

RESEARCH

Open Access



Underestimated effect of intragenic HIV-1 DNA methylation on viral transcription in infected individuals

Sam Kint^{1,2} , Wim Trypsteen¹, Ward De Spiegelaeere³, Eva Malatinkova¹, Sabine Kinloch-de Loes⁴, Tim De Meyer², Wim Van Criekinge^{2†} and Linos Vandekerckhove^{1*†}

Abstract

Background: The HIV-1 proviral genome harbors multiple CpG islands (CpGs), both in the promoter and intragenic regions. DNA methylation in the promoter region has been shown to be heavily involved in HIV-1 latency regulation in cultured cells. However, its exact role in proviral transcriptional regulation in infected individuals is poorly understood or characterized. Moreover, methylation at intragenic CpGs has never been studied in depth.

Results: A large, well-characterized HIV-1 patient cohort ($n = 72$), consisting of 17 long-term non-progressors and 8 recent seroconverters (SRCV) without combination antiretroviral therapy (cART), 15 early cART-treated, and 32 late cART-treated patients, was analyzed using a next-generation bisulfite sequencing DNA methylation method. In general, we observed low level of promoter methylation and higher levels of intragenic methylation. Additionally, SRCV showed increased promoter methylation and decreased intragenic methylation compared with the other patient groups. This data indicates that increased intragenic methylation could be involved in proviral transcriptional regulation.

Conclusions: Contrasting in vitro studies, our results indicate that intragenic hypermethylation of HIV-1 proviral DNA is an underestimated factor in viral control in HIV-1-infected individuals, showing the importance of analyzing the complete proviral genome in future DNA methylation studies.

Keywords: HIV-1, DNA methylation, HIV-1 latency, Epigenetics, Next-generation sequencing, Bisulfite sequencing, Intragenic DNA methylation

Background

Current combination antiretroviral therapy (cART) can successfully control human immunodeficiency virus type 1 (HIV-1) infection and prevent disease progression to the acquired immunodeficiency syndrome (AIDS). However, a cure is not generally achievable due to the establishment of a latent reservoir of proviral HIV-1 DNA which remains dormant and fuels viral rebound upon treatment interruption [1–4]. Therefore, better insight into the mechanisms regulating HIV-1 latency is crucial

in order to interfere with this latency state and to develop cure strategies. The state of HIV-1 latency can be defined as the transcriptional silencing of proviral genes caused by multiple transcriptional blocks after the stable integration of proviral DNA into the host genome [5]. Some of the major silencing mechanisms consist of epigenetic modifications, which have led to several clinical trials investigating the latent viral reservoir reactivation with histone deacetylase inhibitors, albeit with limited success [6–10]. Other epigenetic modifications such as HIV-1 proviral DNA methylation have also been described in HIV-1 transcriptional silencing and have been explored as targets for HIV-1 latency reversing strategy [11–14].

DNA methylation is a well-described epigenetic modification in which a methyl group is added at the number five carbon of the cytosine pyrimidine ring in CpG

* Correspondence: linos.vandekerckhove@ugent.be

[†]Wim Van Criekinge and Linos Vandekerckhove contributed equally to this work.

¹HIV Cure Research Center, Department of Internal Medicine and Pediatrics, Faculty of Medicine and Health Sciences, Ghent University and Ghent University Hospital, Corneel Heymanslaan 10, Medical Research Building 2, 9000 Ghent, Belgium

Full list of author information is available at the end of the article



dinucleotides [15, 16]. This modification plays a role in genome transcription regulation and is crucial in processes such as the development of multicellular organisms, cell differentiation, regulation of gene expression, X-chromosome inactivation, genomic imprinting, and in the suppression of parasitic and other repeat sequences [15–23]. In general, reliable and stable transcriptional silencing is caused if CpG islands (CpGIs)—stretches of DNA that contain an increased frequency of CpG dinucleotides (CG content > 50% and observed/expected CpG ratio > 60%)—in promoter regions are hypermethylated [12, 15, 16, 24, 25]. Methylation of CpGIs within gene bodies (intragenic methylation) has been shown to be involved in regulation of intragenic promoters, alternative splicing, and cellular differentiation, but also in the activation of retroviruses, repetitive elements, and prevention of aberrant transcript production [26–30].

The HIV-1 genome encodes five CpGIs [12]: two are surrounding the promoter region and flanking the HIV-1 transcription start site and several transcription factor binding sites (e.g., TCF-1 α , NF- κ B, SP1) at the 5' long terminal repeat (LTR) region (CpGI LTR in the U3 region of the 5' LTR and CpGI non-coding region (NCR), downstream the HIV-1 5' LTR (Fig. 1)) [12]. Two other CpGIs are located in the *env* gene (CpGI ENV (35% conserved) and CpGI *env-tat-rev* (ETR)), surrounding the HIV-1 antisense open reading frame (Fig. 1) [12, 31]. The fifth CpGI is located in the 3' LTR, where the antisense transcription start site is located [12, 31]. In cultured HIV-1-infected cells, the regulatory role of proviral promoter methylation in viral transcriptional activity is clearly demonstrated: hypermethylation stabilizes HIV-1 latency and demethylating agents can induce activation of HIV-1 transcription [12, 13, 32–34]. However, studies performed on DNA methylation in infected individuals could not reproduce these findings indicating that this in vitro regulation does not apply in vivo [14, 32, 35–38].

To further understand the role of proviral HIV-1 DNA methylation in infected individuals, an NGS-based bisulfite

assay was developed to characterize HIV-1 proviral DNA methylation profiles of both promoter and intragenic regions in the context of a large, well-characterized patient cohort ($n = 72$). This cohort comprises four different patient groups as described by Malatinkova et al. [39]: 15 early cART-treated individuals (ET), 32 late cART-treated individuals (LT), 17 long-term non-progressors (LTNP), and 8 acute seroconverters (SRCV).

Methods

Patient cohorts and DNA samples

HIV-1-positive patients were recruited from two clinical centers, the Ian Charleson Day Centre (Royal Free Hospital, London, UK) and the AIDS Reference Center (Ghent University Hospital, Ghent, Belgium) during the study performed by Malatinkova et al. [39]. Seventy-two HIV-1-positive PBMC samples from that study were selected. Patients were divided into four cohorts based on their disease status (Additional Figure 1). The detailed study design and inclusion criteria have been described previously [39]. Briefly, (1) long-term cART-treated individuals (median treatment time of 10.77 years (interquartile range (IQR), 6.46–12.34 years)) who had initiated treatment during HIV-1 seroconversion (early treated (ET); $n = 15$) or (2) during the chronic phase of the infection (late treated (LT); $n = 32$); (3) cART-naïve long-term non-progressors (LTNPs, $n = 17$) who had maintained HIV-1 viral load (VL) ≤ 1000 copies/ml and CD4+ T cells > 500 cells/mm³ over > 7 years post-infection or (4) cART-naïve seroconverters (SRCV, $n = 8$), who were sampled during the acute phase of the infection. Baseline characteristics and clinical parameters of these cohorts are summarized in Table 1. The Ethical Committees of Ghent University Hospital and the Royal Free Hospital had approved this study (reference numbers: B670201317826 (Ghent) and 13/LO/0729 (London)) with all study subjects giving their written informed consent.

