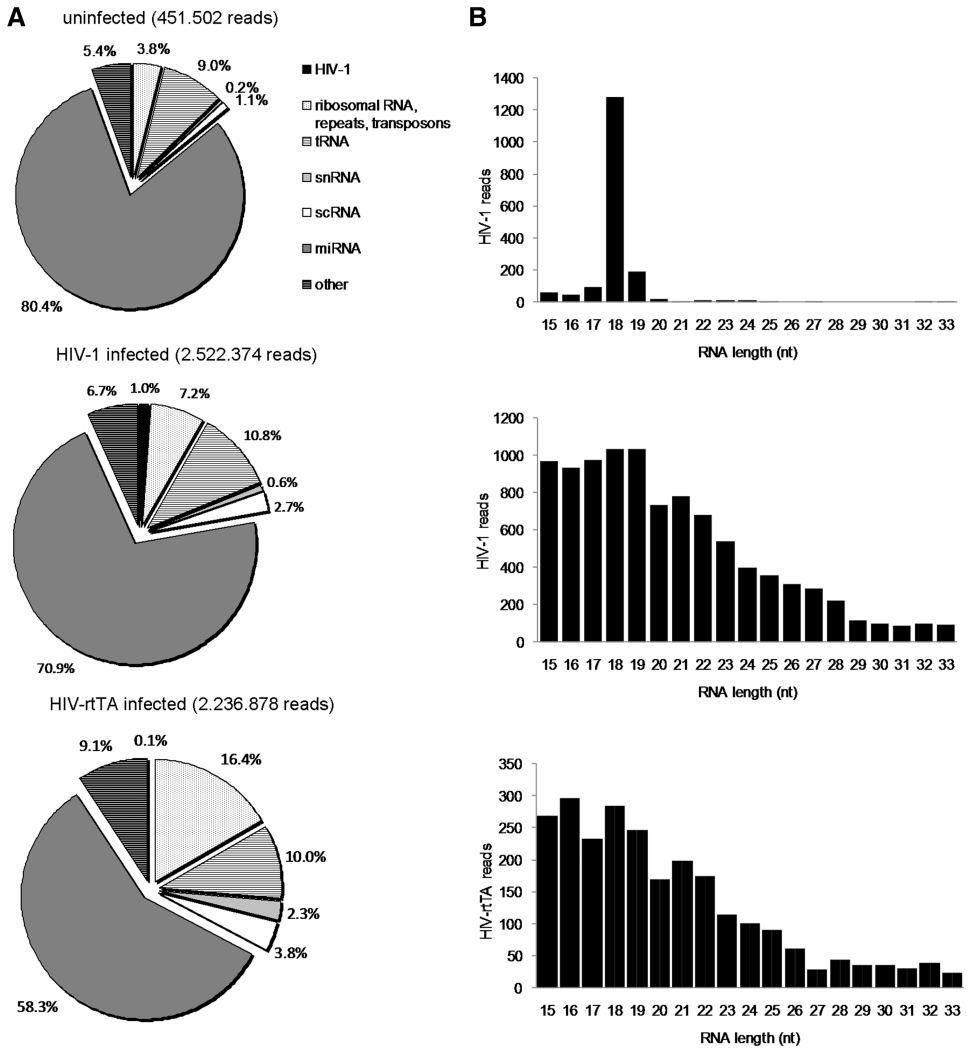

Figure 2. Distribution of the small RNAs. (**A**) Small RNA reads obtained after deep-sequencing of uninfected SupT1 cells, SupT1 cells infected with HIV-1 or HIV-rtTA were annotated. The small RNA reads could be assigned to various RNA classes. The fraction of reads is depicted for each category. The total number of reads per experimental group is indicated above each chart. (**B**) Size distribution of HIV-1-derived small RNAs. Small RNA reads matching HIV-1 or HIV-rtTA sequences were detected by deep-sequencing. The small RNAs vary in size from 15 to 33 nt. The uninfected control was matched against HIV-1 RNA.

samples. Although this small RNA was previously reported by Yeung *et al*. (17) to represent a virus-derived viRNA, it seems more likely that it represents a degradation product of the cellular $tRNA^{Lys3}$ molecule, which acts as primer for HIV-1 reverse transcription by annealing to the complementary PBS (Figure 4A). Several studies have reported processing of tRNA molecules by the RNAi machinery (26,27). Further proof for this scenario comes from the presence of a minor $tRNA^{Lys5a}$ fragment that is known to act as minor primer species in HIV-1 reverse transcription (Figure 4A) (28,29). The 3′ terminal 18–21 nt of $tRNA^{Lys3}$ have a perfect match or near perfect match ($tRNA^{Lys5a}$) with the PBS of HIV-1 RNA (Figure 4B). Thus, the $tRNA^{Lys3/5a}$ primers explain the presence of an HIV-like antisense sequence in uninfected cells.

Interestingly, the analysis of the tRNA-derived small RNA fragments suggests that the tRNA primer concentration is strongly affected by HIV-1 infection (Figure 4C). The relative number of $tRNA^{Lys}$ reads is dramatically reduced during HIV-1 infection, $tRNA^{Lys3}$ drops to 1.4% and $tRNA^{Lys5a}$ to 4.9% of the respective pre-infection levels (arbitrarily set at 100%). The level of all other tRNAs was not as significantly affected, which was plotted for all tRNAs and $tRNA^{Ala}$ as representative. A possible explanation for the severe and specific drop in $tRNA^{Lys}$ levels is that these tRNAs are packaged in newly produced virus particles, although we cannot exclude that HIV-1 infection selectively blocks the processing of the tRNA primer species by the RNAi machinery.

HIV-1 plus-strand derived viRNAs

Further analyses of the HIV-derived sequences from infected cells revealed viRNAs originating primarily from across the positive-stranded RNA genome (Figure 3). Only a small portion (1.9%) of the sequences

