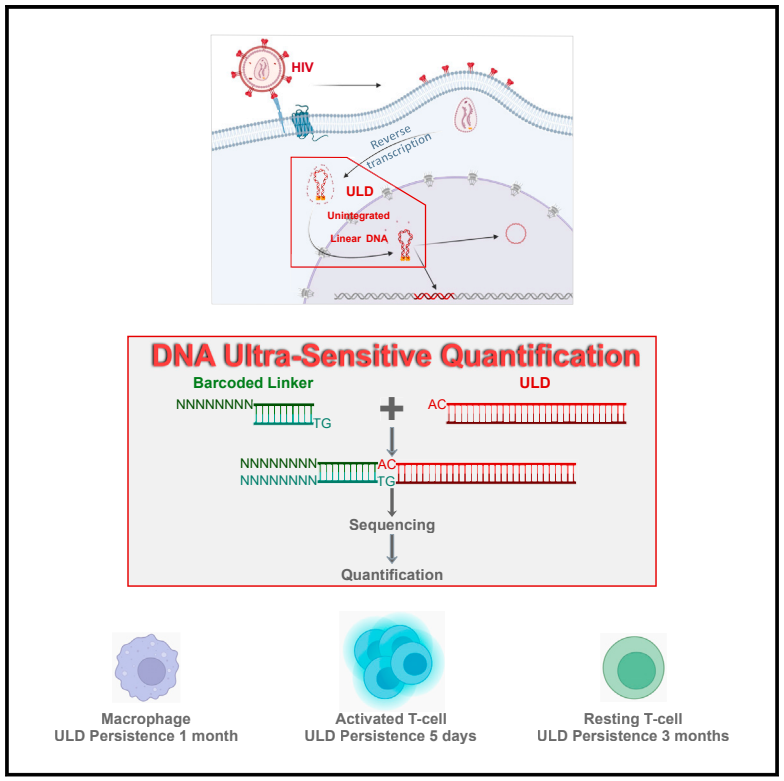
Cell Reports
Methods

Article

DNA ultra-sensitive quantification, a technology for studying HIV unintegrated linear DNA

Graphical abstract



Highlights

- DUSQ allows quantification of processed, unintegrated linear HIV DNA (pULD)
- DUSQ is based on linker-mediated PCR associated with molecular barcodes
- DUSQ has a detection threshold of 1 copy per million cells
- PULD can persist up to 3 months in quiescent T cells *in vitro*

Authors

Hélène Marie Roux, Suzanne Figueiredo, Lucas Sareoua, ..., François Clavel, Rémi Cheynier, Jacques Dutrieux

Correspondence

jacques.dutrieux@inserm.fr

In brief

HIV pre-integrative latency results from the persistence of unintegrated linear DNA (ULD) that can be further integrated into the host genome. Roux et al. develop DUSQ to specifically quantify ULDs with high sensitivity and show their persistence in resting T cells for up to 3 months, depending on cell activation.

Roux et al., 2023, Cell Reports Methods *3*, 100443
April 24, 2023
https://doi.org/10.1016/j.crmeth.2023.100443

CellPress

Cell Reports Methods

CellPress
OPEN ACCESS

Article

DNA ultra-sensitive quantification, a technology for studying HIV unintegrated linear DNA

Hélène Marie Roux,[1] Suzanne Figueiredo,[1] Lucas Sareoua,[1] Maud Salmona,[3,4,5] Juliette Hamroune,[1] Lucie Adoux,[1] Julie Migraine,[2] Allan Hance,[6] François Clavel,[3,5] Rémi Cheynier,[1] and Jacques Dutrieux[1,7,8,*]

[1]Université Paris Cité, Institut Cochin, INSERM U1016, CNRS, UMR8104, 75014 Paris, France
[2]INSERM U1259, Tours, France
[3]Université Paris Cité, Paris, France
[4]INSERM U976, Paris, France
[5]Assistance Publique Hôpitaux de Paris, Hôpital Saint Louis, Laboratoire de Virologie, Paris, France
[6]INSERM U941, Paris, France
[7]Viral DNA Integration and Chromatin Dynamics Network (DyNAVir), France
[8]Lead contact
*Correspondence: jacques.dutrieux@inserm.fr
https://doi.org/10.1016/j.crmeth.2023.100443

MOTIVATION The persistence of ULDs in cells has been observed in several studies by indirect methods and upon T cell activation at different time points post-infection *in vitro*. However, the detection and quantification of integration competent processed (p)ULDs by existing molecular biology techniques remains difficult due to their insufficient specificity and sensitivity. In this study, we developed an ultra-sensitive, specific, and high-throughput technology for pULD quantification called DUSQ (DNA ultra-sensitive quantification), which combines linker-mediated PCR with next-generation sequencing using molecular barcodes and has a detection threshold of 1 copy per million cells.