DNA from aliquots of 10^7 PBMCs was isolated using the DNeasy® Blood & Tissue Kit (Qiagen, The Netherlands, 69504). Sample DNA concentration was determined with

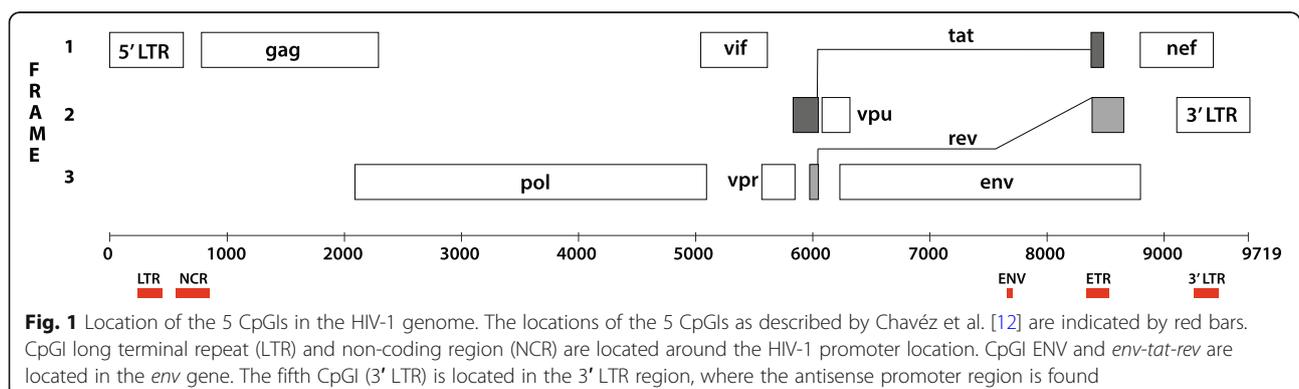


Table 1 Clinical characteristics and viral reservoir markers of the four patient cohorts

	Cohort 1 = ET	Cohort 2 = LTNP	Cohort 3 = LT	Cohort 4 = SRCV
# patients	15	17	32	8
Clinical characteristics				
Age (years)	45 (43–54.5)	49 (38–51)	48 (45–53.25)	37 (27–44.75)
Total cART (years)	11.65 (10.39–11.97)	0 (0–0)	9.80 (6.09–14.73)	0 (0–0)
Total VL suppression (years)	11.18 (9.82–11.37)	9.72 (0–14.67)	6.53 (5–10.42)	0 (0–0)
log VL zenith (copies/ml)	5.74 (5.31–5.88)	2.24 (1.79–2.76)	4.93 (4.24–5.52)	6.15 (5.14–6.31)
CD4 nadir (cells/ μ l)	413.5 (274.5–539.75)	624 (562–693)	154.5 (51.25–266.25)	483.5 (393.75–520.25)
CD4 at collection (cells/ μ l)	961 (737–1129.5)	793 (685–1010)	624.5 (484–885.5)	534 (393.75–617.50)
CD4/CD8	1.12 (0.8–1.47)	0.91 (0.82–1.47)	0.74 (0.6–0.93)	0.62 (0.37–0.87)
Viral reservoir markers				
Total HIV-1 DNA* (c/M PBMC) [40]	88.14 (46.19–124.02)	48.01 (20.16–56.50)	137.01 (56.08–219.20)	1290.48 (519.63–4428.60)
Integrated HIV-1 DNA* (c/M PBMC) [41, 42]	158.00 (122.70–388.55)	28.16 (0–158.41)	586.65 (315.12–918.15)	1802.68 (272.19–3966.55)
CA HIV-1 usRNA (c/M PBMC) [43]	0.79 (0.28–3.12)	0.44 (0.27–3.51)	6.12 (1.80–10.08)	15.47 (0.62–77.60)
2-LTR circles (c/M PBMC) [40]	1.48 (0–3.03)	0.77 (0.65–2.70)	1.32 (0.57–2.18)	15.35 (4.82–24.12)

Values are reported as median (interquartile range), SRCV seroconverters, LTNP long-term non-progressors, PBMCs peripheral blood mononuclear cells, CA cell-associated, usRNA unspliced RNA, cART combination antiretroviral therapy, VL viral load

*Total and integrated HIV-1 DNA measurements are performed using different assays and the absolute copies are therefore not directly comparable. To measure integrated HIV-1 DNA, an Alu-HIV-1 qPCR is used whereas digital PCR is used to determine the total number of HIV-1 DNA copies

the Qubit dsDNA BR (broad range) Assay Kit (Thermo Fisher Scientific, MA, USA, Q32850) on a Qubit 2.0 fluorometer according to the manufacturer's instructions.

Cell culture

Jurkat cells (human T cell leukemia line) and J-Lat 8.4 (Jurkat cells infected with one HIV-1 copy per cell [44]) were cultured in a humidified atmosphere of 37 °C and 5% CO₂ in RPMI 1640 medium with GlutaMAX™ Supplement (Thermo Fisher Scientific, MA, USA, 61870-010), supplemented with 10% FCS and 100 μ g/ml penicillin/streptomycin. The culture medium was renewed every 2 to 3 days. DNA was isolated as described in the previous section.

Primer design

Primers targeting the 4 major HIV-1 CpGIs were designed using 2 online available primer design tools (Methprimer [45] and bisulfite primer seeker (Zymo Research, CA, USA, <https://www.zymoresearch.com/pages/bisulfite-primer-seeker>)). LTR primers were obtained from Trejbalova et al. [13] and ETR₁ primers from Weber et al. [37]. To evaluate primers in silico, the bio-informatics tool developed by Rutsaert et al. [46], estimating the complementarity of each primer combination to all full-length HIV-1 sequences in the Los Alamos National Laboratory (LANL) database (www.hiv.lanl.gov) [47], was adapted: the database was transformed to the bisulfite-treated variant (C→T; CG→CG), nested primer combination analysis was included, as well as analysis of combinations of multiple PCR assays. First,

the in silico analysis was used to evaluate primer combinations that were obtained from literature as well as in-house designed. Primer combinations matching at least 50% of the LANL database and nested combinations with an overlap of at least 2/3 of the matched sequences were retained. Selected primers were in vitro tested using DNA from J-Lat 8.4 [44], diluted in Jurkat DNA at different concentrations to mimic patient samples (10, 000, 5000, 1000, 500, 250, 100 HIV-1 copies per 10⁶ cells). Finally, an additional in silico analysis was used to select 4 or less primer combinations per CpGI that targeted at least 60% of the LANL database. These final primer sequences are listed in Additional File 1.

Bisulfite treatment

A minimum of 5 \times 1 μ g of DNA per patient was bisulfite treated using the Epiect Bisulfite kit (Qiagen, The Netherlands, 59110), which is the least fragmenting commercial bisulfite kit available, according to a previous in-house comparison [18]. We used the standard protocol as provided by the manufacturer. The five aliquots per patient were pooled, and immediately stored at – 20 °C.

Bisulfite-specific PCR

All PCR reactions were performed in triplicate to reduce the probability of preferential amplification of one specific amplicon that would dominate the output. Nested PCR reactions were performed using the FastStart™ Taq DNA Polymerase, 5 U/ μ l (Roche Applied Science, Belgium, 12032953001). A volume containing theoretically at least

ten bisulfite-treated HIV-1 copies (based on the droplet digital PCR measurements as in Malatinkova et al. [39]) was added to the PCR mix containing $10 \times$ PCR buffer, 2.5 U polymerase, 400 nM forward and reverse primers, and 3% DMSO in a final volume of 25 μ l. Each CpGI was amplified with one nested primer combination, and after a failed PCR reaction, the subsequent primer combination was used (Additional File 1). Amplicons were visualized using 3% agarose gel electrophoresis. Depending on the selected primer, we used an in-house optimized PCR amplification protocol or one of the two previously published protocols [13, 37], as described in Additional File 1.

