Title

Assays for precise quantification of total (including short) and elongated HIV-1 transcripts

Authors

Kaiser, Philipp
Joshi, Sunil K
Kim, Peggy
et al.

Publication Date

2017-04-01

DOI

10.1016/j.jviromet.2016.12.017

Peer reviewed

HHS Public Access

Author manuscript

J Virol Methods. Author manuscript; available in PMC 2018 April 01.

Published in final edited form as:

J Virol Methods. 2017 April ; 242: 1–8. doi:10.1016/j.jviromet.2016.12.017.

Assays for precise quantification of total (including short) and elongated HIV-1 transcripts

Philipp Kaiser[a,c], **Sunil K. Joshi**[a], **Peggy Kim**[a], **Peilin Li**[a], **Hongbing Liu**[b], **Andrew P. Rice**[b], **Joseph K. Wong**[a], and **Steven A. Yukl**[a,*]

[a]San Francisco Veterans Affairs (VA) Medical Center and University of California, San Francisco (UCSF), 4150 Clement Street, 111W, San Francisco, CA, 94121, USA

[b]Department of Molecular Virology and Microbiology, Baylor College of Medicine, One Baylor Plaza, MS BCM385, Houston, TX, 77030, USA

Abstract

Despite intensive study, it is unclear which mechanisms are responsible for latent HIV infection *in vivo*. One potential mechanism is inhibition of HIV transcriptional elongation, which results in short abortive transcripts containing the trans-activation response (TAR) region. Because the relative levels of total (including short) and processive transcripts provide measures of HIV transcriptional initiation and elongation, there is a compelling need for techniques that accurately measure both. Nonetheless, prior assays for total transcripts have been semi-quantitative and have seen limited application to patient samples. This manuscript reports the validation of quantitative reverse transcription (RT) droplet digital PCR assays for measurement of total (TAR) and processive (R-U5/gag) HIV transcripts. Traditional RT priming strategies can efficiently detect the TAR region on long HIV transcripts but detect <4% of true short transcripts. The TAR assay presented here utilizes an initial polyadenylation step, which provides an accessible RT priming site and detects short and long transcripts with approximately equal efficiency (70%). By applying these assays to blood samples from 8 ART-treated HIV+ individuals, total HIV transcripts were detected at levels >10-fold higher than elongated transcripts, implying a substantial block to transcriptional elongation *in vivo*. This approach may be applied to other difficult-to-prime RNA targets.

Keywords

HIV-1; Latency; Transcription; Elongation; TAR; Short or abortive transcript

1. Introduction

One of the major barriers to eradication of HIV is the ability of HIV-1 to establish latent infection in CD4+T cells (Chun et al., 1997; Finzi et al., 1997; Wong et al., 1997). The reversible inhibition of HIV gene expression in latently-infected cells prevents death by intrinsic cell defenses or extrinsic immune responses, allowing infectious provirus to persist

*Corresponding author at: 4150 Clement St, 111W, San Francisco, CA 94121, USA., Steven.yukl@ucsf.edu (S.A. Yukl).
[C]Current address: Department of Infectious Diseases, Luzerner Kantonsspital, 6000 Lucerne 16, Switzerland

in long-lived cells and be passed on (and multiply) through cell proliferation. Multiple different mechanisms have been implicated in latency, but despite intensive study, most of which has been performed in various cell line model systems, it is unclear which mechanism or mechanisms predominate *in vivo*. In addition to various blocks at the level of transcriptional initiation, latent HIV infection has been attributed to inhibition of HIV-1 transcriptional elongation, perhaps due to insufficient activity of cellular host elongation factors or the viral transcription factor Tat (Adams et al., 1994; Emiliani et al., 1998; Kao et al., 1987; Lin et al., 2003; Natarajan et al., 2013; Sonza et al., 2002; Taube and Peterlin, 2013; Yukl et al., 2009). The hallmark of inefficient elongation of HIV transcription is the presence of short abortive HIV transcripts (Adams et al., 1994; Budhiraja and Rice, 2013). These non-polyadenylated RNAs comprise a minimum of 58nt from the HIV transcriptional start site (HXB2 positions 456–513) and form a tight stem-loop structure corresponding to the trans-activation response element (TAR) (Fig. 1). Because of their extensive internal base pairing, these abortive transcripts are thought to be partially protected from degradation when released from the transcriptional machinery, and their accumulation has been reported in cell line models (Kao et al., 1987) as well as in peripheral blood lymphocytes of HIV-1 infected individuals on ART (Adams et al., 1994; Adams et al., 1999; Lassen et al., 2004). In the context of HIV cure research and the quest for HIV latency-reversing drugs, there is a compelling need to develop techniques that can accurately quantify both short and processive transcripts in clinical samples, since their relative levels can be used to infer the levels of HIV transcriptional initiation and inhibition of elongation.

The existence of short transcripts can be demonstrated directly in model systems by RNAse protection assays, but detection in clinical samples requires more sensitive, amplification-based methods. Semi-quantitative detection by end-point reverse transcription (RT) polymerase chain reaction (RT-PCR) was used in the seminal papers (Adams et al., 1994; Adams et al., 1999). The subsequent development of more accurate PCR-based quantification techniques, such as real-time PCR (qPCR) and droplet digital PCR (ddPCR), may allow for more reliable quantification of HIV short abortive transcripts. This manuscript reports the development of RT-ddPCR-based assays for quantification of total (including short) and processive transcripts and provides in-depth validation of the efficiencies of both the RT and PCR steps.

Initial attempts to quantify TAR-containing HIV RNAs in clinical samples by RT qPCR gave unexpectedly low readouts, often not exceeding copy numbers of processive transcripts, suggesting virtual non-existence of any short abortive HIV RNA species. Notably, these results were comparable to recent findings of other groups (Anderson et al., 2015, Seattle, USA). However, subsequent control experiments to quantify the TAR region on standards of different lengths suggested a heavy bias towards detection of long standards and highlighted the difficulties in reverse transcription of short abortive HIV transcripts, likely due to their limited length and highly structured nature.

Although it has been well recognized that the efficiency of RT is subject to great variability and difficult to predict (Bustin et al., 2015), RT is often neglected in validation and standardization of RT-PCR assays, as reflected in the widespread use of DNA for external standards. Important factors affecting RT efficiency are the characteristics of the RT enzyme,

the presence of inhibitors, the concentration of background RNA, the abundance of the target, the sequence/structure of the target, and its sequence environment (Sanders et al., 2013; Stahlberg et al., 2004a,b). However, in practical terms, much depends on the choice of an appropriate RT priming strategy. Traditionally, random oligo priming (mostly with hexamers) is chosen to achieve high cDNA yields and the most unbiased representation of RNA sequences. Priming with poly-T oligonucleotides increases specificity for mRNAs at the cost of underrepresentation of targets located far upstream on the mRNA template (Bustin and Nolan, 2004). The use of gene-specific RT primers is reported to allow for the highest sensitivity for a particular target and can be used in one-step RT-PCR procedures. However, priming during RT is inherently less specific than priming during PCR (Haddad et al., 2007). The low priming requirements and low working temperatures of most RT enzymes result in nonspecific priming by short RNA or DNA fragments, RNA-self priming, and also false priming, problems which become particularly evident when strand-specific RT-PCR assays are required (Tuiskunen et al., 2010). Moreover, the presence and location of secondary RNA structures can affect both priming and processivity of RT. Hindrance by secondary structure can be remedied by use of more thermostable enzymes and higher reaction temperatures (Malboeuf et al., 2001) and by choosing priming sites outside the secondary structure. Therefore, nonspecific priming strategies may have advantages for RT of highly structured targets (Bustin and Nolan, 2004). The method presented here introduces an initial polyadenylation step, providing an easy accessible RT priming site for short abortive HIV transcripts, an approach that may be applied to enhance RT efficiency of other difficult-to-prime RNA targets.