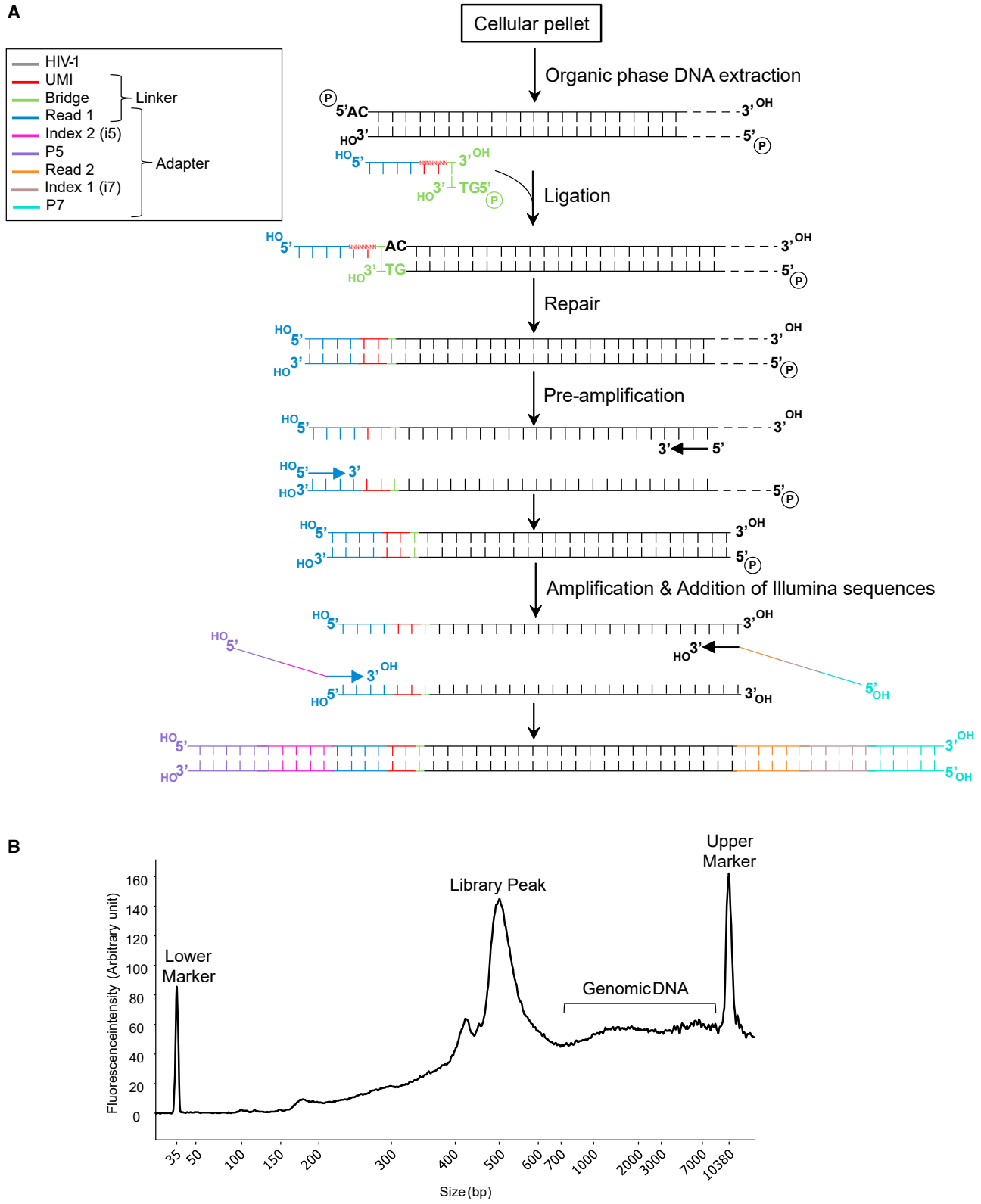
SUMMARY

Unintegrated HIV DNA represents between 20% and 35% of the total viral DNA in infected patients. Only the linear forms (unintegrated linear DNAs [ULDs]) can be substrates for integration and for the completion of a full viral cycle. In quiescent cells, these ULDs may be responsible for pre-integrative latency. However, their detection remains difficult due to the lack of specificity and sensitivity of existing techniques. We developed an ultra-sensitive, specific, and high-throughput technology for ULD quantification called DUSQ (DNA ultra-sensitive quantification) combining linker-mediated PCR and next-generation sequencing (NGS) using molecular barcodes. Studying cells with different activity levels, we determined that the ULD half-life goes up to 11 days in resting $CD4^+$ T cells. Finally, we were able to quantify ULDs in samples from patients infected with HIV-1, providing a proof of concept for the use of DUSQ *in vivo* to track pre-integrative latency. DUSQ can be adapted to the detection of other rare DNA molecules.

INTRODUCTION

The persistence of the human immunodeficiency virus (HIV) genome in cells is one of the major barriers to its eradication in treated patients. After HIV entry into the cell, viral RNA is reverse transcribed into a double-stranded DNA. In most cases, this viral DNA will be integrated and will persist in the cell host genome, thus forming the majority of the viral reservoir. However, unintegrated DNA can also be found either as linear or circular forms. The latter (truncated or not) can harbor 1- or 2-long terminal repeat (LTR; reviewed in Sloan and Wainberg[1]) and have very low levels of transcription[2–4] and no origin of replication. Due to these properties, the only possible substrates for viral integration are the unintegrated linear forms. The viral integrase (IN) homotetramer interacts directly with both extremities of the complete unintegrated linear DNA (ULD),[5,6] forming the intasome. This interaction rapidly leads to a 3′ processing of the ULD due to one of the catalytic activities of the IN. The processed ULD (pULD), which harbors 5′ AC overhangs, can be integrated into the host genome by the IN strand transfer activity.[6] pULD can also be degraded mostly by DNA repair enzymes[7] or circularized by auto-integration, recombination, or host DNA repair machinery.[1]