Sequencing

Bisulfite-treated amplicons were pooled equimolarly and libraries were prepared using the NEBNext UltraII DNA Library Prep Kit for Illumina (NEB, MA, USA, #E7645L/#E7103L). These libraries were sequenced on a MiSeq sequencing system (MiSeq[®] Reagent Kit v3 (600 cycle), MS-102-3003, Illumina). Sequencing reads were trimmed using Trimmomatic (version 0.38), quality controlled using FastQC (version 0.11.8), and subsequently mapped to an in-house developed HIV-1 consensus genome using the Bismark package (version 0.10.1) [48], providing a conversion efficiency estimation and methylation state of all analyzed CpGs.

Statistical analysis

HIV-1-specific amplicons with coverage > 250 were normalized and divided into tiles (blocks of the HIV-1 genome containing the region of interest (LTR or *env*)). Differential methylation analysis per region was performed using the MethylKit package (version 1.6.3) in R (version 3.5.1) [49, 50], including correction for overdispersion. *P* value calculation was performed using the Chi-square test and *p* value correction for multiple testing was performed within each comparison using false discovery rate (FDR) [51, 52].

Spearman rank correlation analysis was performed to explore correlations between DNA methylation (LTR and *env*) and patient characteristics (HIV-1 reservoir and immunological parameters, obtained from Malatinkova et al. [39]). Therefore, methylation data of both regions of every individual was summarized by calculating an *M* value over all CpGs using the formula as described by Du et al. [53]. Using stepwise regression model selection, linear regression models were developed for LTR and *env* methylation densities to determine which independent variables may explain variable DNA methylation in both regions.

Visualization was performed using R (version 3.5.1) with the following packages: PMCMR (version 4.3), Hmisc (version 4.2-0), graphics (version 3.5.1), ggplot2 (version 3.1.0), and corrplot (version 0.84) [50].

Results

In silico, in vitro, and in vivo HIV-1 DNA methylation assay development

Three hundred thirty-eight different nested primer combinations (assays) (13 LTR, 303 NCR, 1 ENV, and 21 ETR) were subjected to an in silico analysis using an adapted version of the bioinformatics tool developed by Rutsaert et al. [46] to estimate the complementarity to the Los Alamos National Library database, resulting in 70 nested PCR assays (2 LTR, 46 NCR, 1 ENV, and 21 ETR, Fig. 2a). The performance of these assays was subsequently tested by PCR amplification in undiluted and diluted J-Lat 8.4 DNA (up to 100 infected cells/10⁶ cells), resulting in 36 assays (2 LTR, 15 NCR, 1 ENV, and 18 ETR) that were capable of generating PCR products at the lowest dilutions (Fig. 2a). After a final in silico analysis, a set of 9 primer combinations (2 LTR, 3 NCR, 1 ENV, and 3 ETR; Fig. 2 and Additional File 1) was selected.

These nine assays were used to determine the HIV-1 methylation profile of HIV-1-positive blood samples. The percentage of patients for whom the primer combinations generated PCR amplicons is listed in Table 2. This data demonstrates a similar trend as expected based on the in silico analysis, being that a certain percentage of HIV-1 sequences would not be detected in patients for certain primer combinations due to HIV-1 sequence variation. The difference between expected amplification percentage and the actual amplification percentage was 7.85%, 1.57%, 10.58%, and 3.57% for LTR, NCR, ENV, and ETR, respectively (Table 2).

SRCV shows increased LTR methylation and decreased *env* methylation

In all four patient cohorts together, average methylation of all CpGs within the LTR region was 2.94% (IQR, 0.19–5.5%). When comparing patient cohorts, we observed significantly higher LTR methylation in SRCV as compared with all the other cohorts (ET, LT, and LTNP) ($\Delta = 6.48\%$; $q = 0.00029$, $\Delta = 4.15\%$; $q = 0.015$, and $\Delta = 5.94\%$; $q = 0.0044$, respectively) (Fig. 3a).

Higher CpG methylation was observed in the *env* region as compared with LTR, averaging 28.86% (IQR, 8.73–39.44%). All cohorts (ET, LT, and LTNP) showed a significantly higher methylation density compared with SRCV ($\Delta = 33.47\%$; $q = 0.013$, $\Delta = 35.32\%$; $q = 0.00017$, and $\Delta = 35.26\%$; $q = 0.028$, respectively) (Fig. 3b).

Correlations between HIV-1 methylation status and reservoir markers

During the explorative correlation analysis, negative correlations were found between the DNA methylation density in the LTR region and the duration of viral suppression ($\rho = -0.34$; $p = 0.020$) and CD4+ T cell count