2. Materials and methods

2.1. Biological samples and RNA extraction

Peripheral blood mononuclear cells (PBMCs) were recovered by Ficoll density gradient centrifugation (Yukl et al., 2010) from leukoreduction filters (healthy donors) or fresh venous blood (8 HIV+ individuals). The study was approved by the local Institutional Review Board of the University of California, San Francisco (UCSF) and the San Francisco VA. All participants gave informed consent. All HIV+ individuals were male, on ART, and had plasma HIV RNA levels <40 copies/ml. Total cellular RNA was recovered by acid guanidinium thiocyanate-phenol-chloroform extraction (Trireagent) and subsequent isopropanol precipitation. RNA was quantified using UV spectrophotometry (NanoDrop 1000).

2.2. Preparation of HIV standards

Validation experiments were performed using HIV plasmid DNA and three different HIV RNA standards: 1) a “short” 58nt TAR loop sequence (HXB2 positions 456–513) corresponding to abortive transcripts; 2) a “long” 2.5 kb HIV RNA standard (positions 326–2844); and 3) a “virion” standard containing full length HIV genomic RNA. “Short” and “long” standards were prepared by *in vitro* transcription. Briefly, the template for the “short” standard RNA was generated by PCR from pNL4-3 (Adachi et al., 1986) using primer TAR F7 (Table 1) tagged with a T7 promoter sequence and TAR R6 followed by gel purification. The template sequence for the “long” standard RNA was generated by nested PCR from

pNL4-3 (HXB2 positions 326–2844). The PCR product was cloned into a pcDNA3.1/V5-His TOPO TA Vector (Invitrogen), expanded in TOP10 cells (Invitrogen), and linearized with Pme1. Subsequently, templates were transcribed using the T7 MEGAscript Kit (Invitrogen). *In vitro* transcribed RNA standards were treated with DNAse I, purified using the miRNeasy Kit (Qiagen), and quantified based on molecular weights and replicate measurement of RNA mass using the NanoDrop 1000. The correct length and integrity of each standard was confirmed using the Agilent RNA 6000 Nano assay. The "virion" standard was prepared from stocks of NL4-3 virus (supernatants from infected cells) by sequential freeze-thaw (to lyse any residual cells but not virions (Fischer et al., 2002; Kaiser et al., 2006)), DNase/RNase treatment (to eliminate free nucleic acids), RNA extraction (using the QIAgen Viral RNA Mini Kit), and subsequent quantification by replicate measurements using the Abbott RealTime HIV-1 Assay. Standards were diluted to the appropriate copy numbers and mixed with cellular RNA extracted from donor PBMCs.

2.3. Reverse transcription procedures

The polyadenylation reaction was carried out in a volume of 20 μl containing 3 μl of 10× Superscript III buffer (Invitrogen) [which results in a final buffer solution of 30 mM TrisHCl (pH 8.4) and 75 mM KCl], 3 μl of 50 mM $MgCl_2$, 1 μl of 10 mM ATP (Epicentre), 2 μl of 4U/μl poly-A polymerase (Epicentre), and 1 μl of 40U/μl RNAseOUT (Invitrogen) at 37.0 °C for 45 min. Subsequently, the reaction mixes were stored at 4.0 °C and the remaining RT reagents were added in a volume of 10 μl containing 1.5 μl of 10 mM dNTPs (Invitrogen), 1.5 μl of 50 ng/μl random hexamers (Invitrogen), 1.5 μl of 50 μM oligo dT15, and 1 μl of 200U/μl Superscript III reverse transcriptase (Invitrogen). To prevent contamination, a short spin was introduced before the addition of RT reagents. The use of oligo dT15 instead of longer dT primers proved to be crucial in order to avoid subsequent inhibition of PCR. For comparison of different priming strategies, the same conditions were used but with water instead of poly-A polymerase and with either no primers (water instead of random hexamers and dT15), random hexamers (water instead of dT15), or gene-specific priming (no hexamers or dT15, but primer TAR R6 at a final concentration of 0.1 μM). For quantification of longer, processive transcripts (R-U5/Gag region), the polyadenylation step was omitted and RT was performed in a volume of 50 μl containing 5 μl of 10× RT buffer, 5 μl of 50 mM $MgCl_2$, 2.5 μl of 50 ng/μl random hexamers, 2.5 μl of 50 μM dT15, 2.5 μl of 10 mM dNTPs, 1.25 μl of 40U/μl RNAseOUT, and 2.5 μl of 200U/μl Superscript III RT. For all assays, RT was performed in a conventional thermocycler at 25.0 °C for 10 min, 50.0 °C for 50 min, followed by an inactivation step at 85.0 °C for 5 min.

2.4. Digital droplet PCR

TAR loop specific primers were designed according to the constraints imposed by target length and HIV subtype B sequence conservation and candidates were evaluated according to the efficiency of the reaction. A published PCR primer-probe combination specific for the R-U5/gag region of the HIV 5′LTR (Kumar et al., 2007; Yukl et al., 2013; Yukl et al., 2014) was used to quantify elongated HIV transcripts. Primer and probe sequences are listed in Table 1. Digital droplet PCR (ddPCR) was performed using the QX100 Droplet Digital qPCR System (BioRad, Hercules, CA). Samples were tested in replicate (at least 2) and each reaction consisted of a 20 μl solution containing 10 μl of ddPCR Probe Supermix, 900 nM of

primers, 250 nM of probe, and 5 μl of undiluted RT product. Droplets were amplified using a 7900 Thermal Cycler (Life Technologies) with the following cycling conditions: 10 min at 95 °C, 45 cycles of 30 s at 95 °C and 59 °C for 60 s, and a final droplet cure step of 10 min at 98 °C. Droplets were read and analyzed using the QuantaSoft software in the absolute quantification mode.

2.5. Calculations and statistics

For any given input of RNA standard, efficiencies of detection were expressed as ratios of absolute cDNA copies detected by ddPCR to the number of input RNA molecules. The overall efficiency of each assay was calculated by plotting the absolute cDNA copies detected by ddPCR versus the input of RNA standard and determining the slope using linear regression analysis. Comparisons between assays were performed using the Wilcoxon signed-rank test. All statistics were performed using GraphPad Prism 5.0.