Several studies have focused on the role of ULDs in HIV persistence. The addition of an IN strand transfer inhibitors (INSTIs) modified the dynamics of viral load decay compared with that seen with reverse-transcriptase inhibitors. Indeed, patients treated with INSTI-based regimens reached



(legend on next page)

an undetectable viral load in 2–4 weeks after treatment initiation, whereas those treated without INSTI reached it in 2–4 months.[8,9] Murray et al. suggested that this difference was due to the presence of integration and replication-competent ULDs in some cells prior to treatment initiation, a phenomenon known as pre-integrative latency. In patients treated without INSTI, these ULDs could progressively integrate and lead to viral production slowing down the viral load decay. This hypothesis was reinforced by the mathematical modeling of HIV RNA decay in patients treated with an INSTI-based regimen.[10]

Although initial studies showed that, in resting T cells, reverse transcription did not reach completion,[11,12] it was demonstrated later on that the achievement of a full-length DNA is possible in these cells but as a much slower process than in activated T cells.[13] However, in resting T cells, the integration step is in fact quite rare,[14] suggesting that pre-integrative latency results from the infection of a cell while in a quiescent state.[15] The persistence of integration and replication-competent ULDs has been observed several times indirectly by quantifying viral production in resting T cells upon activation at different times post-infection. These studies are, however, contradictory, as some concluded that ULDs could no longer be integrated after 6 days post-infection,[16,17] whereas others found that ULDs could still be integrated after 14 or 21 days post-infection.[18] This contradiction is probably due to the use of different experimental approaches for cell activation or differences in cell viability *in vitro*. The latter study also showed that the pre-integration complex (PIC) can be stable for several weeks in close proximity with the centrosome.[18] Moreover, ULDs have been directly detected *in vivo* in untreated patients and *in vitro* in several studies using linker-mediated (LM)-PCR associated with southern blot[13,19] or qPCR.[20] The use of LM-PCR allowed an estimation of ULD half-life in resting T cells of around 1 day.

These conflicting conclusions could originate from a lack of sensitivity and specificity of the used methods to determine pULD half-life. Indeed, the association of LM-PCR with southern blot sensitivity was retrospectively estimated to be around 10^4 ULD copies/10^5 cells,[20] and the one associated with qPCR remains unclear. Therefore, the study of pULDs requires a new quantification method that solves these issues.

In this study, we developed a new technology called DNA ultra-sensitive quantification (DUSQ), which combines LM-PCR, molecular barcoding, and the Illumina next-generation sequencing (NGS) technology. This technology has a detection threshold of 1 pULD copy per sample, and the sequencing step allows the distinct detection and quantification of pULDs and not unprocessed ULDs (uULDs). This increased sensitivity and specificity obtained with DUSQ allowed us to determine the pULD half-life in cell types with various metabolic activities and activation states that were infected *in vitro*. We were also able to validate this technology to detect and quantify pULDs in cells from patients infected with HIV-1.

RESULTS

Molecular basis of ULD quantification using DUSQ

After organic-phase DNA extraction, an LM-PCR is realized (Figure 1A). First, pULDs are ligated to a partially double-stranded DNA adapter. It contains a double-stranded bridge with a 5′ phosphorylated GT overhang on the antisense strand, a 12-base-long single strand composed of random nucleotides called the unique molecular identifier (UMI), and the Illumina Read1 sequence. The UMIs in the ligated adapters give our technology its capacity for DNA ultra-sensitive quantification by acting as molecular barcodes for each ligated pULD, allowing amplification for their detection. Indeed, the fact that the UMIs are 12 nucleotides long means that there should be up to 4^{12} different adapters present during the ligation step, far over the expected number of pULDs in any sample. The ligation is followed by a Klenow fragment repair, resulting in a complete double-stranded DNA. A sizing based on magnetic beads is performed after each of these steps to remove the adapter excess and buffer salts.

A first amplification is done using a forward primer that binds to the Illumina Read1 sequence (present in the adapter) and a reverse primer that binds to a consensus sequence of the HIV-1 5′LTR (subtype B and circulating recombinant forms, which are mostly found in France[21]). A second amplification using primers containing the complete Illumina adapter sequences (including sample indexes, indexes 1 and 2) is performed. After each amplification, the excess of primers is removed by a sizing based on magnetic beads.

Library concentration is estimated by qPCR for each sample, and a library pool is then prepared at a final concentration of 4 nM. The quality of the library pool is verified on a Bioanalyzer chip (Figure 1B), allowing the visualization of a first peak around 415 bp corresponding to the first PCR round, a homogeneous peak around 520 bp corresponding to the final library, and a genomic DNA smear. Illumina sequencing in paired end is then performed. The sequences are generated in Fastq files with trimmed adapters therefore starting with the UMI in 5′.