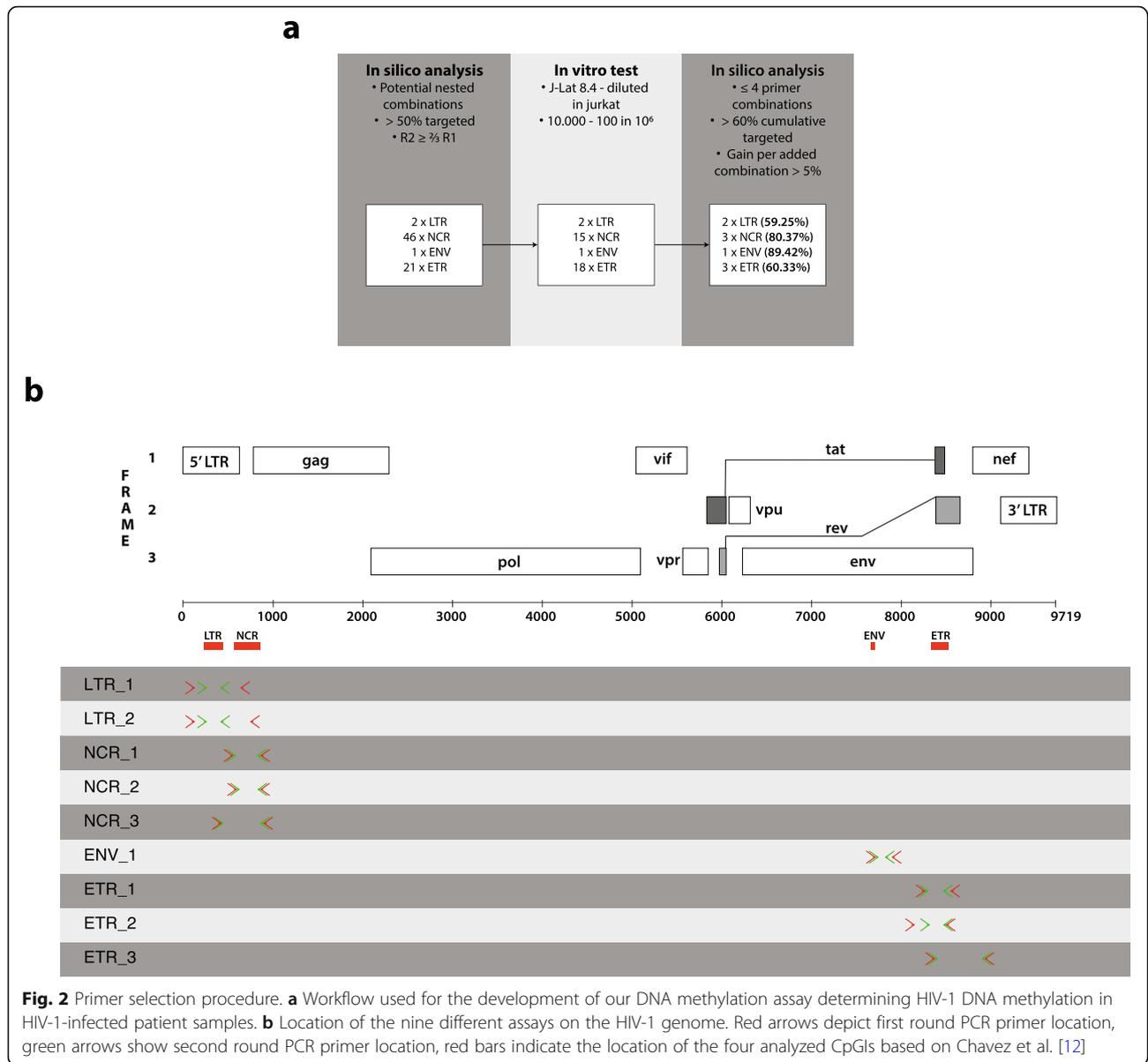


Fig. 2 Primer selection procedure. **a** Workflow used for the development of our DNA methylation assay determining HIV-1 DNA methylation in HIV-1-infected patient samples. **b** Location of the nine different assays on the HIV-1 genome. Red arrows depict first round PCR primer location, green arrows show second round PCR primer location, red bars indicate the location of the four analyzed CpGs based on Chavez et al. [12]

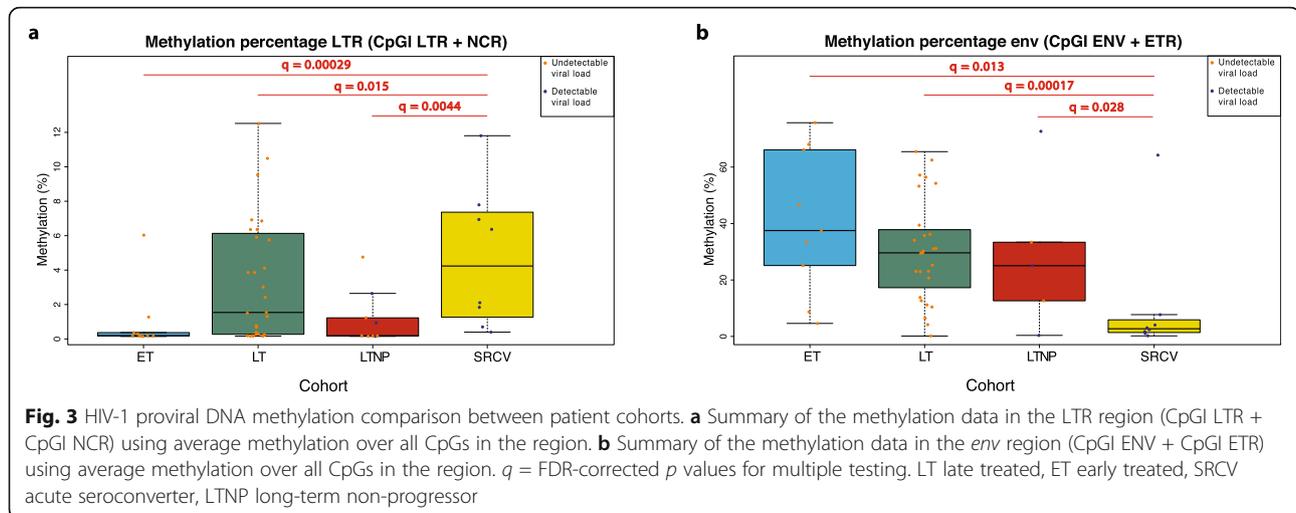
at time of collection ($\rho = -0.27$; $p = 0.043$) (Fig. 4a). However, we observed a significantly positive association for DNA methylation in the *env* region and the CD4 T cell count ($\rho = 0.40$; $p = 0.0045$) and cART duration ($\rho = 0.39$; $p = 0.0055$) (Fig. 4a). Moreover, *env* methylation decreased

with increasing VL levels ($\rho = -0.39$; $p = 0.0063$) and higher CD4+ T cell nadir ($\rho = -0.33$; $p = 0.020$) (Fig. 4a). Based on the linear regression models, the only variable that was independently associated with DNA methylation in the LTR was the duration of VL suppression. Three

Table 2 Performance of the nine final assays compared with the predicted performance using in silico analysis of the primer complementarity

	LTR	NCR	ENV	ETR
# primer combinations	2	3	1	3
% of patients expected to generate amplicons based on in silico analysis*	59.25	80.37	89.42	60.33
% of patients in which DNA was amplified	51.40	81.94	100	63.90
% of patients for which the DNA had sufficient quality to be mapped to HIV-1	48.61	75.00	41.67	63.90

*In silico analysis is based on the bioinformatics primer evaluation tool as described by Rutsaert et al. [46]



variables were independently associated with the *env* methylation: VL, CD4 nadir, and CD4 count at time of sampling (Fig. 4b).