3. Results

3.1. Different RT priming strategies for RNAs containing the HIV TAR loop sequence

Since the TAR loop sequence is present at the 5′ and the 3′ end of any full length HIV transcript, quantification of short abortive transcripts must be inferred by relating measurements of the TAR loop target to those of targets corresponding to processive transcripts downstream of the assumed HIV transcription pausing sites. In order to provide meaningful data, both assays have to be similarly quantitative and sensitive, and the detection of either target has to be independent of the length of the RNA on which it resides. In order to test the efficiency of reverse transcription of the TAR loop sequence in the context of different sequence environments, standard RNAs encompassing HXB2 positions 456–513 (short standard) and 326–2844 (long standard) were converted into cDNA under different RT conditions. Subsequently, cDNA was quantified by ddPCR with primers and probes specific for the TAR loop and the R-U5/gag region (Table 1). The choice of a digital PCR quantification method was based on its advantage of absolute cDNA quantification, which allows for easy comparison of reverse transcription efficiencies.

Detection efficiency and linearity of quantification of the ddPCR step of the TAR and R-U5/gag assays were assessed using standard dilutions of a plasmid containing HIV DNA sequence (HXB2 positions 326–2844). The TAR loop ddPCR assay displayed a detection efficiency of 76%, whereas the R-U5/gag assay had a similar efficiency of 75% ($p = 0.31$; Wilcoxon test) (Fig. 2). Both assays showed excellent linearity over 4 orders of magnitude ($r^2 \geq 0.998$).

Random RT priming is inherently unlikely to yield significant amounts of full-length cDNA. Since abortive transcripts are reported to be as short as 57–59 nucleotides (Kao et al., 1987), which is roughly the minimal amplicon length to host a Taqman assay, randomly primed RT-PCR is bound to underestimate the presence of these HIV RNAs. Although previous studies reported abundant amounts of short abortive transcripts in clinical samples using RT primed by random hexamers (Adams et al., 1994; Adams et al., 1999), based on the above considerations, gene-specific priming appeared more attractive. Surprisingly, RT of the short

standard RNA was extremely inefficient with either random hexamers or specific primers (Fig. 3A). Even with a gene-specific primer, no more than 3.8% of all RNA molecules were detected, whereas the use of random hexamers resulted in even lower (0.1%) detection efficiencies. Conversely, random priming of the long (2519nt) HIV RNA fragment appeared very efficient. Remarkably, even in the absence of any primer for the RT step, a substantial proportion of the long standard was detected, illustrating the low priming requirements for RT, which favors nonspecific or self-priming events (Fig. 3A). The low detection efficiency of the short standard RNA may be explained by the high predicted melting temperature (83.0 °C) of its stem-loop structure, which likely restricts any primer binding. The use of a more thermostable enzyme (ThermoScript, Invitrogen) at up to 70 °C as well as several alternative first strand primers did not increase detection efficiency (data not shown). In contrast, a very high detection efficiency of the long standard RNA suggested that the TAR stem loop secondary structure can be overcome if RT is initiated further downstream on the RNA molecule.

3.2. Generating RT primer binding sites by *in vitro* polyadenylation

In vitro polyadenylation followed by oligo dT primed RT has been used for quantification of microRNAs and other targets too short to accommodate a full Taqman assay (Shi and Chiang, 2005). Based on the observation that the TAR loop sequence can be reverse transcribed efficiently if the reaction is primed beyond the stem loop, we adapted this approach in order to create an accessible RT priming site for short abortive HIV transcripts. Poly-dT priming is known to yield cDNA biased towards the 3′end, especially of long RNAs (Bustin and Nolan, 2004), which may lead to an underrepresentation of both TAR and 5′LTR targets on full-length HIV RNAs. To avoid this problem, a mixture of random hexamers and poly-dT was used for RT priming. This procedure allowed for a dramatic improvement of TAR detection efficiency of the short standard RNA (Fig. 3A), whereas the detection efficiency on the long RNA standard was slightly but not relevantly reduced compared to priming without preceding polyadenylation (median 72% vs. 81%). In summary, introduction of an initial polyadenylation step allowed for reliable detection of TAR with a procedure that works efficiently no matter whether the target sequence is placed on a very short or a long transcript. In contrast to the results for the TAR assay, addition of the polyadenylation step (compared to no polyadenylation) resulted in variable inhibition (from no inhibition to 3-fold) of the R-U5/gag assay when tested on long and virion standards (data not shown). For this reason, subsequent experiments using the R-U5/gag assay were performed without the polyadenylation step.

3.3. Validation experiments

Due to the low frequency of HIV-infected cells, HIV transcripts are rare in clinical samples, especially in cells from patients on ART, so comparatively large amounts of total RNA per assay are needed to measure significant amounts of a particular HIV transcript. Since polyadenylation and RT are separate reactions, which may be saturated at different RNA concentrations, the TAR assay was tested on short and long HIV standard RNAs in a background of different amounts of total cellular RNA from human PBMCs. The efficiency of detection of the short standard was more susceptible to inhibition by background RNA, which occurred at levels ≥1 μg (data not shown), suggesting that both polyadenylation as

well as reverse transcription are rate-limiting steps. Of note, commercial kits for microRNA reverse transcription (such as Qiagen's miScript II RT Kit) that allow polyadenylation and RT as one-step procedure showed divergent RT efficiencies of short and long standard RNA at total RNA inputs of less than 100 ng per reaction (data not shown). Compared to the TAR assay, the R-U5/gag assay (performed without polyadenylation) was less susceptible to inhibition by background RNA and tolerated inputs as high as 5 μg (data not shown).

Detection efficiency and linearity of quantification of the polyadenylation-RT-ddPCR procedure for TAR were assessed using standard dilutions of the short and the long standard RNA in 1 μg PBMC RNA (Fig. 3B). The TAR loop assay had a similar overall detection efficiency of 71% for the short standard and 70% for the long standard RNA ($p = 0.25$, Wilcoxin test). Again, both assays showed excellent linearity ($r^2 > 0.97$) over 3 orders of magnitude (Fig. 3B). These efficiencies are almost as high as those obtained with plasmid DNA standards (Fig. 2), suggesting that under these conditions, reverse transcription is almost complete.

To determine whether this procedure is capable of detecting comparatively small amounts of short abortive transcripts among substantial numbers of processive transcripts, different amounts of short RNA standard were mixed with an amount of long RNA standard that may resemble the levels of full-length transcripts in clinical samples (500 copies per 250 ng of total RNA). Under these circumstances, an excess of short transcripts could be discriminated ($p = 0.0094$, unpaired *t*-test) with as few as 125 copies of short RNA standard in a background of 500 long standards, and quantification of TAR transcripts was linear ($r^2 = 0.98$) in a dilution series of 2000 down to 0 copies of short RNA standard (Fig. 3C). The results of this experiment suggest that short abortive transcripts should be detectable with our procedure when present at levels equal to or higher than elongated transcripts.