DUSQ analysis

For quantification, only the sequences that originate from 5′ to 3′ sequencing are used, following the algorithm described in Figure 2. Sequences are sorted according to their quality score, and those with a Phred score below 20 are discarded. Simultaneously, the UMI of each sequence is trimmed and added at the end of the sequence name using the Fastp tool.[22] The sequences thus obtained are then aligned by the Bowtie2 program[23] with the corresponding pULD reference sequence.

Figure 1. DUSQ library preparation workflow

(A) Scheme of the DUSQ workflow. DNA is extracted by an organic phase method and ligated to a partially double-stranded adapter containing Illumina Read 1 sequence (blue), a unique molecular identifier $(N)_{12}$ (red), and a linker with a 5′ GT cohesive end (green). A reparation is performed to obtain a full-length double-stranded DNA that is pre-amplified using a forward primer specific of the Illumina Read 1 sequence and a reverse primer specific to HIV-1 5′ LTR. A second round of amplification is performed using primers containing the other Illumina sequences. Sequencing is accomplished using the MiSeq technology.

(B) Representative DNA size repartition of 96 samples pooled for sequencing at a concentration of 20 nM evaluated on a 2100 Bioanalyzer. Lower (35 bp) and upper markers (10,380 bp) are indicated. The library peak is around 520 bp, and genomic DNA can be seen at higher molecular sizes.

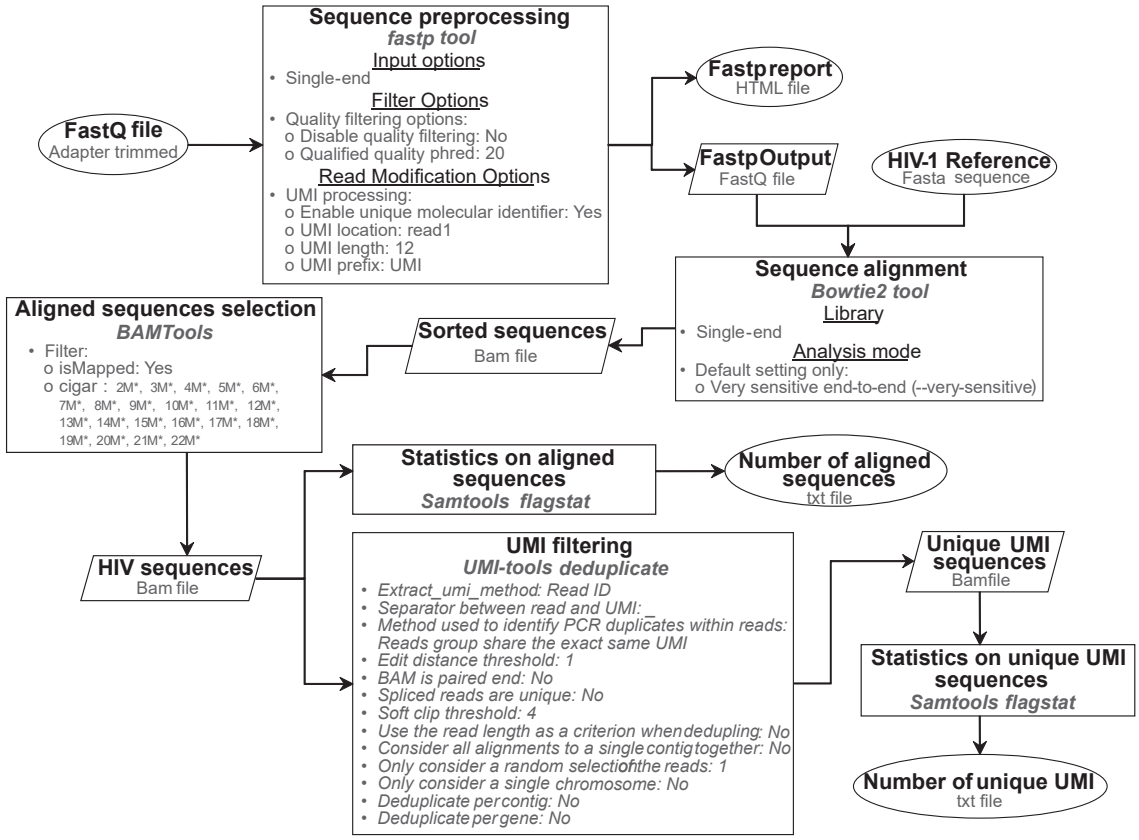


Figure 2. DUSQ analysis workflow

Complete algorithm of the DUSQ bioinformatics analysis workflow. After sequencing, the FastQ files are exported from the MiSeq platform with trimmed adapters. The FastQ files are then preprocessed using the fastp tool,[22] allowing quality filtering and UMI extraction. The output is then aligned on an HIV-1 reference sequence using the Bowtie2[23] tool, and the aligned sequences are filtered depending on their CIGAR string to store only the sequences that are aligned on the first 22 bases (only pULDs) using BAMTools.[24] Aligned sequences are then extracted for UMI filtering and counting using the UMI-tools deduplicate[25] (options that are not mentioned in the algorithm were disabled).