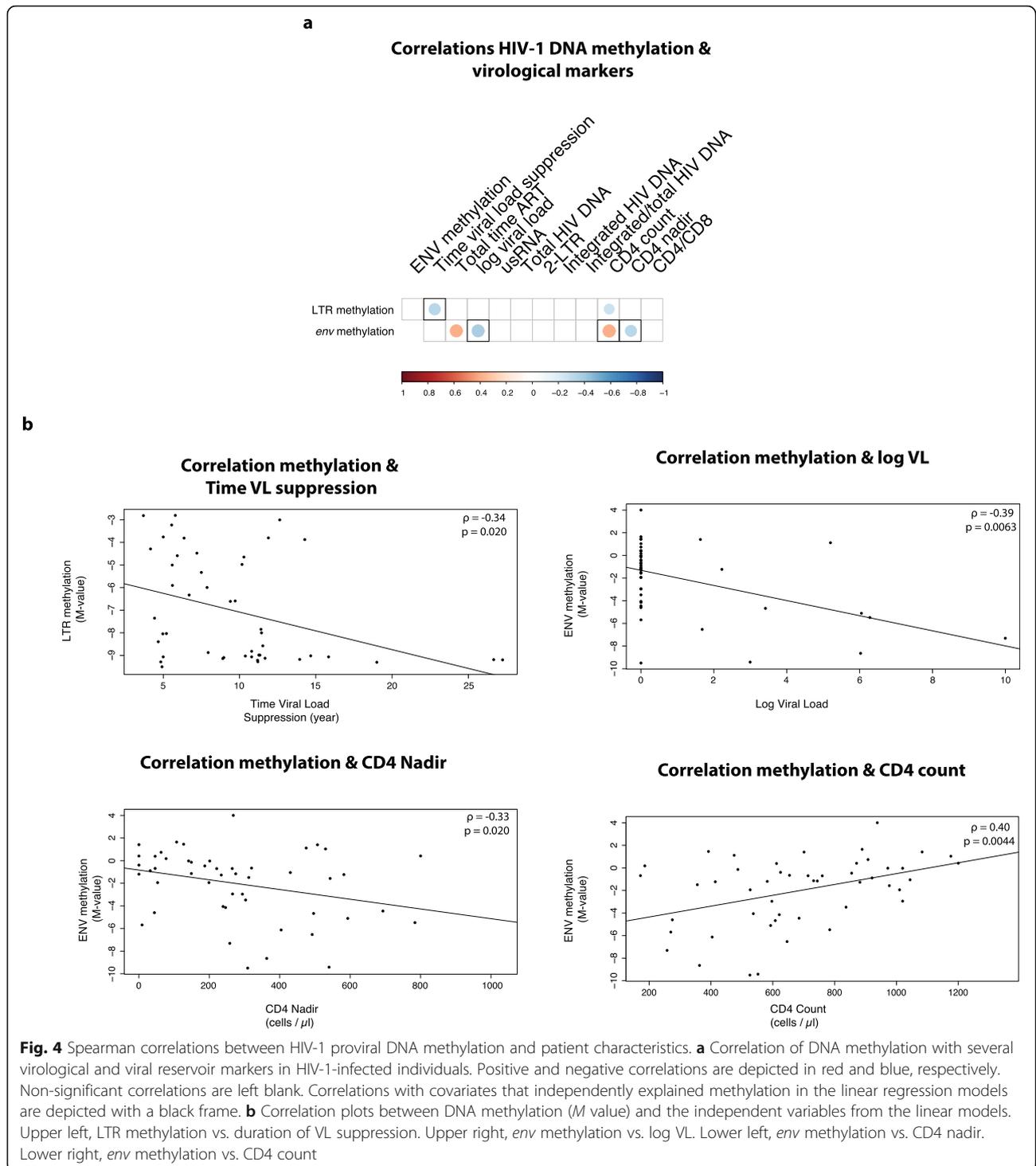
Discussion

The lack of consensus about the role of proviral DNA methylation in HIV-1 transcriptional regulation illustrates the need for a reliable and widely applicable methylation assessment method. In this study, we first described an in silico procedure to accurately predict the complementarity of PCR assays to the HIV LANL database, and an in vitro validation protocol to test the sensitivity of the designed assays. This procedure resulted in nine functional DNA methylation assays, designed against the four most common CpGIs of the HIV-1 provirus, which were consequently used to characterize HIV-1 DNA methylation in a large, well-characterized patient cohort. The in silico analysis was predictive of the number of patient samples leading to successfully amplified PCR products (Table 2), indicating that this is an effective approach to prioritize testing of primer sets in the context of HIV-1 or other pathogens with a high sequence variability. In addition, as shown in the study of Cortés-Rubio et al. [14], by using an NGS-based approach, our method fulfills the need to analyze a large number of proviruses for each patient when compared with the established Sanger sequencing-based methods [54].

Across our four patient cohorts, we have found that the HIV-1 provirus had low amounts of DNA methylation in the promoter region (average 2.94%, IQR 0.19–5.5%) but substantially higher levels of intragenic (*env*) methylation (average 28.86%, IQR 8.73–39.44%). When comparing the differential methylation between the cohorts, only SRCV showed distinct methylation profiles, with increased LTR, and decreased *env* methylation.

Similarly, if patients were divided based on their VL status (detectable VL (VL > 40 HIV-1 copies/ml plasma), comprising all SRCV and 6/17 LTNPs. vs. undetectable VL (VL < 40 HIV-1 copies/ml plasma), comprising ET, LT, and 11/17 LTNPs), individuals with a detectable VL had higher DNA methylation density in the HIV-1 LTR region and a lower density in the *env* region compared with those with an undetectable VL. These observations might indicate that specific methylation profiles may be associated with in vivo HIV-1 transcriptional control and latency maintenance.

Indeed, since the involvement of DNA methylation in HIV-1 latency was first described in 1987 [55], it has been confirmed in HIV-1-infected cultured cells and latency models that promoter methylation density is associated with silencing stability: DNA methylation induction can initiate/stabilize HIV-1 latency, while methylation inhibitors as 5-aza-2'-deoxycytidine (5-aza-CdR) cause HIV-1 reactivation and display clear synergistic effects with other latency reversing agents [11–13, 32–34, 36, 56–58]. These studies reported a major role of promoter DNA methylation in latency regulation, which was in line with the general concept of transcription regulation by DNA methylation: hypermethylation of the promoter region suppresses both basal promoter activity and responses to activating stimuli, and hypomethylation is a transcription mark [57]. However, DNA methylation studies on patient-derived samples have shown—with the exception of some LTNPs—the same trend as in our present observation: low level of DNA methylation in the promoter region, even in patients suppressing VL successfully, therefore not following the predictions from the in vitro experiments [37, 38]. It has been shown that DNA methylation behavior in cell lines is often drastically different from that of in vivo cells due to completely different epigenetic environments and



immortalization, sometimes producing unreliable results in terms of predicting in vivo DNA methylation events [59, 60]. Some studies, however do show increasing LTR DNA methylation over time [13], or dynamic profiles in patients when measured longitudinally [14]. We could not confirm these data since we only measured single time point samples of

patients with similar treatment time/time of virological control (except for the SRCV). The low abundance of DNA methylation in the promoter region of HIV-1 indicates that other (epigenetic) factors as integration site epigenetics or cell type might be more important for transcriptional regulation than promoter methylation.

In previous DNA methylation studies in HIV-1 patients, the focus was on promoter methylation assessment [13, 14, 32, 36–38]. In contrast to promoter methylation, the role of intragenic DNA methylation in general transcriptional regulation is less clearly described [26–30]. Studies outside of the HIV-1 field have suggested that intragenic methylation could have a role in the activation of retroviruses, repetitive elements, alternative splicing, transcription initiation in canonical promoters of embryonic stem cells, and prevention of aberrant transcript production [28–30]. Moreover, intragenic methylation has been shown to be a robust predictor of gene transcription in genes with a CpGI containing promoter [61]. In our study, decreased *env* methylation levels in individuals with active ongoing replication (SRCVs) suggests that intragenic methylation increases in the case of proviral transcriptional silencing, leading to higher methylation in latently infected cells or in those in which viral replication is blocked. Indeed, cART-treated patients and LTNP have lower viral transcription (measured as cell-associated unspliced RNA (CA usRNA)) than SRCV (Table 1) and *env* methylation shows an inverse correlation with CA usRNA within the SRCV cohort ($\rho = -0.81$; $p = 0.014$). Furthermore, intragenic methylation did correlate positively with the CD4+ T cell count, linking high intragenic methylation with viral control. Intragenic methylation was also negatively associated with the VL, a measure that indicates ongoing replication.

In contrast to what was proposed by LaMere et al. [54], we have found no statistical difference between proviral methylation in LTNP with undetectable VL (latent infection) and treated patients (cART-induced suppression) (LTR: $\Delta = 0.85\%$, $q = 0.74$; *env*: $\Delta = 2.29\%$, $q = 0.94$). This could be due to the low number of LTNPs with undetectable VL.