In order to compare the overall efficiency of the RT-ddPCR assays for total (TAR) and processive (R-U5/gag) transcripts, we next tested the efficiency of the TAR and R-U5/gag assays when applied to identical quantities of same virion standard under working assay conditions. Identical amounts of the virion standard (10^2 to 10^5 per RT) were added to the polyadenylation-RT reactions for TAR (with 1 μg cellular RNA) and the RT reactions (no polyadenylation) for R-U5/gag (with 5 μg cellular RNA). After correcting for the differences in the fraction of the RT reaction used for each ddPCR well (5 μl out of 30 μl for TAR assay, vs. 5 μl of 50 μl for R-U5/gag) and the fact that each full length virion has 2 TAR regions but only one R-U5/gag region (so the expected copies for the TAR reaction are 2× the input of virion standard), the overall efficiency of the TAR assay (61%) was similar ($p = 0.63$, Wilcoxon test) to that of the R-U5/gag assay (63%) (Fig. 4A).

Next, we compared the detection efficiency of short standards in the TAR assay to that of long standards in the R-U5/gag assay when both assays are performed under working conditions. The respective standards were added to the polyadenylation-RT reactions for TAR (with 1 μg cellular RNA) and the RT reactions (no polyadenylation) for R-U5/gag (with 5 μg cellular RNA) in amounts calculated to give ≤10–1000 copies per ddPCR well. The efficiency with which the TAR assay detected short standards (69%, Fig. 4B) was slightly lower than the efficiency with which the R-U5/gag assay detected long standards

(88%, Fig. 4C), although the difference did not reach statistical significance (p = 0.5, Wilcoxon test).

3.4. Clinical samples

Finally, the assays for total (TAR) and processive (R-U5/gag) transcripts were applied to total RNA extracted from PBMC from 8 HIV-infected individuals on ART with plasma viral loads below 40 copies/ml. To exclude the presence of contaminating HIV DNA, control reactions containing everything but RT (no RT controls) were performed for both assays; most were completely negative and the remainder gave only one positive droplet, often with fluorescence in both channels consistent with a false positive. After accounting for differences in the RNA input into the RT and the fraction used for ddPCR, HIV RNA copies were normalized to cell numbers using RNA mass, assuming that 1 μg RNA is the equivalent of 10^6 cells(Fischer et al., 1999). In all 8 individuals, HIV RNA molecules containing the TAR loop sequence were detected at levels at least one order of magnitude higher than elongated HIV transcripts encompassing the R-U5/gag region (p = 0.0078, Wilcoxon test; Fig. 5). The ratio of total to processive transcripts ranged from 12 to 165 (mean 46; median 24), implying a substantial block to elongation. A minimum estimate of the number of short transcripts can be calculated using the function (TAR loop RNA copies – 2 * R-U5/gag copies), which assumes that all elongated transcripts are full length and therefore contain 2 TAR regions. Assuming that short abortive HIV transcripts account for the observed excess of TAR loop RNA according to this function, short transcripts represent by far the most abundant HIV nucleic acid species (83–99% of total; median 91%) in patients on ART, implying that only 1–17% of initiated HIV transcripts are elongated beyond the 5′LTR.

4. Discussion

The method presented here allows for RT ddPCR-based quantification of TAR loop containing HIV RNAs in clinical samples with high efficiency and acceptable expenses in terms of reagent costs and hands-on time. Polyadenylation proved crucial to assure efficient reverse transcription of the TAR sequence independently of its RNA sequence context. Notably, the detection efficiency of the assay was validated using RNA templates encompassing the first 58nt of the HIV mRNA sequence, which corresponds to the proposed length of short abortive HIV transcripts as originally observed by Kao et al. (Kao et al., 1987) in RNAse protection assays, as well as using considerably longer RNA templates mimicking elongated HIV transcripts. Reverse transcription of our short RNA templates proved to be strikingly inefficient even using gene-specific RT priming, whereas the TAR loop could be efficiently reverse transcribed if placed on a longer RNA molecule. This suggested highly restricted access to primer binding sites within the TAR loop structure and implied that an intact stem loop can be efficiently overcome once RT is initiated from downstream sequences. Introduction of an *in vitro* polyadenylation step prior to RT increased detection efficiency of the short template dramatically by providing easily accessible RT priming sites downstream of the stem loop structure. This observation suggests that earlier assays using randomlyor specifically-primed RT PCR (Adams et al., 1994; Adams et al., 1999; Lassen et al., 2004; Lin et al., 2003; Natarajan et al., 2013) may

have underestimated levels of short abortive transcripts. In terms of assay design, these findings suggest that for amplification of highly structured RNA targets, first strand synthesis should be initiated downstream of the core structure with random or poly-dT primers or a gene-specific primer deliberately placed in an accessible region. If such a site is not available (for example, because the target is situated very close to the 3′end of the template RNA), the polyadenylation procedure described herein may provide an option to enhance detection efficiency.

One obvious drawback of the procedure is its multistep approach with repeated incubations. Since polyadenylation is indispensable for sufficient RT priming in very short TAR-containing RNAs, it proved to be crucial to separate the polyadenylation temporally from addition of the reverse transcription enzyme in order to create conditions for unbiased cDNA generation. This precluded the use of commercially-available kits providing polyadenylation and reverse transcription in a single step. However, optimization of buffer conditions allowed for simple addition of RT reagents after completion of the polyadenylation reaction without a need for intermediary product purification. Therefore, the procedure should be amenable to partial automatization in a liquid handling robot with built-in thermo element. A second disadvantage of the TAR assay is the comparatively low sample input. Again, due to the need for polyadenylation of every sample RNA molecule, poly-A polymerase activity has to be provided in excess in relation to the sample RNA mass and its availability is restricted by the formulation of the commercial enzyme solution. Empirically, an RNA input of 1 μg per polyadenylation/RT reaction proved to be optimal for minimizing inhibition in a sample mass that can be expected to contain measurable amounts of HIV RNA in samples from HIV infected individuals with suppressed plasma viral load. Finally, polyadenylation may reduce the efficiency of reverse transcription and possibly also enhance the frequency of molecular dropouts during ddPCR, conceivably due to excessive synthesis of poly-dT cDNA, which consumes substrates. In the TAR assay used in this study, inhibition by polyadenylation was of modest degree (>50% detection efficiency) and likely of little relevance for most potential applications. While it would have been ideal to test both the TAR and R-U5/gag assays on aliquots from the same polyadenylation-RT reaction, polyadenylation seemed to have a stronger inhibition on the R-U5/gag assay, which necessitated performing this assay without preceding polyadenylation. However, the polyadenylation-RT-ddPCR assay for TAR and the RT-ddPCR assay for R-U5/gag had essentially identical efficiencies when tested on the same virion standard, and even if the efficiency of the R-U5 gag assay on long standards was slightly higher than the efficiency of the TAR assay in detecting short standards, the effect would be to underestimate the true excess of short transcripts.