These sequences contain the sense strand of the adapter bridge in 5′ (20 bases), an AC, and finally, the first 150 bases of the 5′ LTR of the HIV-1 provirus strain used in the experiment or characterizing the studied patient. To ensure the DUSQ specificity for pULD, a filter is applied to discard the sequences that are not aligned on the 22 first bases according to their CIGAR string.[24] Indeed, as described previously,[13] uULDs can be ligated during LM-PCR (although with a lower efficiency[20]), and as a consequence of the repair step following the ligation, these molecules will harbor 2 successive ACs after the bridge sequence and therefore will not correctly align on the first 22 bases. The UMIs of the aligned sequences are deduplicated by UMI-tools deduplicate[25] so that for each group of identical UMIs, only one sequence is stored on a table as a BAM file. Therefore, each line of the table corresponds to a different UMI, and the lines can be counted using Samtools flagstat.[26]

In order to calculate the number of ULD copies contained in each sample based on the number of UMIs obtained, a control is required to establish a standard curve, which has been generated by serial dilutions of HIV-1-infected MT4R5 cells (sampled 8 h post-infection) in uninfected MT4R5 cells.

Sensitivity and specificity of DUSQ

Ultra-deep sequencing and deduplication of UMI implies the possibility of UMI collision, i.e., the fact that 2 different molecules harbor the same UMI sequence, and thus requires a sufficient UMI length so that the number of possible UMIs exceeds the number of target molecules.[27,28] To make sure that a UMI length of 12 nucleotides was suitable for DUSQ, we verified the base repartition after sequencing for all samples and found a probability of around 25% for each nucleotide at each position (Figure 3A), reflecting the absence of UMI collision.

In order to determine the quantification thresholds of our technology, we performed serial dilutions of HIV-1-infected MT4R5 cells (sampled 8 h post-infection) in uninfected MT4R5 cells and quantified pULDs in 14 replicates of each dilution using DUSQ. The absolute pULD copy number for the undiluted points was determined by qPCR to set the first theoretical copy number on the curve. We observed that no UMI was detected in samples in which there were no pULDs and estimated the DUSQ quantification threshold as 1 pULD copy per million cells (Figure 3B). In order to establish a standard curve for DUSQ, we performed a

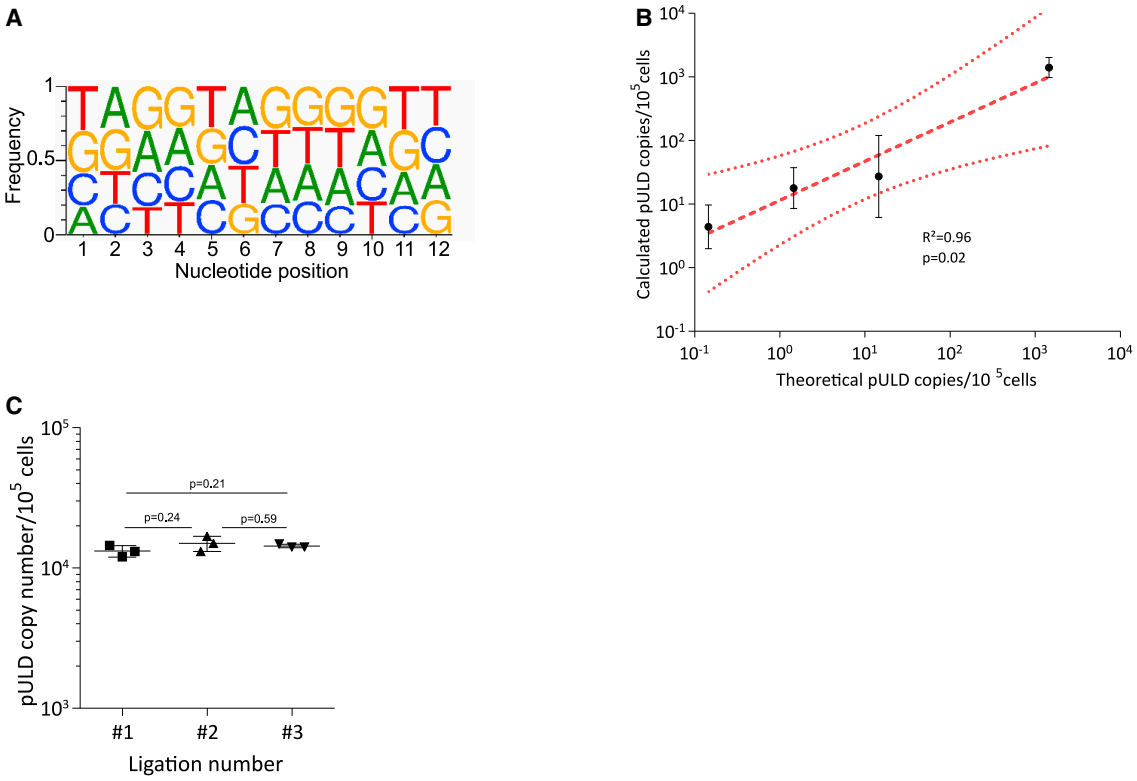


Figure 3. Sensitivity and reproducibility of DUSQ

(A) Representative sequence logo of the probability of finding each nucleotide at each of the 12 positions of the UMIs. These data were generated by Sequence Logo Generator[29] from the sequences issued from MDMs sampled 7 days post-infection after UMI filtering (156,136 reads).