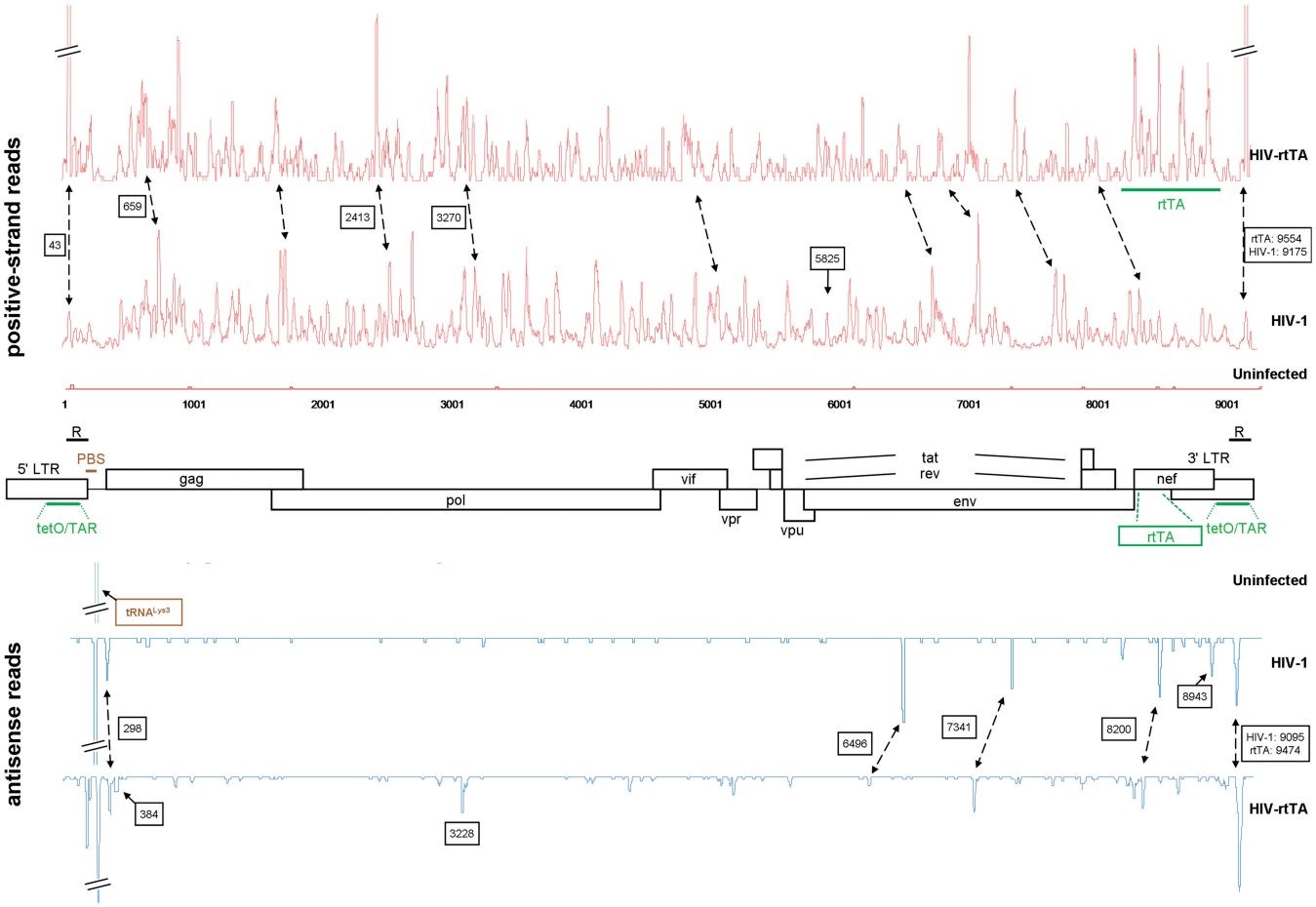

Figure 3. Small RNAs depicted on the HIV-1 RNA genome. The HIV-1 genome is schematically shown in the middle panel. Indicated in green are the insertions and mutations in HIV-rtTA. The plus-strand small RNAs are depicted in red on top of the HIV-1 genome, and the antisense small RNA coverage is plotted below in blue. Peaks represent abundant small RNAs. Dotted lines indicate hotspots shared between HIV-1 and HIV-rtTA. Note that the HIV-rtTA genome is ~400 nt larger than the HIV-1 genome, explaining the shift seen between identical hot spots. Reads and candidate viral viRNAs that were investigated in further detail are marked with the boxed genome position. The primer binding site (PBS) on HIV-1 RNA is indicated in brown and is fully complementary to the major antisense $tRNA^{Lys}$-derived read in uninfected and infected cells.

corresponds to the antisense orientation of the HIV-1 RNA, of which the vast majority represents the cellular tRNA signal. The other antisense reads cluster mostly in the 3′-LTR region of the HIV genome.

We plotted the small RNAs derived from the plus-strand RNA genome of HIV-1 or HIV-rtTA against the viral genome map (Figure 3, upper panel). Close inspection reveals several hot spots of small RNA production. The most intense peaks marked by a boxed genome position in Figure 3 were selected for further investigation. The same overall pattern was observed for HIV-1 and HIV-rtTA, and we used arrows to align some identical signals as the HIV-rtTA genome contains several sequence inserts. There are also some striking differences in the small RNA pattern between HIV-1 and HIV-rtTA, which could be due to the rtTA gene and tetO insertions or the TAR mutations in the latter virus. These genome changes may alter the overall genomic RNA structure, thus leading to different small RNA products. Please note that the R region, which includes the TAR element, is identical at both 5′ and 3′-ends of the HIV-1 RNA genome. This means that we were unable to determine the precise origin of the R-derived small RNA reads and therefore show them at both ends (Figure 3).

The specific small RNA products can be produced from the unspliced or spliced HIV-1 RNAs via Dicer processing or yet unknown mechanisms. In search for candidate HIV-1-encoded vmiRNAs, we focused on hotspot RNA fragments ranging in size from 18 to 21 nt. This selection will remove random RNA degradation products. For the selected hits, the local RNA secondary structure was analyzed for the presence of stem–loop structures with pre-miRNA characteristics. The local RNA structure was determined by Mfold analyses of the small RNA with ~100 nt flanking sequences on each side. This initial screen revealed 16 candidate pre-vmiRNAs. The MiPred miRNA prediction software was subsequently used to obtain a score for each candidate vmiRNA. MiPred is a computational method for prediction of new miRNA candidates, which is presented as a percentage of likelihood (30). Only two candidates passed this stringent

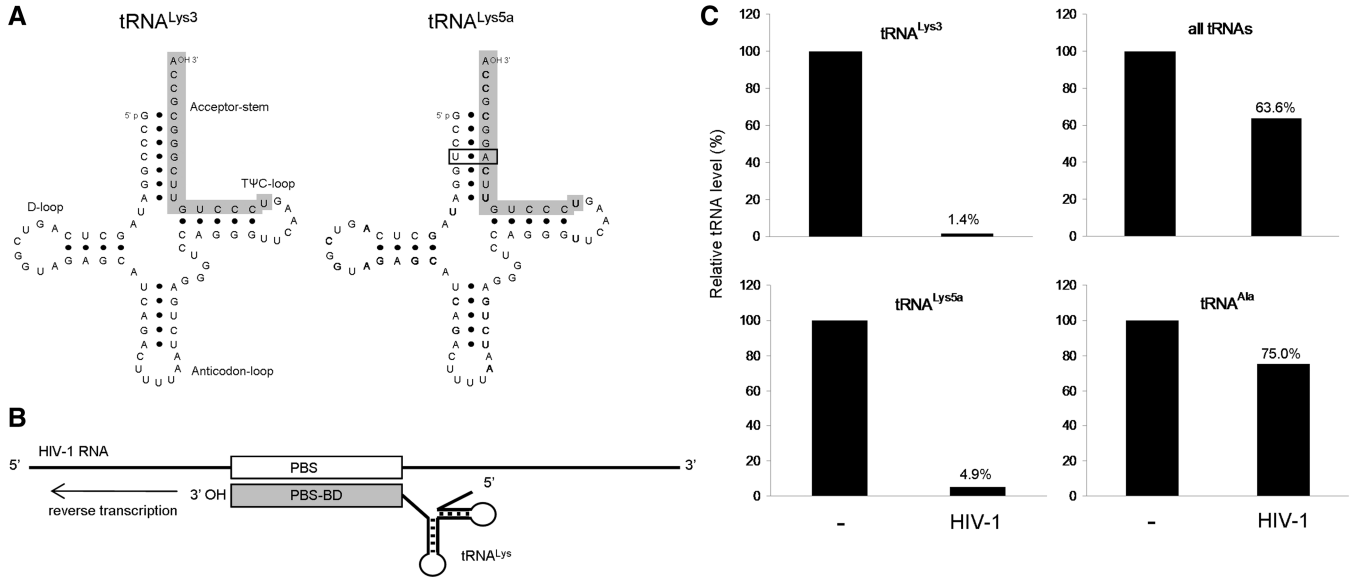

Figure 4. $tRNA^{Lys}$ variants used in HIV-1 reverse transcription. (**A**) The cloverleaf secondary structure of $tRNA^{Lys3}$ and $tRNA^{Lys5a}$. The 3′-end sequence that was predominantly sequenced is marked in gray. This region of the tRNA binds to the PBS of HIV-1 RNA. The boxed nucleotides mark the single base pair difference between the two tRNAs. (**B**) Both $tRNA^{Lys}$ molecules can be used as primer for reverse transcription of the HIV-1 RNA genome. (**C**) tRNA levels found in HIV-1-infected SupT1 cells and uninfected cells. For $tRNA^{Lys3}$ and $tRNA^{Lys5a}$, the reads represent the most abundant short 3′-terminal fragment. For all other tRNAs, we tabulated all tRNA-derived fragments of 16–35 nt. The number of tRNA reads was corrected for the number of ribosomal RNAs reads and the values calculated for uninfected cells were set at 100%.

