Journal of Virological Methods 189 (2013) 53–57

Contents lists available at SciVerse ScienceDirect

Journal of Virological Methods

journal homepage: www.elsevier.com/locate/jviromet

Short communication

Quantitation of HIV DNA integration: Effects of differential integration site distributions on Alu-PCR assays

Troy Brady[a,1], Brendan J. Kelly[b,1], Frances Male[a], Shoshannah Roth[a], Aubrey Bailey[a], Nirav Malani[a], Rik Gijsbers[c], Una O'Doherty[a], Frederic D. Bushman[a,*]

[a] University of Pennsylvania Perelman School of Medicine, Department of Microbiology, 3610 Hamilton Walk, Philadelphia, PA 19104-6076, United States
[b] University of Pennsylvania Perelman School of Medicine, Department of Medicine, Infectious Disease Division, Philadelphia, PA, United States
[c] Division of Molecular Medicine, Katholieke Universiteit Leuven, Leuven, Flanders, Belgium

Article history:
Received 5 June 2012
Received in revised form 23 October 2012
Accepted 9 January 2013
Available online 21 January 2013

Keywords:
HIV
Integration
Alu
Quantitative PCR
LEDGF/p75
Taqman

A B S T R A C T

In many studies of HIV replication, it is useful to quantify the number of HIV proviruses in cells against a background of unintegrated forms of the HIV DNA. A popular method for doing so involves quantitative PCR using one primer complementary to the HIV long terminal repeat (LTR), and a second primer complementary to a cellular Alu repeat, so that PCR product only forms from templates where a provirus is integrated in the human genome near an Alu repeat. However, several recent studies have identified conditions that alter distributions of HIV integration sites relative to genes. Because Alu repeats are enriched in gene rich regions, this raises the question of whether altered integration site distributions might confound provirus abundance measurements using the Alu-PCR method. Here modified versions of the HIV tethering protein LEDGF/p75 were used to retarget HIV integration outside of transcription units, and show that this has a negligible effect on Alu-PCR quantitation of proviral abundance. Thus altered integration targeting, at least to the degree achieved here, is not a major concern when using the Alu-PCR assay.

The Alu-PCR method (Butler et al., 2001, 2002; O'Doherty et al., 2002) allows quantitation of HIV proviruses in cells despite the presence of unintegrated forms of HIV DNA (Butler et al., 2001; Farnet and Haseltine, 1991). Amplification is carried out using primers complementary to the HIV LTR and cellular Alu repeats, so that only proviruses integrated near Alu repeats in the human genome support amplification (Fig. 1). The progression of the PCR can be quantified using the Taqman PCR method or molecular beacons. For increased sensitivity, a two-step nested PCR method can be used (O'Doherty et al., 2002). Because each provirus in a cell resides a defined distance from the nearest Alu repeat, and because this distance varies for each provirus in the population, the control used for absolute quantitation must be chosen carefully. To match the heterogeneity in the analyate, standard DNA is prepared from heavily infected cells after long term culture, the long term culture being important to dilute out unintegrated DNA. The abundance of integrated proviruses can then be determined in purified genomic DNA from these cultures, and the heavily infected cell DNA diluted serially into uninfected cell DNA to make a standard curve. After analysis by Alu-PCR, comparison of the unknown to the standard allows estimation of the absolute number of proviruses in a genomic DNA sample (Butler et al., 2001; O'Doherty et al., 2002).

It is now clear that various treatments can alter HIV integration target site selection, raising the question of how this influences the Alu-PCR assay. HIV integration in human cells occurs most commonly in active transcription units (Schröder et al., 2002). HIV integration is promoted by binding of the cellular LEDGF/p75 protein (product of the PSIP1 gene) to HIV integrase protein (Cherepanov et al., 2003; Emiliani et al., 2005; Llano et al., 2006; Maertens et al., 2003; Turlure et al., 2004), and depletion of LEDGF/p75 reduces the targeting of integration to transcription units (Ciuffi et al., 2005; Marshall et al., 2007; Shun et al., 2007). LEDGF/p75 appears to act via a simple tethering mechanism, in which LEDGF/p75 binds to integrase with its C-terminal domain and to chromatin in transcription units with its N-terminal domains. Three reports have shown that re-engineering the LEDGF/p75 tether to contain a CBX/HP1β heterochromatin binding domain promotes integration outside of transcription units (Ferris et al., 2010; Gijsbers et al., 2010; Silvers et al., 2010). Further studies suggest that modulating the levels of additional cellular genes or cell growth status can also have detectable effects on integration frequency in transcription units (Barr et al., 2006; Ciuffi