(B) Determination of the DUSQ quantification threshold on serial dilutions of HIV-1-infected MT4R5 cells in uninfected MT4R5 cells. Mean ± SD of n = 14 replicates is represented. The linear fitness is represented by the dotted line with its 95% confidence interval. The linear fitness coefficient R^2 and its p value are indicated.

(C) Reproducibility of pULD quantification by DUSQ. MT4R5 cells were treated with dolutegravir 2 h pre-infection and infected by NL-4.3 during 24 h. Three DUSQ quantifications were performed on each of n = 3 independent ligations. Each dot represents a DUSQ quantification, and the mean ± SD is represented. All the quantifications gave data within a 2-fold range, and no significant difference was observed between the 3 ligations (p values are indicated, unpaired two-tailed t test).

linear regression (R^2 = 0.96; p = 0.02), which allowed for the quantification of pULDs in *in vitro* and *in vivo* samples.

DUSQ reproducibility was assessed on MT4R5 cells pre-treated with an INSTI (dolutegravir [DTG]) and infected with HIV-1 NL-4.3 strain for 24 h. Three independent ligations were performed, and DUSQ quantification was done in triplicate on each of these (Figure 3C). We did not observe any significant differences between each experiment, and the 9 values obtained in this experiment clustered within a 2-fold range (mean of 14,166.7 pULD copies/10^5 cells; SD ± 895; $IC_{95\%}$ [10,087; 19,563]) reflecting a high reproducibility of our pULD quantification technology.

Proof of concept for ULD quantification using DUSQ

It is likely that pULD stability depends, at least partially, on the state of quiescence of the host cell. In order to study this concept, we determined the pULD half-life in cell culture models with various levels of quiescence. Cells were pre-treated with an INSTI (DTG) 2 h before infection to allow accumulation of pULDs. A reverse-transcriptase inhibitor (nevirapine [NVP]) was then added 1 or 5 days post-infection (depending on the cell type) to inhibit new rounds of infection, thus allowing us to follow one pool of pULDs through time.

To make sure that cells were infected and that the pULDs had been synthesized, we quantified 2LTR circles by qPCR (Figure 4, left panels). The quantification of pULDs was then performed by DUSQ, and their half-lives were calculated by non-linear regression (Figure 4, middle and right panels).

We first used a T cell line (MT4R5) as a model of highly activated cells (Figure 4A). In order to avoid pULD dilution by cell proliferation, the latter was abolished by gamma irradiation prior to infection (Figure S1A). As seen with the 2LTR circles levels, the infectivity was not hindered by the irradiation. In this model, we determined a pULD half-life of 0.5 days (r^2 = 0.98; $IC_{95\%}$ [0.4; 0.6]).

We then used monocyte-derived macrophages (MDMs; Figure S1B) from two different healthy donors (Figure 4B) as a model for moderate cell activity. A pre-treatment with viral-like particles containing the SIVmac251 Vpx[30] protein allowed a very high level of infection, as shown by the 2LTR circle level. Although