Table 1. HIV-encoded vmiRNA candidates

vmiRNA candidates	HIV-1 region	No. of reads	Size sequence (nt)	ΔG pre-vmiRNA (kcal/mol)	miRNA score[a] (%)	Most abundant sequence	Sequence conservation[b] (%)	References
43/9175	TAR	53	19–22	−36.80	Real (78)	43–62 (16×)	95	(32) (31)
659	gag	175	18–21	−18.82	Pseudo (82)	666–685 (55×)	75	(31)
2413	pol	85	17–23	−23.80	Real (60)	2413–2432 (39×)	99	New
3270	pol	80	17–21	−18.20	Pseudo (75)	3274–3292 (38×)	82	New
5825	env	27	15–21	−17.80	Pseudo (66)	5831–5845 (7×)	92	New

[a]Calculated by MiPred (30).
[b]Percentage of strains with ≤2 nt substitutions compared to our prototype LAI.

test, vmiRNA-43/9175 in the TAR region and vmiRNA-2413 in the pol gene (Table 1). Hairpin motifs that do not obtain the threshold score of a real miRNA are classified as pseudo-miRNAs: vmiRNA-659, 3270 and 5825. These five candidate vmiRNAs and their properties are listed in Table 1 and the predicted pre-vmiRNA hairpin structures are shown in Figure 5. We noticed a high level of sequence conservation among HIV-1 strains for the very likely candidates, vmiRNA-43/9175 and vmiRNA-2413 (95% and 99%, respectively). Sequence conservation was determined with the QuickAlign tool of the Los Alamos HIV sequence database as the percentage of 371 HIV-1 subtype B isolates with two or less mismatches compared to the prototype LAI isolate. Interestingly, vmiRNA-43/9175 and vmiRNA-659 that we identified via the hotspot small RNAs were previously predicted by computer analyses (31), and the former has been verified experimentally (18,24).

Previous studies on vmiRNAs in HIV-1-infected cells revealed that the TAR hairpin present at the extreme 5′ and 3′-ends of all HIV-1 transcripts is a substrate for Dicer and processed into small fragments of miRNA size (17,33,34). Primarily small RNAs originating from the 3′ side of the hairpin were described, which were predicted to target apoptosis related genes, suggesting that TAR-derived vmiRNAs may prevent apoptosis of the infected cell in order to boost virus production (32). We detected the abundant TAR-derived vmiRNA-43/9175 that is marked in red in the TAR stem–loop models in Figure 6. Most other RNA fragments were 20–22 nt and originated from the 3′ side of the TAR hairpin (Figure 6 and Table 2), in agreement with the previous reports. The relative accumulation of TAR-derived small RNAs was significantly higher in cells infected with HIV-rtTA than wild-type HIV-1 (4.3% versus 0.4% of the total HIV-1-specific reads). This effect may be due to the multiple inactivating point mutations present in the TAR motif of HIV-rtTA (marked in black boxes in Figure 6). These sequence changes or their structural impact also likely explains the observed changes in actual TAR-processing products (Table 2).

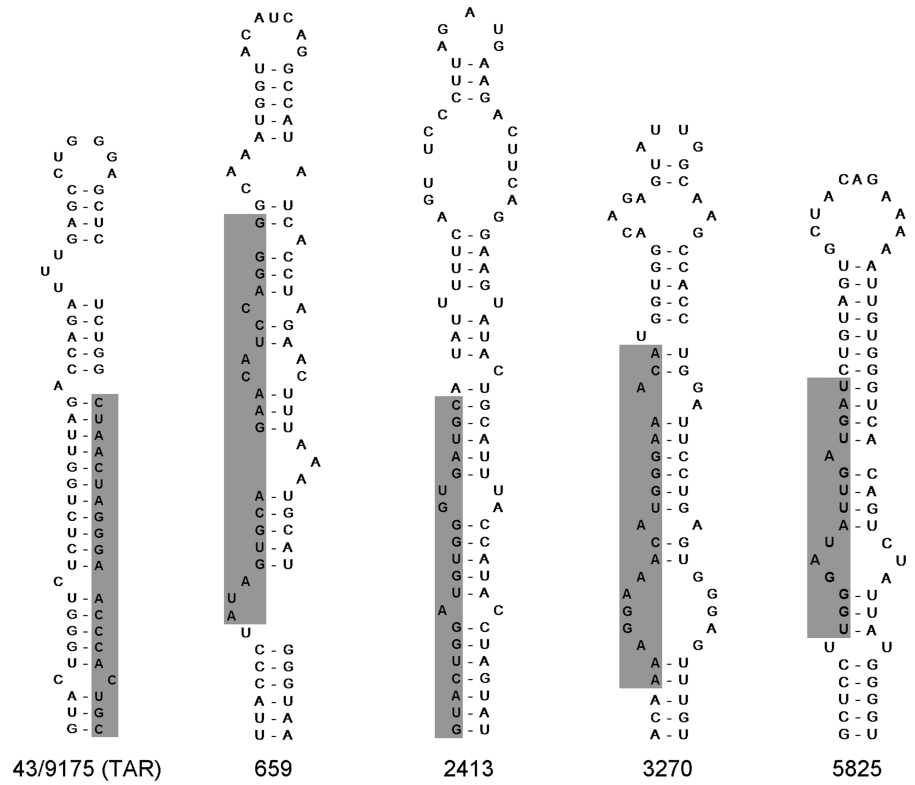

Figure 5. Secondary structure of candidate HIV-1-encoded pre-vmiRNAs. The number of the pre-miRNA indicates the position on the HIV-1 RNA genome. The most abundant read scored by deep-sequencing is marked in gray. The RNA structures were predicted by the Mfold algorithm. The 43/9175 pre-vmiRNA represents the 3′TAR motif.

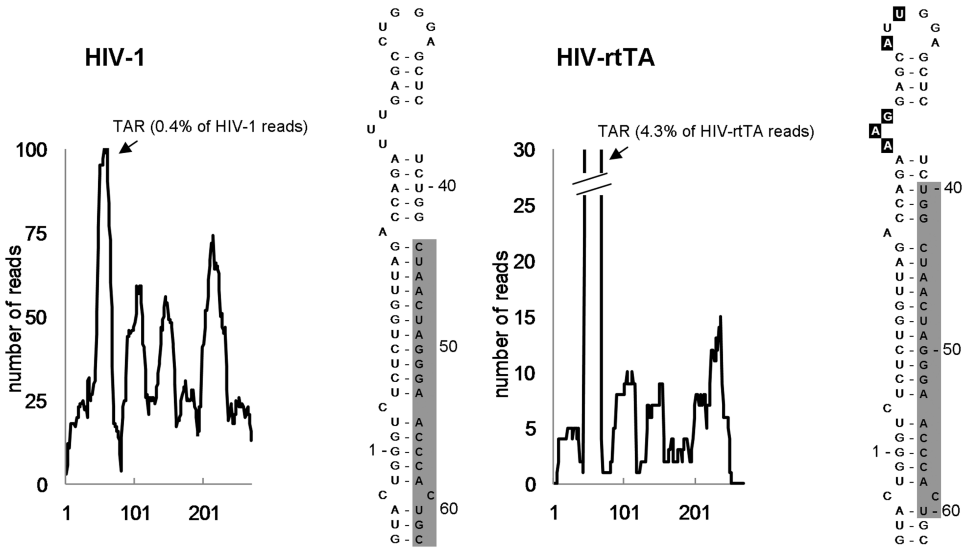

Figure 6. Multiple small RNA products are processed from the TAR hairpin. More TAR-derived small RNA reads were obtained from SupT1 cells infected with HIV-rtTA than HIV-1. The reads are different in sequence, all originating from the 3′ arm of the hairpin, but shifting toward the 3′-end of the hairpin for HIV-rtTA. The most abundant sequence is marked in gray in the 3′ TAR hairpin for each virus. Mutations within the TAR hairpin of HIV-rtTA are indicated as black boxed nucleotides.