Abbreviations: HIV, human immunodeficiency virus; LEDGF/p75, human lens epithelium-derived growth factor; LTR, long terminal repeat; PBS, primer-binding site; RT, reverse transcriptase.

* Corresponding author.
E-mail address: bushman@mail.med.upenn.edu (F.D. Bushman).

[1] Both the authors contributed equally.

0166-0934/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jviromet.2013.01.004

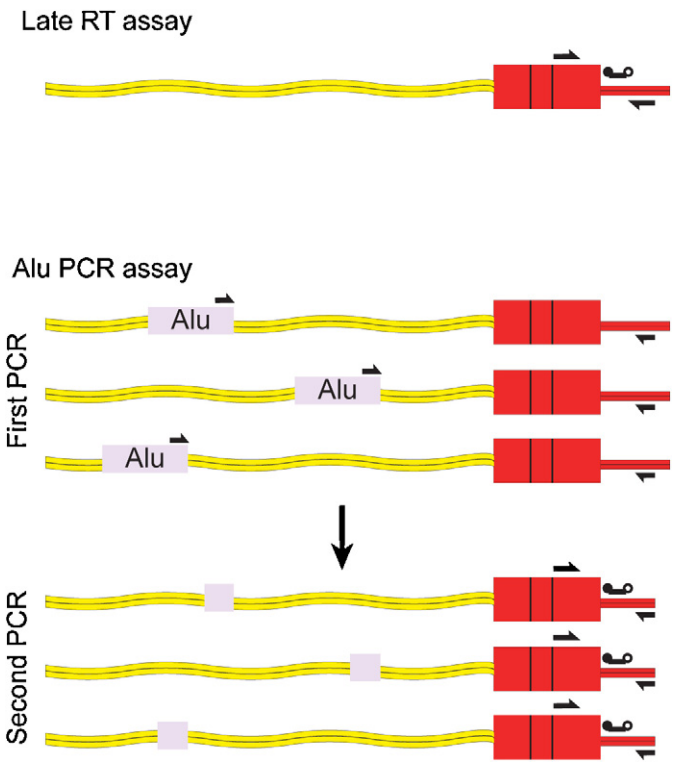

Fig. 1. Methods for quantifying proviral abundance using the Alu-PCR assay. Arrows indicate the positions of PCR primers, the line with two balls indicates the Taqman probe.

et al., 2006; Ocwieja et al., 2011; Schaller et al., 2011). Because Alu repeats are predominantly located in gene-rich regions (Lander, 2001; Venter, 2001), not randomly distributed in chromosomes, it is thus possible that factors altering integration targeting would confound Alu-PCR quantitation of proviral abundance.

In this study, infections were carried out in the presence of LEDGF/p75 knockdowns, CBX1-LEDGF fusions, or controls, to generate genomic DNA with different distributions of HIV integration sites (Gijsbers et al., 2010) and quantify the effects on the Alu-PCR assay. LEDGF/p75 was knocked down using optimized siRNAs, then genes encoding three altered tethering proteins introduced. These included the LEDGF/p75 C-terminal region fused to CBX1, a version of this chimera defective for integrase binding (the D366N mutation; Cherepanov et al., 2004; Llano et al., 2006), and an intact version of LEDGF/p75 expressed from a modified gene engineered to be insensitive to the siRNAs used for knock down. Cells were challenged with a GFP-bearing NL4-3 based HIV vector (Lu et al., 2004b), then vector supernatants were removed, and cells were allowed to grow for 14 days to dilute out unintegrated forms of the HIV DNA.