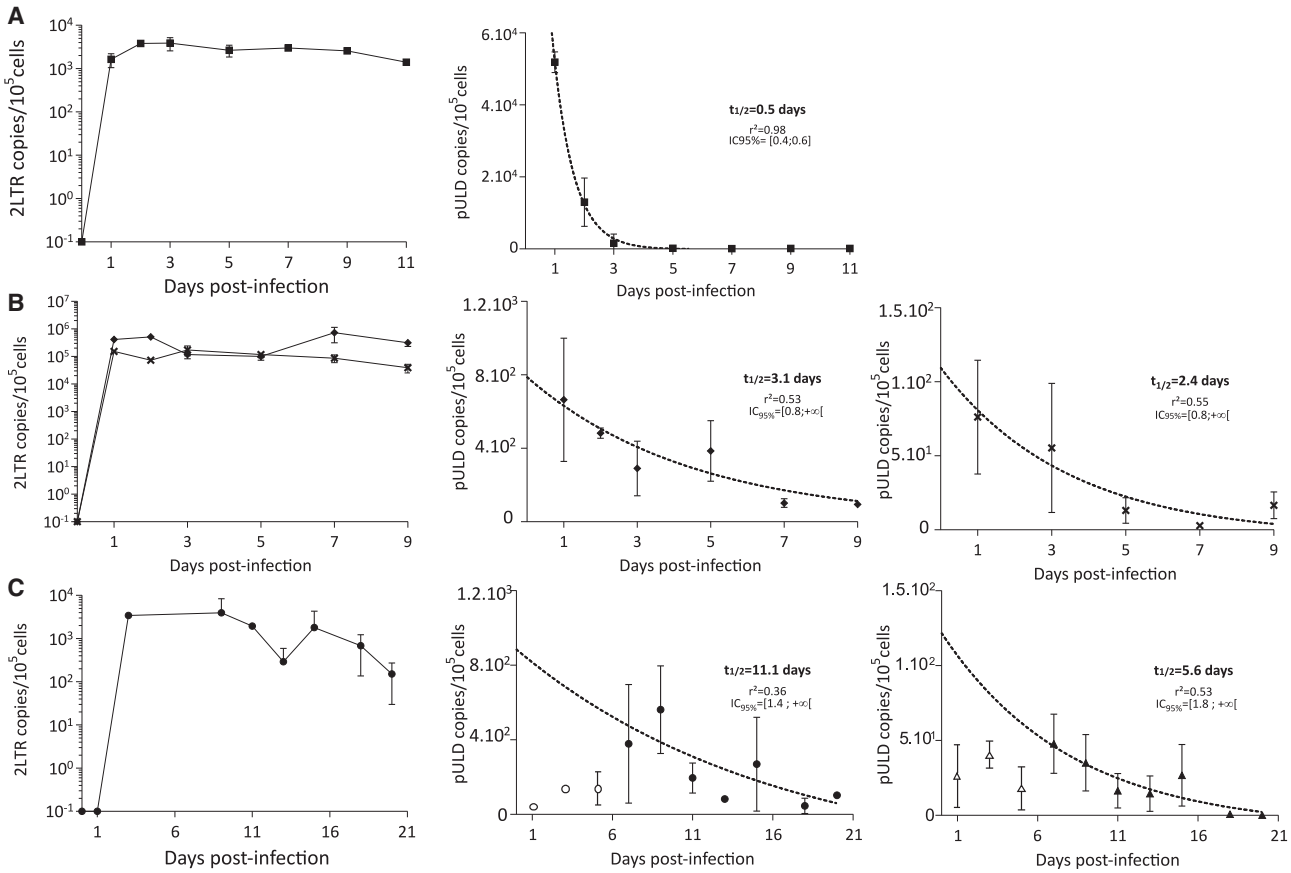


Figure 4. *In vitro* determination of pULDs half-life in MT4R5 cells, macrophages, and resting CD4⁺ T cells

Quantification of 2LTR circles (left panel) and pULDs (middle and right panels) over time post-infection by (A) NL-4.3 in MT4R5 cell line (squares), (B) NL-AD8 in monocyte-derived macrophages from 2 different healthy donors (diamonds and crosses), and (C) NL-4.3 in resting CD4⁺ T cells from 2 different healthy donors (circles and triangles; empty symbols represent the pULD quantification before the addition of NVP; the 2LTR quantification could not be determined for the second experiment). Mean ± SD of n = 3 independent replicates is represented. pULD half-life has been determined using a one-phase decay non-linear regression; the goodness of the fit is indicated with the r^2 coefficient and the half-life $IC_{95\%}$ is indicated.

the infection rate seems to be higher in MDMs than in MT4R5 cells, it is difficult to compare them due to the use of different virus strains, with the R5 tropism of the strain used on MDMs (NL-AD8) being more adapted to this cell type. In this model, we determined a pULD half-life of 3.1 days (Figure 4B, middle panel; r^2 = 0.53; $IC_{95\%}$ [0.8; +∞[) and 2.4 days (Figure 4B, right panel; r^2 = 0.55; $IC_{95\%}$ [0.8; +∞[).

Finally, we examined the pULD half-life in resting CD4⁺ T cells, defined as $CD4^+CD25^-CD69^-HLA\text{-}DR^-$ cells purified from peripheral blood mononuclear cells (PBMCs) from two different healthy donors (Figure 4C). They were cultured in the presence of interleukin-7 (IL-7) to promote long-term survival and to enhance the infectivity of HIV-1[31] without affecting their resting state (Figures S1C and S2). The infection rate of these cells was similar to that observed in MT4R5 cells but with a 1 day delay in time. This observation is consistent with the fact that reverse transcription is slower in quiescent cells, as previously described.[13] Interestingly, the pULD half-lives determined for these cells were 11.1 (Figure 4C, middle panel; r^2 = 0.36; $IC_{95\%}$ [1.4; +∞[) and 5.6 days (Figure 4C, right panel; r^2 = 0.53 $IC_{95\%}$[1.8; +∞[).

Therefore, DUSQ allowed us to demonstrate that the pULD half-life in resting T cells is up to 20-fold higher than in MT4R5s and up to 4-fold higher than in MDMs.

Finally, to assess DUSQ suitability for pULD quantification in cells from patients infected with HIV-1, DNA was extracted from PBMCs and/or spleen mononuclear cells (SMCs) from 6 untreated patients chronically infected with HIV-1 (Figure 5). In these samples, pULDs were not detected using LM-PCR associated with qPCR (data not shown). Interestingly, using DUSQ, pULDs were detectable in all samples, showing a pULD level of 49.9 in PBMCs (SEM ± 34.3) and 423.9 in SMCs (SEM ± 357).

DISCUSSION

Although integration is essential for virus replication in the cell, it does not always occur following reverse transcription of the viral genome. Indeed, a significant proportion of viral DNA found in cells is in a non-integrated form: 35% of total DNA in untreated patients and 20% of total DNA in treated patients with undetectable viremia.[32] Several studies have demonstrated the ability of