Several vmiRNA are processed by the RNAi machinery

To investigate the role of the HIV-1-encoded miRNA candidates, we cloned them with 100-nt flanking sequence on both sides in expression constructs under control of the CMV promoter. First, the vmiRNAs were tested in luciferase reporter assays to assess whether the vmiRNAs can inhibit gene expression in a sequence-specific manner (Figure 7A). A specific luciferase reporter was constructed for each vmiRNA and its

Table 2. TAR-derived small RNA fragments

Position[a] (nt length)	HIV-1	HIV-rtTA
37–57 (21)	–	23×
38–54 (17)	6×	2×
38–58 (21)	7×	25×
39–54 (16)	7×	–
40–54 (15)	–	5×
40–59 (20)	–	8×
40–60 (21)	11×	28×[b]
40–61 (22)	–	10×
41–61 (21)	6×	–
43–62 (20)	16×[b]	–

[a]5′R coordinates only.
[b]Most abundant fragment.

activity in the absence of inhibitor was set at 100%. The vmiRNA targets, fully complementary to the potentially processed mature vmiRNA, together with ~15 nt of the neighboring sequence was cloned in the 3′-UTR of the luciferase reporter construct. Three of the five vmiRNAs showed target inhibition, vmiRNA-659 and 2413 were not able to knockdown gene expression of their target. All vmiRNAs were less active than the human miRNA-643 that served as a positive control. These results indicate that at least three of the HIV-1 vmiRNAs candidates are processed by the RNAi machinery into functional miRNAs.

Next, we studied the production of small RNAs from the candidate vmiRNA constructs. We decided to examine the small RNA production from the well characterized TAR hairpin (vmiRNA-43/9175), vmiRNA-659 and 3270. We chose the latter two because they represent a poorly active (659) and fairly active (3270) vmiRNA. HEK 293T cells were transfected with the vmiRNA constructs or the pLAI HIV-1 genome and we examined small RNA production by northern blotting. As negative controls, pBluescript (pBS) and an unrelated vmiRNA construct were used. Analysis confirmed the processing of TAR-derived vmiRNA-43/9175 in cells transfected with the HIV-1 molecular clone (pLAI) and the vmiRNA-43/9175 expression construct, with a fragment of ~58 nt corresponding to the TAR hairpin and a processed fragment of ~22 nt (Figure 7B, left panel). Notably, the TAR precursor was abundantly present in pLAI-transfected cells, while almost absent in the vmiRNA-43/9175 transfected cells. This could be due to the presence of TAR at both the 5′- and 3′-ends of the pLAI transcript. Poorly processive transcription from the viral LTR promoter may also yield a surplus of short TAR transcripts. Furthermore, the Tat protein is also produced by pLAI and subsequent binding to TAR may inhibit Dicer function. In contrast, the vmiRNA-43/9175 transcripts encode only a single TAR structure, the CMV promoter yields processive transcription complexes and the Tat protein is not present. These changes may explain the accumulation of the ~22 nt small RNA. For vmiRNA-659, we were not able to detect any small RNA products (data not shown), which corresponds with the poor activity in the luciferase assay (Figure 7A). For the vmiRNA-3270, we observed distinct small RNA products of ~42, ~32 and to a lesser extent of ~22 nt (Figure 7B, right panel). Although present in extremely small amounts, the 22 nt fragments could represent RNAi processing products. Small RNA products were not detected in pLAI-transfected cells, which could be due to lower expression from the LTR promoter compared to the CMV promoter. The vmiRNA-3769 is located in the pol gene and is thus expressed exclusively in the unspliced HIV-1 RNA. To test the impact of the vmiRNAs on HIV-1 production, we co-transfected the vmiRNA constructs with the HIV-1 molecular clone pLAI, but no significant effect on HIV-1 production was observed (results not shown). It is possible that the vmiRNAs affect the expression of cellular mRNAs, which should be addressed in future studies.

HIV-1 antisense-derived vsiRNAs

Relatively, few sequence reads were found that correspond to antisense HIV-1 sequences (Figure 3, lower panel). Similar antisense signals that cluster near the 3′-LTR region were detected in HIV-1 and HIV-rtTA-infected cells. Interestingly, higher levels of antisense versus sense RNAs were retrieved from the cells infected with HIV-rtTA (17.9%) than HIV-1 (1.9%). More than a quarter of these HIV-rtTA reads originated from the antisense TAR element, which in fact is also likely to adopt a hairpin structure (35,36). Peaks that map to the R region are indicated in both the 5′R and 3′R, but may be derived from only one of these regions. The majority of 3′-end antisense RNAs are likely to originate from dsRNA formed between HIV-1 sense and antisense transcripts that originate from a promoter inside the HIV-1 genome or a cellular promoter that is located downstream of the integrated HIV-1 provirus (Figure 1). Such an antisense promoter is encoded in the genome of several retroviruses including T-cell lymphotropic virus type 1, feline immunodeficiency virus, Friend murine leukemia and Moloney murine leukemia virus (37–41), but also HIV-1 (42,43). Antisense transcripts from a downstream cellular promoter have been described for integrated endogenous retroviral elements (16,44,45). If the latter scenario is true for HIV-1, the R-derived vsiRNAs are more likely derived from the 3′R.

HIV-1 production is affected by antisense vsiRNA

We hypothesized that the HIV-1-derived antisense sequences detected by SOLiD deep sequencing might represent vsiRNA processing products of dsRNA intermediates (Figure 1). We selected eight of these antisense fragments for further analyses (numbered in Figure 3, lower panel). The corresponding vsiRNAs were generated using *in vitro* transcription. In co-transfection experiments with the pLAI molecular clone, we tested these vsiRNAs for their ability to inhibit HIV-1 production (Figure 8A). Virus production was measured as the CA-p24 concentration in the supernatant and the value obtained in co-transfection with the control siLuc was set at 100%. The positive siNef control demonstrated potent inhibition, but the various vsiRNAs also triggered profound