Applying this method to PBMC samples from HIV-infected individuals on ART, TAR loop containing RNAs were detected at levels one to two orders of magnitude higher than levels of HIV transcripts elongated to the end of the 5′LTR, suggesting that there is a substantial block to elongation of HIV transcription *in vivo* and that levels of HIV transcriptional initiation are considerably higher than previously recognized by measurement of more distal HIV RNA targets. Prior reports on short abortive transcripts and their relation to elongated transcripts are rare and mainly qualitative (Adams et al., 1994; Lassen et al., 2004; Lin et al., 2003). The only study providing quantitative data reported considerably lower levels of short

abortive transcripts in clinical samples (Adams et al., 1999). This may be due to the fact that the sample set used by Adams et al. included both patients on and off ART, whereas our participants all had suppressed viral loads. Furthermore, as shown herein, the random RT priming strategy used in that study most probably underestimates the abundance of short HIV transcripts.

The excess of short HIV transcripts in PBMCs from HIV+ patients on ART invites consideration of old and new questions regarding their origin and relevance for HIV pathogenesis. Whether these transcripts originate from aborted nascent HIV transcripts or whether they predominantly represent persisting remnants from degradation of longer transcripts remains at least partly unanswered due to lack of data on the kinetics of their decay in living cells (Karn, 1999). Although processing of HIV TAR RNAs into functional microRNAs by the RNAse III Dicer has been demonstrated (Ouellet et al., 2008), TAR's stable stem-loop structure might pose a bottleneck to nuclease digestion, which should result in even higher levels of short transcripts after a bout of HIV expression. However, with this caveat, short abortive HIV transcripts may be interpreted – as originally suggested (Kao et al., 1987) – as a measure of transcription initiation rates in the absence of sufficient factors supporting transcriptional elongation, and their levels relative to processive transcripts should act as indicators of specific blocks to HIV transcription, which might prove to be a valuable tool for studying *in vivo* mechanisms of HIV latency and effects of putative latency-reversing agents.

Acknowledgments

The authors thank the study participants and the staff at the Infectious Diseases Clinic of the San Francisco VA, including Mai Vu, Harry Lampiris, and Reginald Hillarie. The following reagent was obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: pNL4-3 from Dr. Malcolm Martin. This work was supported by the U.S. Department of Veterans Affairs (IK2 CX000520-01 [to S.Y.], I01 BX000192 [J.W.]), the National Institute of Diabetes and Digestive and Kidney Diseases (N.I.D.D.K.) at the National Institutes of Health (1R01 DK108349-01 [S.Y.]), the National Institute of Allergy and Infectious Diseases at the National Institutes of Health (R56 AI116342, R33 AI116218 [J.W.], U19 AI096109 [J.W.], P30 AI036211 [A.R.], and R21 AI116173 [A.R.]), the American Foundation for AIDS Research (amfAR) Institute for HIV Cure Research (109301 [S.Y., J.W.]), and the Swiss National Science Foundation (PBZHP3_147260 [P.K.]).

References

Adachi A, Gendelman HE, Koenig S, Folks T, Willey R, Rabson A, Martin MA. Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. J Virol. 1986; 59:284–291. [PubMed: 3016298]

Adams M, Sharmeen L, Kimpton J, Romeo JM, Garcia JV, Peterlin BM, Groudine M, Emerman M. Cellular latency in human immunodeficiency virus-infected individuals with high CD4 levels can be detected by the presence of promoter-proximal transcripts. Proc Natl Acad Sci U S A. 1994; 91:3862–3866. [PubMed: 8171003]

Adams M, Wong C, Wang D, Romeo J. Limitation of tat-associated transcriptional processivity in HIV-infected PBMC. Virology. 1999; 257:397–405. [PubMed: 10329550]

Anderson, EM., Gorelick, R., Hill, S., Rhem, C., Kearney, M., Mellors, JW., Coffin, JM., Piatak, M., Maldarelli, F. Investigating Mechanisms of HIV Persistence using Droplet Digital PCR Approaches. Seattle, USA. CROI; 2015.

Budhiraja S, Rice AP. Reactivation of latent HIV: do all roads go through P-TEFb? Future Virol. 2013; 8

Bustin SA, Nolan T. Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. J Biomol Tech. 2004; 15:155–166. [PubMed: 15331581]

Bustin S, Dhillon HS, Kirvell S, Greenwood C, Parker M, Shipley GL, Nolan T. Variability of the reverse transcription step: practical implications. Clin Chem. 2015; 61:202–212. [PubMed: 25361949]

Chun TW, Stuyver L, Mizell SB, Ehler LA, Mican JA, Baseler M, Lloyd AL, Nowak MA, Fauci AS. Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy. Proc Natl Acad Sci U S A. 1997; 94:13193–13197. [PubMed: 9371822]

Emiliani S, Fischle W, Ott M, Van Lint C, Amella CA, Verdin E. Mutations in the tat gene are responsible for human immunodeficiency virus type 1 postintegration latency in the U1 cell line. J Virol. 1998; 72:1666–1670. [PubMed: 9445075]

Finzi D, Hermankova M, Pierson T, Carruth LM, Buck C, Chaisson RE, Quinn TC, Chadwick K, Margolick J, Brookmeyer R, Gallant J, Markowitz M, Ho DD, Richman DD, Siliciano RF. Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. Science. 1997; 278:1295–1300. [PubMed: 9360927]

Fischer M, Huber W, Kallivroussis A, Ott P, Opravil M, Luthy R, Weber R, Cone RW. Highly sensitive methods for quantitation of human immunodeficiency virus type 1 RNA from plasma, cells, and tissues. J Clin Microbiol. 1999; 37:1260–1264. [PubMed: 10203467]

Fischer M, Wong JK, Russenberger D, Joos B, Opravil M, Hirschel B, Trkola A, Kuster H, Weber R, Günthard HF. Study ftSHC. Residual cell-associated unspliced HIV-1 RNA in peripheral blood of patients on potent antiretroviral therapy represents intracellular transcripts. Antiviral Ther. 2002; 7:91–1031.

Haddad F, Qin AX, Giger JM, Guo H, Baldwin KM. Potential pitfalls in the accuracy of analysis of natural sense-antisense RNA pairs by reverse transcription-PCR. BMC Biotechnol. 2007; 7:21. [PubMed: 17480233]

Kaiser P, Niederöst B, Joos B, von Wyl V, Opravil M, Weber R, Günthard HF, Fischer M. Equal amounts of intracellular and virion-enclosed hepatitis C virus RNA are associated with peripheral-blood mononuclear cells in vivo. J Infect Dis. 2006; 194:1713–1723. [PubMed: 17109344]

Kao SY, Calman AF, Luciw PA, Peterlin BM. Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product. Nature. 1987; 330:489–493.