In general, the lack of promoter DNA methylation in HIV-1 proviral genomes in vivo suggests that this modification is of subordinate importance in the regulation of the viral life cycle compared with the more abundant, yet less studied intragenic DNA methylation. Our observations indicate that intragenic DNA methylation could be a late event during infection. Methylation of the proviral genome may occur stochastically during years of viral control, yet act as a stable epigenetic mark once established. This may subsequently affect transcription, including splicing, of viral transcripts, which could affect viral replication by interaction with transcriptional elongation (*tat*) or export of viral RNA (*rev*). Nevertheless, additional in vitro and in vivo experiments targeting the (intragenic) DNA methylation are required to evaluate the exact impact on the HIV-1 life cycle. Especially temporal changes of intragenic methylation would be very informative, yet our study was limited by the lack of

longitudinal sampling. Other limitations include the fact that although the cohort size was much larger than previous studies [13, 14, 32, 36–38], the patient groups described here were not balanced, not in size, nor for sex, and age. Additionally, we did no specific CD4+ T cells selection. The use of PBMCs could potentially mask differential methylation since it is shown that LRAs have cell-type specific effects, indicating cell-type specific epigenetic profiles [62]. Moreover, due to the targeted nature of the methodology, it does not allow to provide information about integration site methylation or replication competence of the analyzed provirus. Finally, we did not provide information about the fifth CpGI (3' LTR), nor did we analyze non-CpGI CpGs.

Conclusions

Altogether, our study illustrates the underestimation of the role of intragenic proviral DNA methylation in patient samples. Previous studies have mainly focused on LTR methylation and have interpreted LTR methylation as a transcriptional regulatory factor, ignoring any potential role of *env* methylation [13, 35, 38]. We suggest that both *env* and LTR methylation are involved in HIV-1 transcription regulation and that *env* methylation could be an important predictor of viral transcription in vivo. However, we also suggest that proviral promoter methylation is hindered/inhibited in all HIV-1-positive patients, especially those on cART, but that its density still influences viral transcription rate.

The exact functions of DNA methylation of these two regions should be clarified by performing additional experiments using longitudinal follow-up studies to monitor proviral DNA methylation dynamics within patients, starting early during infection, and ideally continuing over a period of multiple years of cART. Different CD4+ T cell types should be analyzed separately to avoid cell-type dependent bias of the data. If HIV-1-positive patients were to undergo treatment interruption, DNA methylation profiles should also be monitored in order to understand the methylation dynamics during viral rebound. Moreover, proviral intragenic non-CpGI methylation analysis could also provide a better understanding of HIV-1 latency regulation by DNA methylation. Here, we do provide a useful tool to help design and estimate the sample size needed in these studies. Altogether, these insights should be of paramount importance when looking at the various strategies to control HIV-1 after discontinuation of cART and for the HIV-1 cure field.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s13148-020-00829-1>.

Additional file 1. Primers and PCR experiments.

Additional file 2: Figure 1. Overview of patient cohorts included in this study. Patients are divided into four groups based on their disease state: early treated, late treated, Long-Term Non-Progressor and acute seroconverter. Arrows depict moment of sampling. PHI = Primary HIV-1 Infection; cART = combination Anti-Retroviral Therapy.

Abbreviations

5-aza-CdR: 5-aza-2'-deoxycytidine; AIDS: Acquired immunodeficiency syndrome; CA usRNA: Cell-associated unspliced RNA; cART: Combination antiretroviral therapy; CpGs: CpG islands; ET: Early cART-treated individuals; ETR: *env-tat-rev*; HIV-1: Human immunodeficiency virus type 1; IQR: Interquartile range; LANL: Los Alamos National Laboratory; LT: Late cART-treated individuals; LTNP: Long-term non-progressors; LTR: Long terminal repeat; NCR: Non-coding region; NGS: Next-generation sequencing; PBMC: Peripheral blood mononuclear cells; qPCR: Quantitative real-time PCR; VL: Viral load

Acknowledgements

The following reagent was obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, and NIH, J-Lat Full Length Clone (clone 8.4) from Dr. Eric Verdin. This study was in part performed with the support of The Foundation for AIDS Research (AmFAR), USA (grant ID 108314-51-RGR), HIV-ERA (130442 SBO, EURECA), Research Foundation Flanders SBO project SAPHIR (S000319N) and by The National Institute of Health (NIH), USA (grant R01 AI134419). Linos Vandekerckhove was supported by the Research Foundation Flanders (FWO), Belgium and he received a fundamental clinical mandate (1.8.020.09.N.00). Eva Malatinkova was funded by the Agency for Innovation by Science and Technology of the Flemish Government, Belgium (IWT, grant ID 111286) and is currently funded by Special Research Fund Ghent University, Belgium (BOF, grant ID 01P06816). The study did not have pharmaceutical firm sponsorship.

Authors' contributions

SK designed the study, performed experiments, analyzed, interpreted and visualized data, and drafted the manuscript. SKdeL and LV recruited patients. EM, SKdeL, and WDS contributed to the data collection. WDS contributed to the study design. WT contributed to the data interpretation and manuscript writing. TDM contributed to the statistical data analysis. LV and WWC designed the study, supervised experiments, and contributed to the manuscript writing. All co-authors revised and edited the manuscript. The author(s) read and approved the final manuscript".

Funding

This study was supported by AmFAR, HIV-ERA, NIH, FWO, IWT, and BOF. Funders had no role in the study design, data collection, analysis and interpretation, the writing of the report, or the decision to submit the work for publication.

Availability of data and materials

Datasets used during the current study are available from the corresponding author on request.

Ethics approval and consent to participate

Patient written informed consent was obtained from all the study participants. The study was approved by the Ethical Committee of Ghent University Hospital (reference number B670201317826) and Royal Free Hospital (reference number 13/LO/0729).

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Author details

¹HIV Cure Research Center, Department of Internal Medicine and Pediatrics, Faculty of Medicine and Health Sciences, Ghent University and Ghent University Hospital, Corneel Heymanslaan 10, Medical Research Building 2, 9000 Ghent, Belgium. ²Biobix, Department of Data Analysis and Mathematical Modelling, Faculty of Bio-science Engineering, Ghent University, Coupure Links 653, 9000 Ghent, Belgium. ³Department of Morphology, Faculty of Veterinary Medicine, Ghent University, Salisburylaan 133, 9820 Merelbeke,

Belgium. ⁴Division of Infection and Immunity, Royal Free Hospital, Royal Free Campus, University College London, Pont St, Hampstead, London NW3 2QG, UK.