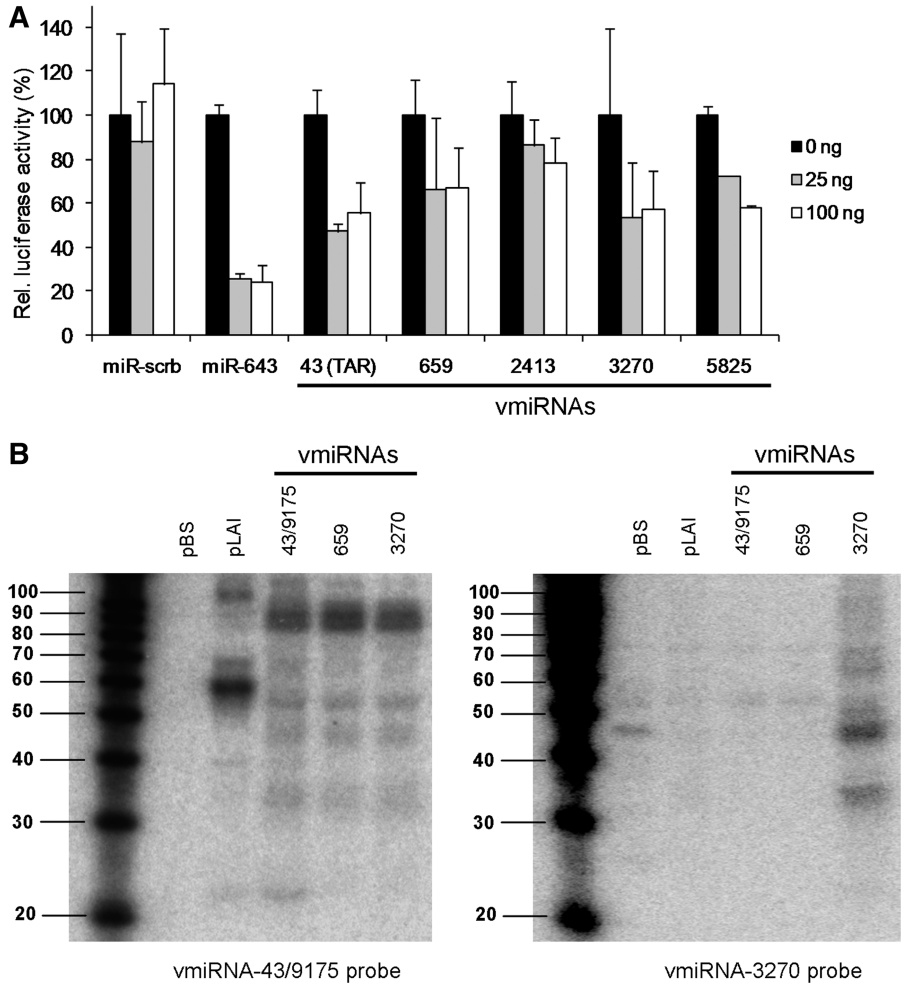

Figure 7. HIV-1 vmiRNA are functionally active. (**A**) Expression of vmiRNAs leads to target gene knockdown. 293T cells were co-transfected with 25 ng of the respective firefly luciferase reporter plasmid (encoding the vmiRNA target sequence), 0.5 ng of renilla luciferase plasmid, and various amounts of the corresponding vmiRNA constructs. Relative luciferase activity was determined as the firefly to renilla ratio. A scrambled miRNA (miRNA-scrb) construct served as negative control and the human encoded miRNA-643 was used as positive control. The values measured without miRNA were set at 100% for each luciferase reporter. (**B**) Northern blot analysis of the HIV-1-specific vmiRNAs. Small RNA was isolated from 293T cells transfected with the HIV-1 molecular clone pLAI and the indicated vmiRNA constructs. LNA oligonucleotide probes were used that are complementary to vmiRNA-43/9175 and 3270. Small RNA isolated from cells transfected with the irrelevant vmiRNA and pBS served as negative controls.

inhibition of virus production. Additionally, after co-transfection experiments with the pLAI and vsiRNA 7341 and vsiRNA 8200 a 5′-RACE PCR was performed to detect vsiRNA-mediated cleavage of the viral transcript. In both cases cleavage of the viral transcript was detected at the position of the vsiRNA. Next, we investigated whether we could neutralize this effect of the added vsiRNAs by inhibiting them with antagomirs, e.g. specific LNAs (Figure 8B). Antagomir (LNA) 9095 effectively antagonizes the inhibitory effect of vsiRNA 9095 whereas the control LNA molecule that targets an irrelevant sequence could not. Virus production without antagomir and vsiRNA was set at 100%. For the other four vsiRNAs, we determined the absolute amount of virus production in the absence and presence of the respective LNAs. The antagomirs could efficiently inhibit the added vsiRNAs, thus increasing virus production. We next wanted to probe whether inhibitory vsiRNAs are produced in naturally HIV-infected cells. However, cells with an integrated HIV-1 provirus cannot easily be obtained because active virus production leads to massive cell death. We therefore screened virus production in cells transfected with pLAI and the vsiRNA antagomirs. CA-p24 production as measured with the control LNA molecule was set at 100% (Figure 8C). All vsiRNA-antagomirs triggered a 2- to 4-fold increase in virus production, suggesting that endogenously produced vsiRNAs restrict HIV-1 gene expression.

DISCUSSION

Ever since the mechanism of RNAi was described as antiviral mechanism in plants, researchers have addressed the question whether or not RNAi contributes to the antiviral defense in other organisms, including mammals.

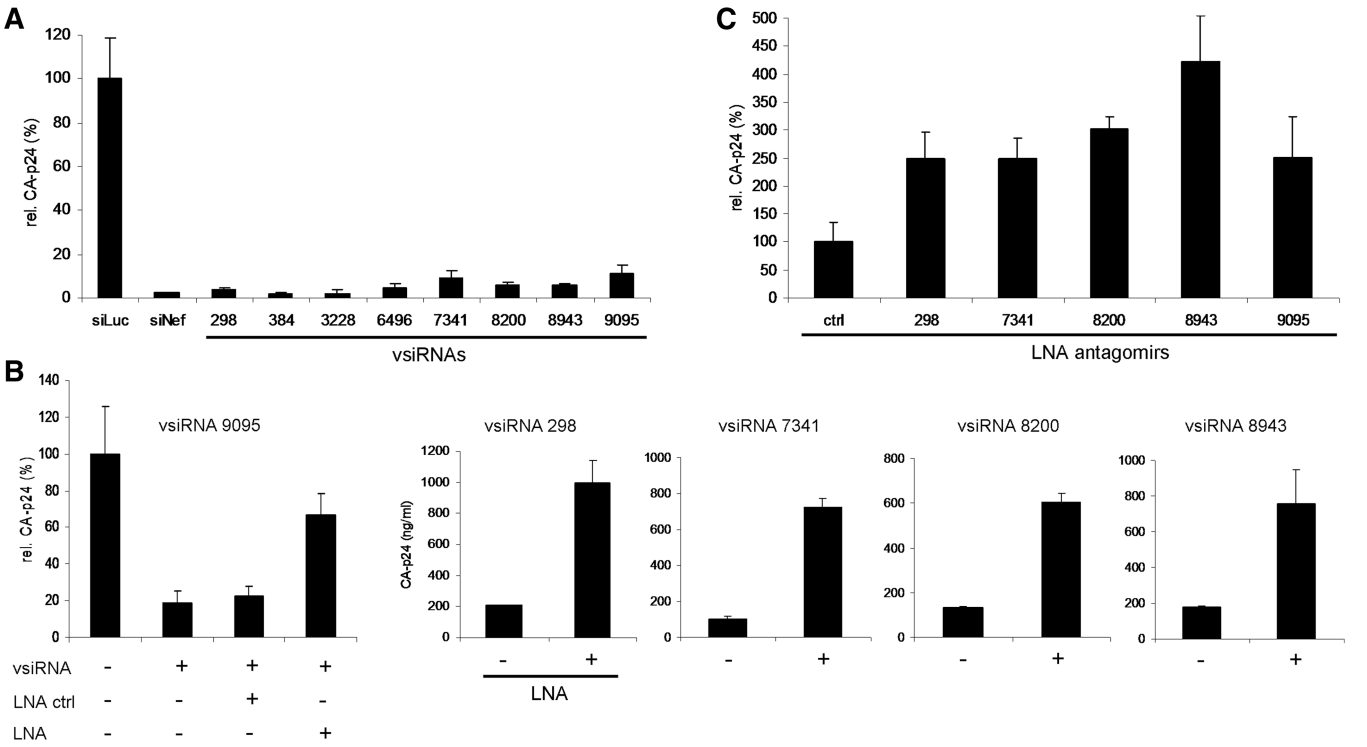

Figure 8. Virus production is inhibited by vsiRNAs. (**A**) 293T cells were co-transfected with 100 ng pLAI, 0.5 ng renilla luciferase plasmid and 50 nM of the indicated vsiRNA. CA-p24 levels in the culture supernatant were measured and renilla luciferase expression was measured to control for the transfection efficiency. The siRNA targeting firefly luciferase (siLuc) served as negative control to set virus production at 100% and the siRNA targeting the Nef region of HIV-1 was used as a positive control (siNef). (**B**) Both vsiRNAs and the corresponding antagomir (LNA) were co-transfected to investigate whether the antagomirs can antagonize the inhibitory effect of the vsiRNA. Virus production without antagomir or vsiRNA was set at 100%. An LNA control targeting an irrelevant sequence was used as a negative control. (**C**) The ability of the LNA antagomirs to block the endogenous vsiRNAs and to affect virus production was tested. An irrelevant vsiRNA antagomir (ctrl) was used as negative control and virus production in this setting was arbitrarily set at 100%.