Karn J. Tackling tat J Mol Biol. 1999; 293:235–254. [PubMed: 10550206]

Kumar AM, Borodowsky I, Fernandez B, Gonzalez L, Kumar M. Human immunodeficiency virus type 1 RNA levels in different regions of human brain: quantification using real-time reverse transcriptase-polymerase chain reaction. J Neurovirol. 2007; 13:210–224. [PubMed: 17613711]

Lassen KG, Bailey JR, Siliciano RF. Analysis of human immunodeficiency virus type 1 transcriptional elongation in resting CD4+ T cells in vivo. J Virol. 2004; 78:9105–9114. [PubMed: 15308706]

Lin X, Irwin D, Kanazawa S, Huang L, Romeo J, Yen TS, Peterlin BM. Transcriptional profiles of latent human immunodeficiency virus in infected individuals: effects of Tat on the host and reservoir. J Virol. 2003; 77:8227–8236. [PubMed: 12857891]

Malboeuf CM, Isaacs SJ, Tran NH, Kim B. Thermal effects on reverse transcription: improvement of accuracy and processivity in cDNA synthesis. Biotechniques. 2001; 30:1074–1078. 1080, 1082. passim. [PubMed: 11355343]

Natarajan M, Schiralli Lester GM, Lee C, Missra A, Wasserman GA, Steffen M, Gilmour DS, Henderson AJ. Negative elongation factor (NELF) coordinates RNA polymerase II pausing, premature termination, and chromatin remodeling to regulate HIV transcription. J Biol Chem. 2013; 288:25995–26003. [PubMed: 23884411]

Ouellet DL, Plante I, Landry P, Barat C, Janelle ME, Flamand L, Tremblay MJ, Provost P. Identification of functional microRNAs released through asymmetrical processing of HIV-1 TAR element. Nucleic Acids Res. 2008; 36:2353–2365. [PubMed: 18299284]

Sanders R, Mason DJ, Foy CA, Huggett JF. Evaluation of digital PCR for absolute RNA quantification. PLoS One. 2013; 8:e75296. [PubMed: 24073259]

Shi R, Chiang VL. Facile means for quantifying microRNA expression by real-time PCR. Biotechniques. 2005; 39:519–525. [PubMed: 16235564]

Sonza S, Mutimer HP, O'Brien K, Ellery P, Howard JL, Axelrod JH, Deacon NJ, Crowe SM, Purcell DF. Selectively reduced tat mRNA heralds the decline in productive human immunodeficiency

virus type 1 infection in monocyte-derived macrophages. J Virol. 2002; 76:12611–12621. [PubMed: 12438587]

Stahlberg A, Hakansson J, Xian X, Semb H, Kubista M. Properties of the reverse transcription reaction in mRNA quantification. Clin Chem. 2004a; 50:509–515. [PubMed: 14726469]

Stahlberg A, Kubista M, Pfaffl M. Comparison of reverse transcriptases in gene expression analysis. Clin Chem. 2004b; 50:1678–1680. [PubMed: 15331507]

Taube R, Peterlin M. Lost in transcription: molecular mechanisms that control HIV latency. Viruses. 2013; 5:902–927. [PubMed: 23518577]

Tuiskunen A, Leparc-Goffart I, Boubis L, Monteil V, Klingstrom J, Tolou HJ, Lundkvist A, Plumet S. Self-priming of reverse transcriptase impairs strand-specific detection of dengue virus RNA. J Gen Virol. 2010; 91:1019–1027. [PubMed: 19940062]

Wong JK, Hezareh M, Gunthard HF, Havlir DV, Ignacio CC, Spina CA, Richman DD. Recovery of replication-competent HIV despite prolonged suppression of plasma viremia. Science. 1997; 278:1291–1295. [PubMed: 9360926]

Yukl S, Pillai S, Li P, Chang K, Pasutti W, Ahlgren C, Havlir D, Strain M, Gunthard H, Richman D, Rice AP, Daar E, Little S, Wong JK. Latently-infected CD4+ T cells are enriched for HIV-1 Tat variants with impaired transactivation activity. Virology. 2009; 387:98–108. [PubMed: 19268337]

Yukl SA, Shergill AK, McQuaid K, Gianella S, Lampiris H, Hare CB, Pandori M, Sinclair E, Gunthard HF, Fischer M, Wong JK, Havlir DV. Effect of raltegravir-containing intensification on HIV burden and T-cell activation in multiple gut sites of HIV-positive adults on suppressive antiretroviral therapy. AIDS. 2010; 24:2451–2460. [PubMed: 20827162]

Yukl SA, Boritz E, Busch M, Bentsen C, Chun TW, Douek D, Eisele E, Haase A, Ho YC, Hütter G, Justement JS, Keating S, Lee TH, Li P, Murray D, Palmer S, Pilcher C, Pillai S, Price RW, Rothenberger M, Schacker T, Siliciano J, Siliciano R, Sinclair E, Strain M, Wong J, Richman D, Deeks SG. Challenges in detecting HIV persistence during potentially curative interventions: a study of the Berlin patient. PLoS Pathog. 2013; 9:e1003347. [PubMed: 23671416]

Yukl SA, Kaiser P, Kim P, Li P, Wong JK. Advantages of using the QIAshredder instead of restriction digestion to prepare DNA for droplet digital PCR. Biotechniques. 2014; 56:194–196. [PubMed: 24724845]

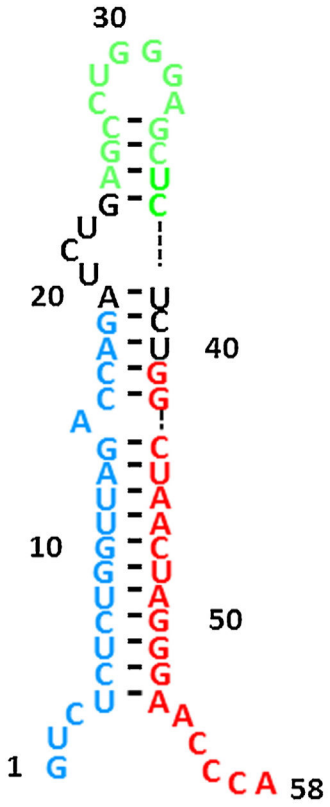

Fig. 1.