To verify altered integration targeting, genomic DNA was isolated from cells, and DNA containing integration sites amplified using ligation-mediated PCR. PCR products were then sequenced using the 454/Roche pyrosequencing method and human DNA flanking integrated proviruses mapped to the human genome (Ciuffi et al., 2009). The resulting integration site data sets are listed in supplementary Table 1.

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jviromet.2013.01.004.

To assess the differences in proviral distribution between samples, the relationship of integration sites to nearby genomic features was analyzed (Fig. 2) and found to parallel results from previous studies with these reagents (Gijsbers et al., 2010). Briefly, wild type HeLaP4-CCR5 cells infected with HIV-1 showed favored integration in transcription units, whereas reduction in LEDGF/p75 expression diminished the proportion of integration sites in transcription units (Fig. 2A). Back-complementing the depleted cells with an LEDGF/p75 allele insensitive to the siRNA restored integration in transcription units to WT levels. In the cells depleted for LEDGF/p75 but containing the CBX1-LEDGF fusion, integration in transcription units was no longer favored. In contrast, control cells encoding a version of the CBX1-LEDGF fusion containing an amino-acid substitution at the integrase binding site (D366N) showed no difference from the LEDGF/p75 knockdown cells.

Detailed comparisons of integration site distributions over genomic features and epigenetic marks (Fig. 2B and C) also showed patterns paralleling those reported previously (Gijsbers et al., 2010; Marshall et al., 2007; Shun et al., 2007). For associations with genomic features such as gene density, transcription units, CpG islands, and DNase sensitive sites, integration in CBX1-LEDGF cells appeared globally shifted toward a more random distribution compared to other conditions (Fig. 2B). For associations with epigenetic features (Fig. 2C), the CBX1-LEDGF sites showed many opposite associations compared to the other datasets. For example, all cell types except CBX1-LEDGF favored integration near marks associated with transcription, including H4K20me3 and H3K79me3. CBX1-LEDGF expressing cells, however, did not favor integration near these marks. Similarly, HIV integration near regions with H3K9me3, a mark associated with centromeric heterochromatin, transcriptionally silent regions and regions of CBX1 binding, was favored only in the presence of CBX1-LEDGF fusions and disfavored for all other conditions. Together, these changes confirm a significant shift in the HIV integration profile in CBX1-LEDGF cells compared to the other cell types.

Fig. 2D presents a summary of the mean number of Alu repeats in each cell line, as assessed in a range of genomic window sizes (100 bp to 10 kb) around integration sites in each data set. In a 1 kb window, the wild type control cells had a mean of 0.47 annotated Alu repeats surrounding each integration site. The mean number of Alu repeats in the LEDGF/p75 knockdown cells was not significantly different from wild type in the 1 kb window, but the cells knocked down for LEDGF/p75 and containing the CBX1-LEDGF fusion were significantly different (mean 0.29 Alu repeats per 1 kb window). Whether there were significant differences in the number of Alu repeats around integration sites was assessed using a Kruskal–Wallis test followed by pairwise Wilcoxon rank sum tests. The CBX1-LEDGF cell line differed significantly from wild type cells in all windows tested, but the LEDGF KD line differed significantly from wild type cells only in the 10 kb window (Bonferroni-corrected $p < 0.05$ for all tests). The differences in numbers of Alu repeats were consistently less than two-fold over all windows.

To investigate how differences in proximity of proviruses to Alu repeats affected the Alu-PCR assay in samples containing many proviruses, results were compared for two Taqman PCR assays that quantify the copy number of integrated proviruses per cell (Fig. 3). For Alu-independent quantitation, a one-step assay was used with primers and probe binding to internal sequences within the HIV vector DNA, specifically in the R-U5 and PBS regions of HIV (the "late RT" amplicon) (Butler et al., 2001). Results were compared to those with the two-step version of the Alu-PCR protocol (Agosto et al., 2007; O'Doherty et al., 2002). In the first step, PCR was carried out using primers binding to an Alu repeat and HIV *gag*. In the second step, the late RT Taqman amplicon was used to quantify the amount of Alu-PCR product formed. Importantly, the same procedure was applied to control DNA comprised of a dilution series of long-term infected cell DNA, so that the differential abundance due to the preceding Alu-PCR step could be quantified.