Although the RNAi machinery is well conserved in eukaryotes and fully functional in mammals, researchers initially failed to detect the accumulation of short viRNAs in virus-infected mammalian cells. These viRNAs are considered the hallmark of antiviral RNAi. To counter the antiviral RNAi response, plant viruses encode RSSs (46). These RSSs were shown to suppress antiviral RNAi via sequestration of antiviral siRNAs, protection of long virus-specific dsRNA from processing into siRNAs or direct inhibition of specific components of the RNAi pathway. Despite the lack of viRNA detection in mammalian cells, some mammalian viruses do encode RSSs, which suggests that RNAi also plays a role as antiviral defense mechanism in mammals (47). Virus-derived small RNAs are likely rare in mammalian cells due to the absence of RNA-dependent RNA polymerase that amplifies siRNA production and thus the antiviral RNAi response.

The deep-sequencing technology is now enabling a more sensitive search for those antiviral viRNA molecules. One of the first indications in support of antiviral RNAi in mammals was the discovery that transposon-specific siRNAs, so-called endo-siRNAs, are produced in mouse oocytes to block transposon activity (16). Because retroviruses such as HIV-1 are related to endogenous retrotransposons, it is not unlikely that similar virus-specific viRNAs are produced in HIV-1-infected cells to modulate virus replication and to possibly contribute to viral latency. viRNAs were recently identified via deep sequencing in mammalian cells infected with HIV-1, Dengue virus, vesicular stomatitis virus, polio virus, hepatitis C virus and West Nile virus (15,17,18). Small virus-specific RNAs appear to accumulate in mammalian-infected cells only to ~1%, whereas virus-specific RNAs can accumulate to ~20% of the total of small RNA pool in plants (48). This difference in expression could explain why RNAi has a more significant role in antiviral responses in plants and insects compared to mammals and that viRNAs are more easily detected in those organisms. Although viRNAs were detected in early deep-sequencing studies in infected mammalian cells, most virus-specific sequence reads were observed only once. This greatly reduces the ability to perform quantitative comparisons. In this study, we used the latest deep-sequencing technology SOLiD^{TM} 3 Plus System to identify HIV-1-derived small RNAs in infected T cells. Using this extremely sensitive-sequencing technology, we were able to identify millions of sequence reads, most of which represent cellular miRNAs. The expression of some of these cellular miRNAs is strongly influenced by HIV-1 infection (Schopman *et al*, manuscript in preparation). Importantly, using this technology

we found similar sequence reads in two independent HIV-1-infected T cell cultures, demonstrating the robustness of this method. Furthermore, similar results were obtained in preliminary experiments with HIV-1-infected primary T lymphocytes.

Consistent with previous studies, we detected a small percentage of the reads (1%) that correspond to virus-specific small RNAs (15,17). The majority of these sequences were found to correspond to the HIV-1 plus-strand RNA, and we reasoned that abundant species could represent virus-encoded miRNAs, vmiRNAs. We propose that these small RNAs are processed from the viral RNA by Drosha and Dicer (Figure 1), although we cannot exclude other mechanisms (e.g. Dicer-mediated cleavage of dsRNA intermediates as proposed for vsiRNA generation). Mfold and MiPred analyses of the region encoding these abundant 18–21 nt sequence reads resulted in the identification of 5 vmiRNA candidates, including the previously described TAR vmiRNA (17,33). Three of these vmiRNAs are functional in terms of suppressing a luciferase reporter when made from an expression cassette and 2 vmiRNAs produced processed ~22 nt fragments that were detected on Northern blot. For the TAR vmiRNA we primarily detected the mature miRNA fragment derived from the 3′ side of the TAR hairpin, with some local variation in the actual cleavage sites. Interestingly, we found that TAR vmiRNAs were significantly overrepresented in cells infected with the HIV-rtTA variant. The HIV-rtTA variant contains TAR mutations that prevent Tat binding, which could enable more efficient recognition of this hairpin by the Dicer enzyme. The Tat protein has been shown to suppress RNAi in human cells (33,49,50). Changes in the TAR hairpin could lead to differences in processing by Drosha and/or Dicer, thus resulting in different reads (positions, numbers and length). Alternatively, exchange of the transcriptional components of the Tat-TAR system (HIV-1) by those of the Tet-on system (HIV-rtTA) may lead to non-processive transcription, thus increasing the relative concentration of short 5′-TAR transcripts. Previous studies showed that HIV-1 can manipulate the host cell RNAi pathway in order to promote viral latency and/or to suppress antiviral immune defenses (32,51). We hypothesize that the vmiRNAs originating from HIV-1 RNA are likely to affect virus replication in a positive manner. Although our preliminary results indicated that vmiRNA overexpression does not significantly affect virus production, further detailed analyses should reveal the biological functions of these HIV-1-specific vmiRNAs.

Another interesting finding was the profound antisense signal of a specific 18 nt RNA fragment that is fully complementary to the PBS sequence of the HIV-1 RNA genome. The cellular $tRNA^{Lys3}$ molecule binds the PBS to prime reverse transcription of the HIV-1 RNA genome into DNA, which can subsequently integrate into the host genome (29). Yeung *et al.* (17) were the first to report this specific small RNA in HIV-1-infected cells and they suggested that it represents a Dicer cleavage product that may have antiviral activity. However, we detected the same tRNA-derived small RNA in uninfected control cells. Therefore, this signal most likely represents processed tRNA rather than an siRNA processed by Dicer upon tRNA binding to the PBS. Recent studies have confirmed that cellular tRNAs can be processed by the RNAi machinery to yield 18 nt 3′-terminal fragments in the absence of virus (26). Whatever the origin, this RNA can potentially target HIV-1 transcripts for degradation by the RNAi machinery. Further analyses of the tRNA reads revealed a minor $tRNA^{Lys5a}$ variant besides the major $tRNA^{Lys3}$ primer species. This finding confirms previous biochemical and virological studies and this difference in efficiency of primer usage correlates with the intracellular concentration of the two tRNA molecules (28,52).

The accumulation of small antisense RNAs is restricted to the 3′-end of the HIV-1 genome. These vsiRNAs are likely derived from dsRNA intermediates that are processed by the RNAi machinery, although we cannot exclude alternative mechanisms of vsiRNA production. There are two possible routes for synthesis of such HIV-1 dsRNA molecules and antisense transcripts. Morris and co-workers (43) reported that the 3′-terminal part of the HIV-1 DNA genome encodes promoter activity that could generate antisense HIV-1 transcripts. Alternatively, the antisense transcripts could be produced via transcription from a cellular promoter of a neighboring gene at the site of provirus integration (Figure 1). The encoded HIV-1 antisense sequences are likely terminated at the polyadenylation signal in the HIV-1 genome. This explains why most antisense reads are restricted to the 3′ terminal segment. The viral sense and antisense transcripts could form extended dsRNA, which is a substrate for Dicer to produce vsiRNAs. We determined that transfection of *in vitro* made vsiRNAs transcripts potently inhibits virus production. Conversely, blocking of the vsiRNAs made in HIV-1 infected cells with LNA antagomirs stimulated virus production. We and others previously identified HIV-1 Tat to exhibit RSS activity and reported that Tat-mediated RNAi suppression is required for optimal virus production (49,50,47,53). These results are consistent with the notion that RNAi plays a role in innate antiviral defense and that HIV-1 needs to counter this mechanism. In this study, we identified the vsiRNA candidates in HIV-1-infected T cells that could be involved in the antiviral RNAi response.