Structure of the HIV TAR stem loop within a 58nt short abortive transcript target as predicted by RNAStructure (by Dr. David H. Mathews, http://rna.urmc.rochester.edu). A free energy of −27.5 kcal/mole was predicted, corresponding to a melting temperature of approximately 83.0 °C under typical reaction conditions. Stem-loop base paring is indicated by bold lines. The sequences corresponding to the PCR forward primer, reverse primer, and fluorescent-labelled probe are highlighted in blue, red, and green, respectively. Numbering of bases begins with the HIV mRNA start site (HXB2 position 456). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

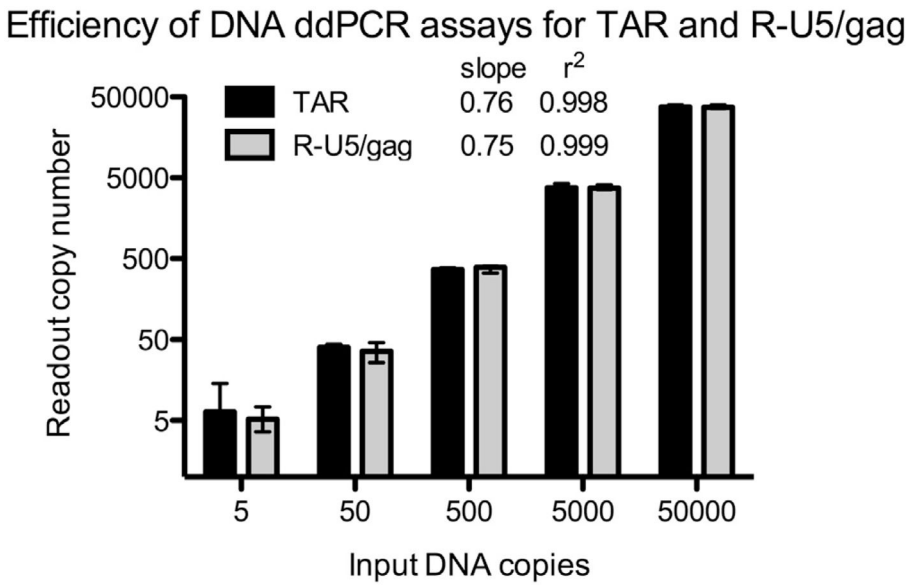

Fig. 2.

Detection efficiency and linearity of the ddPCR assays for DNA containing the TAR loop and R-U5/gag regions. Different amounts of a plasmid containing both target sequences were quantified on a background of 1 μg of total human DNA. Bars indicate median values and range of 4 replicates. Detection efficiency and linearity of each assay are expressed by the slope and r^2, as determined by linear regression, of the plot of readout copies versus calculated input of DNA standard per ddPCR well.

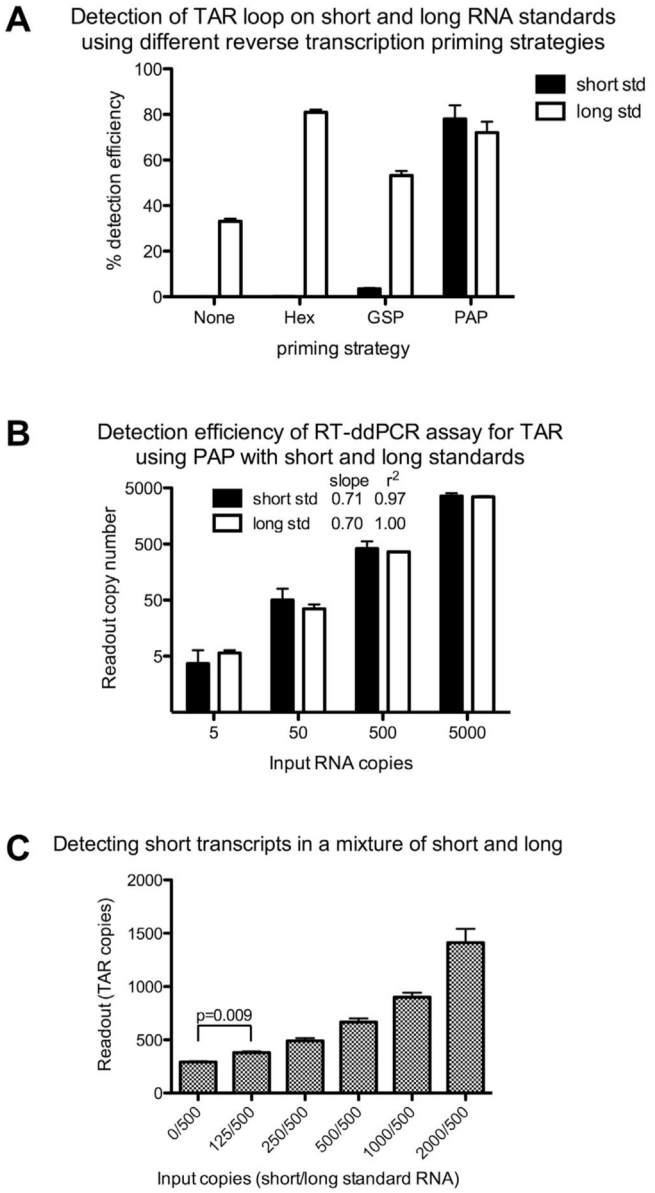

Fig. 3.

Detection of the TAR loop target on short and long HIV RNA standards. 3A: Efficiency of different priming strategies for reverse transcription. Short (58nt) HIV RNA standards (black columns) or long (2519nt) HIV RNA standards (white columns) were spiked into 1 μg of total human PBMC RNA and reverse-transcribed in the absence of any dedicated first strand primer (none) or with priming by random hexamers (HEX), a TAR-loop gene-specific reverse primer (GSP), or poly-dT15 and random hexamer primers after *in vitro* polyadenylation with poly-A polymerase (PAP). Subsequently, cDNA was quantified using ddPCR specific for the TAR loop. Efficiencies of detection are expressed as ratios of copies detected to the calculated input of 2500 copies per ddPCR well. 3B: Detection efficiency and linearity of the polyadenylation-RT-ddPCR assay for TAR. Different quantities of short HIV RNA standards (black columns) or long HIV RNA standards (white columns) were spiked

into 1 μg of total human PBMC RNA, polyadenylated, reverse transcribed, and assayed for the TAR region using ddPCR. Detection efficiency and linearity are expressed by the slope and r^2, as determined by linear regression, of the plot of measured cDNA copies versus calculated input of RNA standard per ddPCR well. 3C: Detecting short transcripts in a mixture of short and long targets. Different amounts of short standard (ranging from 0 to 2000 copies) were mixed with 500 copies of long standard in a background of 1 μg of total human PBMC RNA and measured using the polyadenylation-RT-ddPCR assay for TAR. Numbers on the x axis indicate the inputs of short standard and long standard RNA copy number equivalents per ddPCR reaction. The y axis indicates the total TAR copies measured by ddPCR. For all graphs, bars indicate median values and range of 4 replicates.