Received: 27 December 2019 Accepted: 16 February 2020

Published online: 28 February 2020

References

- Barré-Sinoussi F, Ross AL, Delfraissy J-F. Past, present and future: 30 years of HIV research. *Nat Rev Microbiol*. 2013;11:877–83. <https://doi.org/10.1038/nrmicro3132>.
- Archin NM, Sung JM, Garrido C, Soriano-Sarabia N, Margolis DM. Eradicating HIV-1 infection: seeking to clear a persistent pathogen. *Nat Rev Microbiol*. 2014;12:750–64. <https://doi.org/10.1038/nrmicro3352>.
- Buzón MJ, Massanella M, Llibre JM, Esteve A, Dahl V, Puertas MC, et al. HIV-1 replication and immune dynamics are affected by raltegravir intensification of HAART-suppressed subjects. *Nat Med*. 2010;16:460–5. <https://doi.org/10.1038/nm.2111>.
- De Scheerder M-A, Vrancken B, Dellicour S, Schlub T, Lee E, Shao W, et al. HIV rebound is predominantly fueled by genetically identical viral expansions from diverse reservoirs. *Cell Host Microbe*. 2019. <https://doi.org/10.1016/j.chom.2019.08.003>.
- Yukl SA, Kaiser P, Kim P, Telwate S, Joshi SK, Vu M, et al. HIV latency in isolated patient CD4+ T cells may be due to blocks in HIV transcriptional elongation, completion, and splicing. *Sci Transl Med*. 2018;10:1–16. <https://doi.org/10.1126/scitranslmed.aap9927>.
- Abner E, Jordan A. HIV "shock and kill" therapy: in need of revision. *Antiviral Res*. 2019;166:19–34. <https://doi.org/10.1016/j.antiviral.2019.03.008>.
- Archin NM, Liberty AL, Kashuba AD, Choudhary SK, Kuruc JD, Crooks AM, et al. Administration of vorinostat disrupts HIV-1 latency in patients on antiretroviral therapy. *Nature*. 2012;487:482–5. <https://doi.org/10.1038/nature11286>.
- Rasmussen TA, Tolstrup M, Brinkmann CR, Olesen R, Erikstrup C, Solomon A, et al. Panobinostat, a histone deacetylase inhibitor, for latent virus reactivation in HIV-infected patients on suppressive antiretroviral therapy: a phase 1/2, single group, clinical trial. *Lancet HIV*. 2014;1:e13–21. [https://doi.org/10.1016/S2352-3018\(14\)70014-1](https://doi.org/10.1016/S2352-3018(14)70014-1).
- Elliott JH, Wightman F, Solomon A, Ghneim K, Ahlers J, Cameron MJ, et al. Activation of HIV transcription with short-course vorinostat in HIV-infected patients on suppressive antiretroviral therapy. *PLoS Pathog*. 2014;10:e1004473. <https://doi.org/10.1371/journal.ppat.1004473>.
- Søgaard OS, Graversen ME, Leth S, Olesen R, Brinkmann CR, Nissen SK, et al. The depsipeptide romidepsin reverses HIV-1 latency in vivo. *PLOS Pathog*. 2015;11:e1005142. <https://doi.org/10.1371/journal.ppat.1005142>.
- Gutekunst KA, Kashanchi F, Brady JN, Bednarik DP. Transcription of the HIV-1 LTR is regulated by the density of DNA CpG methylation. *J Acquir Immune Defic Syndr*. 1993;6:541–9.
- Chávez L, Kauder S, Verdin E. In vivo, in vitro, and in silico analysis of methylation of the HIV-1 provirus. *Methods*. 2011;53:47–53. <https://doi.org/10.1016/j.jymeth.2010.05.009>.
- Trejbalová K, Kovářová D, Blažková J, Machala L, Jilich D, Weber J, et al. Development of 5' LTR DNA methylation of latent HIV-1 provirus in cell line models and in long-term-infected individuals. *Clin Epigenet*. 2016;8:19. <https://doi.org/10.1186/s13148-016-0185-6>.
- Cortés-Rubio CN, Salgado-Montes de Oca G, Prado-Galbarro FJ, Matías-Florentino M, Murakami-Ogasawara A, Kuri-Cervantes L, et al. Longitudinal variation in human immunodeficiency virus long terminal repeat methylation in individuals on suppressive antiretroviral therapy. *Clin Epigenetics*. 2019;11:134. <https://doi.org/10.1186/s13148-019-0735-9>.
- Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet*. 2010;11:204–20. <https://doi.org/10.1038/nrg2719>.
- Attwood JT, Yung RL, Richardson BC. DNA methylation and the regulation of gene transcription. *Cell Mol Life Sci*. 2002;59:241–57. <https://doi.org/10.1007/s00018-002-8420-z>.
- Reik W. Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature*. 2007;447:425–32. <https://doi.org/10.1038/nature05918>.
- Kint S, De Spiegelaere W, De Kesel J, Vandekerckhove L, Van Crielinge W. Evaluation of bisulfite kits for DNA methylation profiling in terms of DNA fragmentation and DNA recovery using digital PCR. *PLoS One*. 2018;13:e0199091. <https://doi.org/10.1371/journal.pone.0199091>.

19. Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell*. 1999;99:247–57. [https://doi.org/10.1016/S0092-8674\(00\)81656-6](https://doi.org/10.1016/S0092-8674(00)81656-6).
20. Pradhan S, Bacolla A, Wells RD, Roberts RJ. Recombinant human DNA (cytosine-5) methyltransferase. *J Biol Chem*. 1999;274:33002–10. <https://doi.org/10.1074/jbc.274.46.33002>.
21. Denisenko O, Mar D, Trawczynski M, Bomszyk K. Chromatin changes trigger laminin genes dysregulation in aging kidneys. *Aging (Albany NY)*. 2018;10:1133–45. <https://doi.org/10.18632/aging.101453>.
22. Yoder JA, Walsh CP, Bestor TH. Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet*. 1997;13:335–40. [https://doi.org/10.1016/S0168-9525\(97\)01181-5](https://doi.org/10.1016/S0168-9525(97)01181-5).
23. Matzke MA, Mette MF, Aufsatz W, Jakowitsch J, Matzke AJM. Host defenses to parasitic sequences and the evolution of epigenetic control mechanisms. *Genetica*. 1999;107:271–87. https://doi.org/10.1007/978-94-011-4156-7_27.
24. Bird AP. DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Res*. 1980;8:1499–504. <https://doi.org/10.1093/nar/8.7.1499>.
25. Schorderet DF, Gartler SM. Analysis of CpG suppression in methylated and nonmethylated species. *Proc Natl Acad Sci U S A*. 1992;89:957–61.
26. Shukla S, Kavak E, Gregory M, Imshimizu M, Shutinoski B, Kashlev M, et al. CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature*. 2011;479:74–9. <https://doi.org/10.1038/nature10442>.
27. Jeziorska DM, Murray RJ, De Gobbi M, Gaentzsch R, Garrick D, Ayyub H, et al. DNA methylation of intragenic CpG islands depends on their transcriptional activity during differentiation and disease. *Proc Natl Acad Sci U S A*. 2017;114:E7526–35. <https://doi.org/10.1073/pnas.1703087114>.
28. Neri F, Rapelli S, Krepelova A, Incarnato D, Parlato C, Basile G, et al. Intragenic DNA methylation prevents spurious transcription initiation. *Nature*. 2017;543:72–7. <https://doi.org/10.1038/nature21373>.
29. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet*. 2012;13:484–92. <https://doi.org/10.1038/nrg3230>.
30. Teissandier A, Bourc'his D. Gene body DNA methylation conspires with H3K36me3 to preclude aberrant transcription. *EMBO J*. 2017;36:1471–3. <https://doi.org/10.15252/embj.201796812>.
31. Cassan E, Arigon-Chifolleau A-M, Mesnard J-M, Gross A, Gascuel O. Concomitant emergence of the antisense protein gene of HIV-1 and of the pandemic. *Proc Natl Acad Sci U S A*. 2016;113:11537–42. <https://doi.org/10.1073/pnas.1605739113>.
32. Blazkova J, Trejbalova K, Gondois-Rey F, Halfon P, Philibert P, Guiguen A, et al. CpG methylation controls reactivation of HIV from latency. *PLoS Pathog*. 2009;5:e1000554. <https://doi.org/10.1371/journal.ppat.1000554>.
33. Kauder SE, Bosque A, Lindqvist A, Planelles V, Verdin E. Epigenetic regulation of HIV-1 latency by cytosine methylation. *PLoS Pathog*. 2009;5:e1000495. <https://doi.org/10.1371/journal.ppat.1000495>.
34. Bouchat S, Delacourt N, Kula A, Darcis G, Van Driessche B, Corazza F, et al. Sequential treatment with 5-aza-2'-deoxycytidine and deacetylase inhibitors reactivates HIV-1. *EMBO Mol Med*. 2016;8:117–38. <https://doi.org/10.15252/emmm.201505557>.
35. Ho YC, Shan L, Hosmane NN, Wang J, Laskey SB, Rosenbloom DIS, et al. Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. *Cell*. 2013;155:540–51. <https://doi.org/10.1016/j.cell.2013.09.020>.
36. Blazkova J, Murray D, Justement JS, Funk EK, Nelson AK, Moir S, et al. Paucity of HIV DNA methylation in latently infected, resting CD4+ T cells from infected individuals receiving antiretroviral therapy. *J Virol*. 2012;86:5390–2. <https://doi.org/10.1128/JVI.00040-12>.
37. Weber S, Weiser B, Kemal KS, Burger H, Ramirez CM, Korn K, et al. Epigenetic analysis of HIV-1 proviral genomes from infected individuals: Predominance of unmethylated CpG's. *Virology*. 2014;449:181–9. <https://doi.org/10.1016/j.virol.2013.11.013>.
38. Palacios JA, Pérez-Piñar T, Toro C, Sanz-Minguela B, Moreno V, Valencia E, et al. Long-term nonprogressor and elite controller patients who control viremia have a higher percentage of methylation in their HIV-1 proviral promoters than aviremic patients receiving highly active antiretroviral therapy. *J Virol*. 2012;86:13081–4. <https://doi.org/10.1128/JVI.01741-12>.
39. Malatinkova E, De Spiegelaere W, Bonczkowski P, Kiselina M, Vervisch K, Trypsteen W, et al. Impact of a decade of successful antiretroviral therapy initiated at HIV-1 seroconversion on blood and rectal reservoirs. *Elife*. 2015;4:e09115. <https://doi.org/10.7554/eLife.09115>.
40. Malatinkova E, Kiselina M, Bonczkowski P, Trypsteen W, Messiaen P, Vermeire J, et al. Accurate quantification of episomal HIV-1 two-long terminal repeat circles by use of optimized DNA isolation and droplet digital PCR. *J Clin Microbiol*. 2015;53:699–701. <https://doi.org/10.1128/JCM.03087-14>.
41. Liszewski MK, Yu JJ, O'Doherty U. Detecting HIV-1 integration by repetitive-sampling Alu-gag PCR. *Methods*. 2009;47:254–60. <https://doi.org/10.1016/j.ymeth.2009.01.002>.
42. De Spiegelaere W, Malatinkova E, Lynch L, Van Nieuwerburgh F, Messiaen P, O'Doherty U, et al. Quantification of integrated HIV DNA by repetitive-sampling Alu-HIV PCR on the basis of poisson statistics. *Clin Chem*. 2014;60:886–95. <https://doi.org/10.1373/clinchem.2013.219378>.
43. Kiselina M, Pasternak AO, De Spiegelaere W, Vogelaers D, Berkhout B, Vandekerckhove L. Comparison of droplet digital PCR and seminested real-time PCR for quantification of cell-associated HIV-1 RNA. *PLoS One*. 2014;9:e85999. <https://doi.org/10.1371/journal.pone.0085999>.
44. Jordan A, Bisgrove D, Verdin E. HIV reproducibly establishes a latent infection after acute infection of T cells in vitro. *EMBO J*. 2003;22:1868–77. <https://doi.org/10.1093/emboj/cdg188>.
45. L-Cc L, Dahiya R. MethPrimer: designing primers for methylation PCRs. *Bioinformatics*. 2002;18:1427–31. <https://doi.org/10.1093/bioinformatics/18.11.1427>.
46. Rutsaert S, De Spiegelaere W, Van Hecke C, De Scheerder M-A, Kiselina M, Vervisch K, et al. In-depth validation of total HIV-1 DNA assays for quantification of various HIV-1 subtypes. *Sci Rep*. 2018;8:17274. <https://doi.org/10.1038/s41598-018-35403-6>.
47. Foley B, Leitner T, Apetrei C, Hahn B, Mizrahi I, Mullins J, et al. HIV sequence compendium 2016. *Theor Biol Biophys Group, Los Alamos Natl Lab NM, LA-UR-16-25625*. 2016.
48. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics*. 2011;27:1571–2. <https://doi.org/10.1093/bioinformatics/btr167>.
49. Akalin A, Korkmaksson M, Li S, Garrett-Bakelman FE, Figueroa ME, Melnick A, et al. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol*. 2012;13:R87. <https://doi.org/10.1186/gb-2012-13-10-r87>.
50. R Development Core Team. R: a language and environment for statistical computing. *R Found Stat Comput*. 2008;1:2673. <https://doi.org/10.1007/978-3-540-74686-7>.
51. McCullagh P, Nelder JA. Generalized linear models. 2nd dition. London: Chapman and Hall; 1989. <https://doi.org/10.1007/978-1-4899-3242-6>.
52. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple. vol. 57; 1995.
53. Du P, Zhang X, Huang C-C, Jafari N, Kibbe WA, Hou L, et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics*. 2010;11:587. <https://doi.org/10.1186/1471-2105-11-587>.
54. LaMere SA, Chaillon A, Huynh C, Smith DM, Gianella S. Challenges in quantifying cytosine methylation in the HIV provirus. *MBio*. 2019;10:e02268–18. <https://doi.org/10.1128/mBio.02268-18>.
55. Bednarik DP, Mosca JD, Raj NB. Methylation as a modulator of expression of human immunodeficiency virus. *J Virol*. 1987;61:1253–7.
56. Bednarik DP, Cook JA, Pitha PM. Inactivation of the HIV LTR by DNA CpG methylation: evidence for a role in latency. *EMBO J*. 1990;9:1157–64.
57. Ishida T, Hamano A, Koiwa T, Watanabe T. 5' long terminal repeat (LTR)-selective methylation of latently infected HIV-1 provirus that is demethylated by reactivation signals. *Retrovirology*. 2006;3:69. <https://doi.org/10.1186/1742-4690-3-69>.
58. Schulze-Forster K, Götz F, Wagner H, Kröger H, Simon D. Transcription of HIV1 is inhibited by DNA methylation. *Biochem Biophys Res Commun*. 1990;168:141–7. [https://doi.org/10.1016/0006-291X\(90\)91685-L](https://doi.org/10.1016/0006-291X(90)91685-L).
59. Nestor CE, Ottaviano R, Reinhardt D, Cruickshanks HA, Mjoseng HK, McPherson RC, et al. Rapid reprogramming of epigenetic and transcriptional profiles in mammalian culture systems. *Genome Biol*. 2015;16:11. <https://doi.org/10.1186/s13059-014-0576-y>.
60. Cruickshanks HA, McBryan T, Nelson DM, VanderKraats ND, Shah PP, van Tuyn J, et al. Senescent cells harbour features of the cancer epigenome. *Nat Cell Biol*. 2013;15:1495–506. <https://doi.org/10.1038/ncb2879>.
61. Hu Y, Huang K, An Q, Du G, Hu G, Xue J, et al. Simultaneous profiling of transcriptome and DNA methylome from a single cell. *Genome Biol*. 2016;17:88. <https://doi.org/10.1186/s13059-016-0950-z>.

62. Grau-Expósito J, Luque-Ballesteros L, Navarro J, Curran A, Burgos J, Ribera E, et al. Latency reversal agents affect differently the latent reservoir present in distinct CD4+ T subpopulations. *PLoS Pathog.* 2019;15:e1007991. <https://doi.org/10.1371/journal.ppat.1007991>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