Deep-sequencing technology provides a powerful tool to unravel the function of small RNAs in cell biology. Only few studies have used this technology to study small RNAs in virus-infected cells. Here, we presented the first results of a high-throughput small RNAs-sequencing project in HIV-1-infected cells. The combined results suggest that numerous small virus derived RNAs are produced in HIV-1-infected T cells that can influence the viral replication cycle.

ACKNOWLEDGEMENTS

We thank Stephan Heynen for performing CA-p24 ELISA. N.S., F.B., B.B. and J.H. participated in the design of the study. N.S., Y.P.L. and T.B. performed

the experiments. M.W. and Av.K. analyzed the data and N.S., B.B. and J.H. drafted the manuscript.

FUNDING

This research is sponsored by ZonMw (Translational Gene Therapy Program) and NWO-CW (Top grant). Funding for open access charge: NWO-CW.

Conflict of interest statement. None declared.

REFERENCES

1. Waterhouse,P.M., Wang,M.B. and Lough,T. (2001) Gene silencing as an adaptive defence against viruses. *Nature*, **411**, 834–842.
2. Voinnet,O. (2001) RNA silencing as a plant immune system against viruses. *Trends Genet*., **17**, 449–459.
3. Wilkins,C., Dishongh,R., Moore,S.C., Whitt,M.A., Chow,M. and Machaca,K. (2005) RNA interference is an antiviral defence mechanism in Caenorhabditis elegans. *Nature*, **436**, 1044–1047.
4. Wang,X.H., Aliyari,R., Li,W.X., Li,H.W., Kim,K., Carthew,R., Atkinson,P. and Ding,W. (2006) RNA interference directs innate immunity against viruses in adult Drosophila. *Science*, **312**, 452–454.
5. Segers,G.C., Zhang,X., Deng,F., Sun,Q. and Nuss,D.L. (2007) Evidence that RNA silencing functions as an antiviral defense mechanism in fungi. *Proc. Natl Acad. Sci. USA*, **104**, 12902–12906.
6. de Vries,W. and Berkhout,B. (2008) RNAi suppressors encoded by pathogenic human viruses. *Int. J. Biochem. Cell Biol*., **40**, 2007–2012.
7. Meister,G. and Tuschl,T. (2004) Mechanisms of gene silencing by double-stranded RNA. *Nature*, **431**, 343–349.
8. Boss,I.W., Plaisance,K.B. and Renne,R. (2009) Role of virus-encoded microRNAs in herpesvirus biology. *Trends Microbiol*., **17**, 544–553.
9. Pfeffer,S., Zavolan,M., Grasser,F.A., Chien,M., Russo,J.J., Ju,J., John,B., Enright,A.J., Marks,D., Sander,C. *et al.* (2004) Identification of virus-encoded microRNAs. *Science*, **304**, 734–736.
10. Umbach,J.L., Kramer,M.F., Jurak,I., Karnowski,H.W., Coen,D.M. and Cullen,B.R. (2008) MicroRNAs expressed by herpes simplex virus 1 during latent infection regulate viral mRNAs. *Nature*, **454**, 780–783.
11. Wu,Q., Luo,Y., Lu,R., Lau,N., Lai,E.C., Li,W.X. and Ding,S.W. (2010) Virus discovery by deep sequencing and assembly of virus-derived small silencing RNAs. *Proc. Natl Acad. Sci. USA*, **107**, 1606–1611.
12. Fischer,W., Ganusov,V.V., Giorgi,E.E., Hraber,P.T., Keele,B.F., Leitner,T., Han,C.S., Gleasner,C.D., Green,L., Lo,C.C. *et al.* (2010) Transmission of single HIV-1 genomes and dynamics of early immune escape revealed by ultra-deep sequencing. *PLoS ONE*, **5**.
13. Reese,T.A., Xia,J., Johnson,L.S., Zhou,X., Zhang,W. and Virgin,H.W. (2010) Identification of novel microRNA-like molecules generated from herpesvirus and host tRNA transcripts. *J. Virol*., **84**, 10344–10353.
14. Metzker,M.L. (2010) Sequencing technologies - the next generation. *Nat. Rev. Genet*., **11**, 31–46.
15. Parameswaran,P., Sklan,E., Wilkins,C., Burgon,T., Samuel,M.A., Lu,R., Ansel,K.M., Heissmeyer,V., Einav,S., Jackson,W. *et al.* (2010) Six RNA viruses and forty-one hosts: viral small RNAs and modulation of small RNA repertoires in vertebrate and invertebrate systems. *PLoS. Pathog*., **6**, e1000764.
16. Watanabe,T., Totoki,Y., Toyoda,A., Kaneda,M., Kuramochi-Miyagawa,S., Obata,Y., Chiba,H., Kohara,Y., Kono,T., Nakano,T. *et al.* (2008) Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature*, **453**, 539–543.
17. Yeung,M.L., Bennasser,Y., Watashi,K., Le,S.Y., Houzet,L. and Jeang,K.T. (2009) Pyrosequencing of small non-coding RNAs in HIV-1 infected cells: evidence for the processing of a viral-cellular double-stranded RNA hybrid. *Nucleic Acids Res*., **37**, 6575–6586.
18. Lefebvre,G., Desfarges,S., Uyttebroeck,F., Munoz,M., Beerenwinkel,N., Rougemont,J., Telenti,A. and Ciuffi,A. (2011) Analysis of HIV-1 expression level and sense of transcription by high-throughput sequencing of the infected cell. *J. Virol*., **85**, 6205–6211.
19. Peden,K., Emerman,M. and Montagnier,L. (1991) Changes in growth properties on passage in tissue culture of viruses derived from infectious molecular clones of HIV-1_{LAI}, HIV-1_{MAL}, and HIV-1_{ELI}. *Virology*, **185**, 661–672.
20. Zhou,X., Vink,M., Klaver,B., Berkhout,B. and Das,A.T. (2006) Optimization of the Tet-On system for regulated gene expression through viral evolution. *Gene Ther*., **13**, 1382–1390.
21. Schopman,N.C., Ter Brake,O. and Berkhout,B. (2010) Anticipating and blocking HIV-1 escape by second generation antiviral shRNAs. *Retrovirology*, **7**, 52.
22. Ruijter,J.M., Thygesen,H.H., Schoneveld,O.J., Das,A.T., Berkhout,B. and Lamers,W.H. (2006) Factor correction as a tool to eliminate between-session variation in replicate experiments: application to molecular biology and retrovirology. *Retrovirology*, **3**, 1–8.
23. Schopman,N.C., Liu,Y.P., Konstantinova,P., Ter Brake,O. and Berkhout,B. (2010) Optimization of shRNA inhibitors by variation of the terminal loop sequence. *Antiviral Res*., **86**, 204–211.
24. Berkhout,B., Silverman,R.H. and Jeang,K.T. (1989) Tat trans-activates the human immunodeficiency virus through a nascent RNA target. *Cell*, **59**, 273–282.
25. Verhoef,K., Marzio,G., Hillen,W., Bujard,H. and Berkhout,B. (2001) Strict control of human immunodeficiency virus type 1 replication by a genetic switch: Tet for Tat. *J. Virol*., **75**, 979–987.
26. Cole,C., Sobala,A., Lu,C., Thatcher,S.R., Bowman,A., Brown,J.W., Green,P.J., Barton,G.J. and Hutvagner,G. (2009) Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs. *RNA*, **15**, 2147–2160.
27. Haussecker,D., Huang,Y., Lau,A., Parameswaran,P., Fire,A.Z. and Kay,M.A. (2010) Human tRNA-derived small RNAs in the global regulation of RNA silencing. *RNA*, **16**, 673–695.
28. Schopman,N.C., Heynen,S., Haasnoot,J. and Berkhout,B. (2010) A miRNA-tRNA mix-up: tRNA origin of proposed miRNA. *RNA. Biol*., **7**, 573–576.
29. Abbink,T.E. and Berkhout,B. (2008) HIV-1 reverse transcription initiation: a potential target for novel antivirals? *Virus Res*., **134**, 4–18.
30. Jiang,P., Wu,H., Wang,W., Ma,W., Sun,X. and Lu,Z. (2007) MiPred: classification of real and pseudo microRNA precursors using random forest prediction model with combined features. *Nucleic Acids Res*., **35**, W339–W344.
31. Bennasser,Y., Le,S.Y., Yeung,M.L. and Jeang,K.T. (2004) HIV-1 encoded candidate micro-RNAs and their cellular targets. *Retrovirology*, **1**, 43.
32. Klase,Z., Winograd,R., Davis,J., Carpio,L., Hildreth,R., Heydarian,M., Fu,S., McCaffrey,T., Meiri,E., yash-Rashkovsky,M. *et al.* (2009) HIV-1 TAR miRNA protects against apoptosis by altering cellular gene expression. *Retrovirology*, **6**, 18.
33. Klase,Z., Kale,P., Winograd,R., Gupta,M.V., Heydarian,M., Berro,R., McCaffrey,T. and Kashanchi,F. (2007) HIV-1 TAR element is processed by Dicer to yield a viral micro-RNA involved in chromatin remodeling of the viral LTR. *BMC. Mol. Biol*., **8**, 63.
34. Ouellet,D.L., Plante,I., Landry,P., Barat,C., Janelle,M.E., Flamand,L., Tremblay,M.J. and Provost,P. (2008) Identification of functional microRNAs released through asymmetrical processing of HIV-1 TAR element. *Nucleic Acids Res*., **36**, 2353–2365.
35. Berkhout,B., Vastenhouw,N.L., Klasens,B.I. and Huthoff,H. (2001) Structural features in the HIV-1 repeat region facilitate strand transfer during reverse transcription. *RNA*, **7**, 1097–1114.
36. Klasens,B.I.F., Huthoff,H.T., Das,A.T., Jeeninga,R.E. and Berkhout,B. (1999) The effect of template RNA structure on