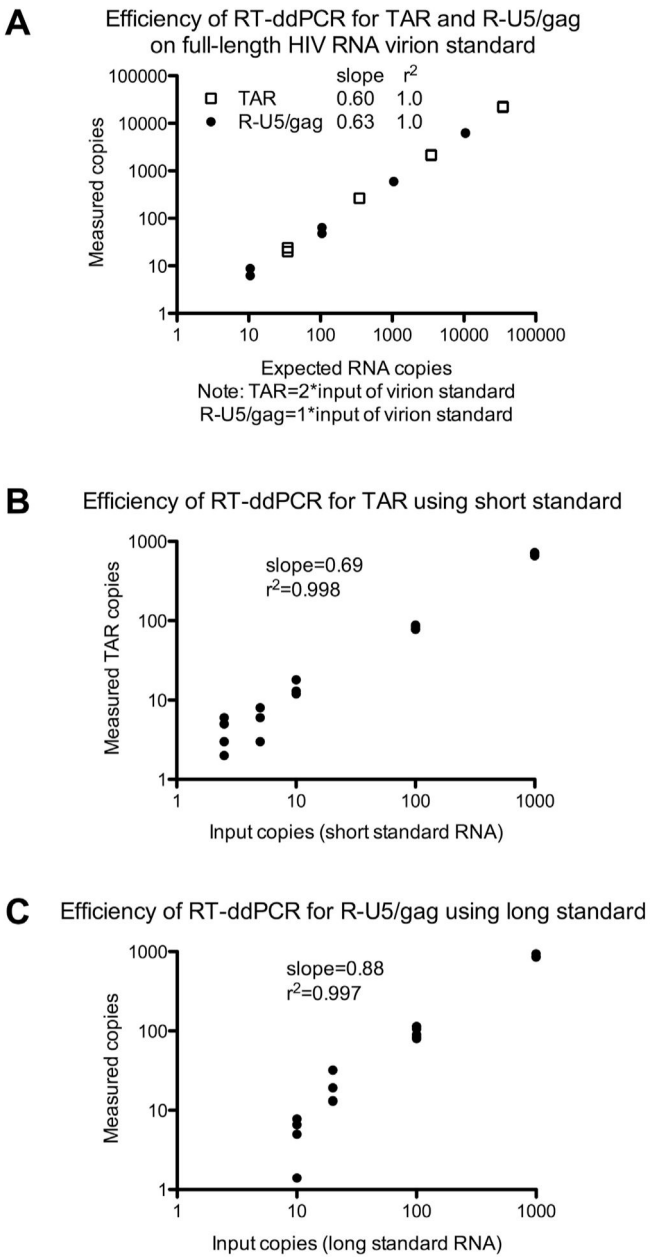

Fig. 4.

Comparing the efficiencies of the assays for total (TAR) and processive (R-U5/gag) transcripts. 4A: Efficiency of polyadenylation-RT-ddPCR assay for TAR and RT-ddPCR assay for R-U5/gag on full-length virion standard. Identical amounts of the virion standard (10^2 to 10^5 per RT) were added to the polyadenylation-RT reactions for TAR (with 1 μg cellular RNA) and the RT reactions (no polyadenylation) for R-U5/gag (with 5 μg cellular RNA). The x axis shows the expected copies per ddPCR well, as calculated from the input of virion standard into the RT reaction, the fraction of the RT reaction used per ddPCR well, and the fact that the virion standard contains two identical TAR regions (one each at the 5′ and 3′ LTR) but only one R-U5/gag region, so the TAR assay would be expected to give 2

copies per virion standard while the R-U5/gag assay should give only one copy per virion standard. The y axis shows the copies measured per ddPCR well (2 replicates) using the TAR assay (open squares) or R-U5/gag assay (black circles). 4B: Efficiency of polyadenylation-RT-ddPCR assay for TAR using the short HIV RNA standard. Short HIV RNA standard was added to the polyadenylation-RT reactions (with 1 μg cellular RNA) in amounts calculated to give 5, 10, 100, or 1000 copies per ddPCR well. The y axis shows the TAR copies measured per ddPCR well (3–4 replicates). 4C: Efficiency of RT-ddPCR assay for R-U5/gag using the long HIV RNA standard. The long HIV RNA standard was added to the RT reactions (with 5 μg cellular RNA) in amounts calculated to give 10, 20, 100, or 1000 copies per ddPCR well. The y axis shows the measured R-U5/gag copies per ddPCR well (minimum 4 replicates). For 4A–C, detection efficiency and linearity are expressed by the slope and r^2, as determined by linear regression, of the plot of measured versus expected copies per ddPCR well.

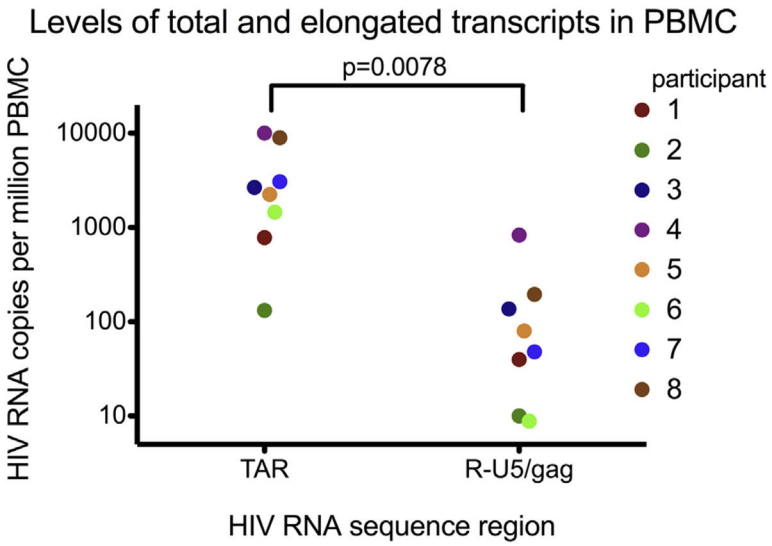

Fig. 5.

Levels of total (including short) and elongated HIV transcripts in PBMC of ART-treated HIV+ adults. Total RNA was extracted from PBMC of 8 HIV-infected individuals on ART with plasma viral loads <40 copies/ml using Trireagent. One μg of total RNA was used for the polyadenylation-RT-ddPCR assay for TAR (total HIV transcripts), and 5 μg was used for the RT-ddPCR assay for R-U5/gag (elongated HIV transcripts). After accounting for differences in the RNA input into the RT and the fraction used for ddPCR, HIV RNA copies were normalized to cell numbers using RNA mass, assuming that 1 μg RNA is the equivalent of 10^6 cells. Each data point represents the mean of two replicate ddPCR wells. In all 8 individuals, HIV RNA molecules containing the TAR loop sequence were detected at levels at least an order of magnitude higher than elongated HIV transcripts encompassing the R-U5/gag region.

Table 1

Nucleotide sequences of primers and dual-labelled probes used for RT-ddPCR.

Name	Purpose	Conservation Subtype B	HIV HXB2 positions	Sequence
TAR-F7	F primer TAR	81.1%	456 → 474	GTCTCTCTGGTTAGACCAG
TAR-R6	R primer TAR	75.6%	496 ← 513	TGGGTTCCCTAGYTAGCC
TAR-P3	Probe primer TAR	96.2%	480 → 492	FAM-AGCCTGGGAGCTC-MGB-NFQ
5′LTR-F	F primer R-U5/gag	90.0%	522 → 543	GCCTCAATAAAGCTTGCCTTGA
5′LTR-R	R primer R-U5/gag	83.9%	626 ← 643	GGGCGCCACTGCTAGAGA
5′LTR-P	Probe primer R-U5/gag	68.2%	559 ← 584	FAM-CCAGAGTCACACAACAGACGGGCACA-NFQ