elongation by HIV-1 reverse transcriptase. *Biochim. Biophys. Acta*, **1444**, 355–370.
37. Larocca,D., Chao,L.A., Seto,M.H. and Brunck,T.K. (1989) Human T-cell leukemia virus minus strand transcription in infected T-cells. *Biochem. Biophys. Res. Commun.*, **163**, 1006–1013.
38. Gaudray,G., Gachon,F., Basbous,J., Biard-Piechaczyk,M., Devaux,C. and Mesnard,J.M. (2002) The complementary strand of the human T-cell leukemia virus type 1 RNA genome encodes a bZIP transcription factor that down-regulates viral transcription. *J. Virol.*, **76**, 12813–12822.
39. Cavanagh,M.H., Landry,S., Audet,B., Arpin-Andre,C., Hivin,P., Pare,M.E., Thete,J., Wattel,E., Marriott,S.J., Mesnard,J.M. *et al.* (2006) HTLV-I antisense transcripts initiating in the 3′LTR are alternatively spliced and polyadenylated. *Retrovirology*, **3**, 15.
40. Briquet,S., Richardson,J., Vanhee-Brossollet,C. and Vaquero,C. (2001) Natural antisense transcripts are detected in different cell lines and tissues of cats infected with feline immunodeficiency virus. *Gene*, **267**, 157–164.
41. Rasmussen,M.H., Ballarin-Gonzalez,B., Liu,J., Lassen,L.B., Fuchtbauer,A., Fuchtbauer,E.M., Nielsen,A.L. and Pedersen,F.S. (2010) Antisense transcription in gammaretroviruses as a mechanism of insertional activation of host genes. *J. Virol.*, **84**, 3780–3788.
42. Landry,S., Halin,M., Lefort,S., Audet,B., Vaquero,C., Mesnard,J.M. and Barbeau,B. (2007) Detection, characterization and regulation of antisense transcripts in HIV-1. *Retrovirology*, **4**, 71.
43. Morris,K.V., Santoso,S., Turner,A.M., Pastori,C. and Hawkins,P.G. (2008) Bidirectional transcription directs both transcriptional gene activation and suppression in human cells. *PLoS Genet.*, **4**, e1000258.
44. Ghildiyal,M. and Zamore,P.D. (2009) Small silencing RNAs: an expanding universe. *Nat. Rev. Genet.*, **10**, 94–108.
45. Yang,N. and Kazazian,H.H. Jr (2006) L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. *Nat. Struct. Mol. Biol.*, **13**, 763–771.
46. Kasschau,K.D. and Carrington,J.C. (1998) A counterdefensive strategy of plant viruses: suppression of posttranscriptional gene silencing. *Cell*, **95**, 461–470.
47. Haasnoot,J., Westerhout,E.M. and Berkhout,B. (2007) RNA interference against viruses: strike and counterstrike. *Nat. Biotechnol.*, **25**, 1435–1443.
48. Qi,X., Bao,F.S. and Xie,Z. (2009) Small RNA deep sequencing reveals role for Arabidopsis thaliana RNA-dependent RNA polymerases in viral siRNA biogenesis. *PLoS ONE*, **4**, e4971.
49. Haasnoot,J., de Vries,W., Geutjes,E.J., Prins,M., de Haan,P. and Berkhout,B. (2007) The Ebola virus VP35 protein is a suppressor of RNA silencing. *PLoS Pathog.*, **3**, e86.
50. Bennasser,Y., Le,S.Y., Benkirane,M. and Jeang,K.T. (2005) Evidence that HIV-1 encodes an siRNA and a suppressor of RNA silencing. *Immunity*, **22**, 607–619.
51. Huang,J., Wang,F., Argyris,E., Chen,K., Liang,Z., Tian,H., Huang,W., Squires,K., Verlinghieri,G. and Zhang,H. (2007) Cellular microRNAs contribute to HIV-1 latency in resting primary CD4(+) T lymphocytes. *Nat. Med.*, **13**, 1241–1247.
52. Das,A.T., Klaver,B. and Berkhout,B. (1995) Reduced replication of human immunodeficiency virus type 1 mutants that use reverse transcription primers other than the natural tRNA(3Lys). *J. Virol.*, **69**, 3090–3097.
53. Qian,S., Zhong,X., Yu,L., Ding,B., de Haan,P. and Boris-Lawrie,K. (2009) HIV-1 Tat RNA silencing suppressor activity is conserved across kingdoms and counteracts translational repression of HIV-1. *Proc. Natl Acad. Sci. USA*, **106**, 605–610.
54. Michael,N.L., Vahey,M.T., D'Arcy,L., Ehrenberg,P.K., Mosca,J.D., Rappaport,J. and Redfield,R.R. (1994) Negative-strand RNA transcripts are produced in human immunodeficiency virus type 1-infected cells and patients by a novel promoter downregulated by Tat. *J. Virol.*, **68**, 979–987.