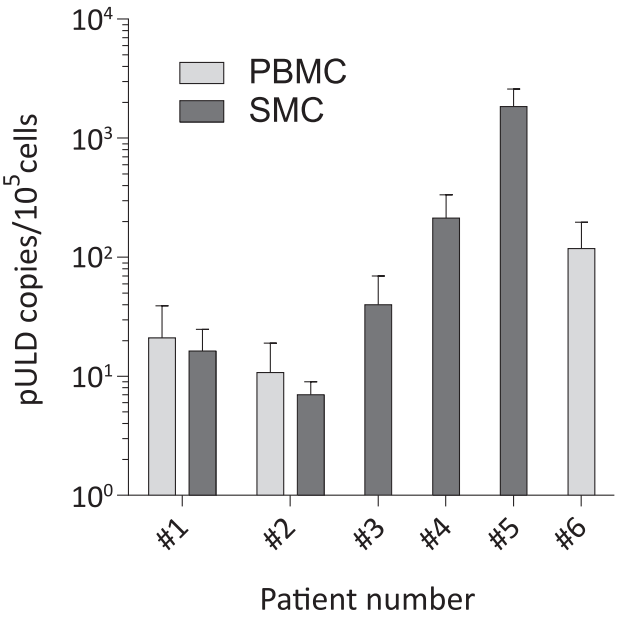


Figure 5. *In vivo* quantification of pULDs in splenocytes and blood of patients chronically infected with HIV-1

Quantification of pULDs in PBMCs or SMCs of 6 patients chronically infected with HIV-1. Mean ± SEM of n = 3 (patients #1, #2, 3, and #4) or n = 8 (patients #5 and #6) independent replicates is represented.

quiescent cells that do not contain integrated DNA to produce virus upon activation.[14,16–18] This observation can only be explained by the pre-integrative latency phenomenon, i.e., the ability of ULD to remain in the quiescent cell in the pre-integrated state and then to integrate and be expressed upon cell activation. To this day, the direct study of this phenomenon by the detection and quantification of pULDs remains difficult due to the lack of sensitivity and specificity of existing methods such as LM-PCR combined with southern blot or qPCR.[13,20]

In this study, we developed an ultra-sensitive, specific, and high-throughput technology for pULD quantification, DUSQ, which combines LM-PCR with NGS through the use of UMIs as molecular barcodes. To avoid any risk of collision during the deduplication of the PCR amplicons, we chose 12-nucleotide-long UMIs. In addition, sequencing allows us to differentiate uULDs from pULDs and therefore ensures the high specificity of DUSQ in regard to pULDs. Indeed, a uULD could be ligated to an adapter even though it has a blunt end in 5′, as observed previously.[13] However, ligated uULDs will present an ACAC sequence at the junction with the bridge due to the addition and the repair of the single-stranded GT of the linker and thus cannot be aligned with the reference sequence. In terms of sensitivity, DUSQ can detect as few as 1 copy of pULD per million cells. Moreover, the use of Illumina NGS allows a high throughput due to an extremely large library of potential indexes and the different sequencing depths available.

These properties of DUSQ allowed us to determine, in different cell types and according to their metabolic activity, the half-life of pULDs. Indeed, the persistence of pULDs in the cell depends on the impossibility of their integration in quiescent cells[14] and can be hindered by their circularization or their degradation. Thus, by mimicking the impossibility of integration by using an INSTI in our experiments, we were able to observe in a highly active cell line (MT4R5) in which proliferation had been abolished a pULD half-life of 0.5 days. In the organism, tissue-resident macrophages represent an important reservoir[33–37] and could also participate in pre-integrative latency. In our experiments, we demonstrated that the half-life of pULDs in MDMs was 5-fold longer than that observed in MT4R5s, demonstrating the ability of MDM cells to host this particular latency. $CD4^+$ T cells are the major target of HIV-1 infection, and their resting counterparts are the most likely to harbor long-term pre-integrative latency. In our experimental system, we were able to observe a half-life of pULDs in these cells up to 11 days, almost 4- and 20-fold longer than in MDMs and in the T cell line, respectively. This observation is consistent with previous studies where the induction of integration by resting T cell activation allowed viral production beyond 14 or even 21 days post-infection.[14,18] These results imply the fact that pre-integrative latency is not only a result of limited integration in quiescent cells. Indeed, when integration is inhibited in highly active cells, pULD half-life is still much shorter than in quiescent cells, suggesting that cell activity may play a role in their circularization and degradation and therefore in their persistence.

Finally, we were able to detect and quantify pULDs in mononuclear cells from blood and spleen samples from patients chronically infected with HIV-1. Although we cannot determine whether the pULDs we saw in these specimens belong to pre-integrative latency or simply originate from recent reverse transcription due to the lack of treatment in these patients, these results provide a proof of concept for the capacity of DUSQ to quantify pULDs in patient samples, allowing to use such a technology to track pre-integrative latency in future cohorts. Thus, DUSQ could be used to determine whether the second slope of viral load decay observed in patients initiating antiretroviral therapy without INSTI[9,10] would indeed be due to pre-integrative latency.

The DUSQ technology will allow a better understanding of the role of pre-integrative latency in viral DNA persistence and reservoir. Moreover, using DUSQ for the study of pULD dynamics within the infected cell could bring new insights into the early stages of the HIV viral cycle such as reverse transcription efficiency, capsid integrity, or nuclear import. Finally, DUSQ can also be easily adapted to the detection of any other rare DNA molecules that are not easily quantifiable by existing molecular biology techniques.

Limitations of the study

As a proof of concept for the use of DUSQ to quantify pULDs in patients, we included blood and spleen samples from patients who were untreated and chronically infected with HIV-1. Virological data were not available for these historical samples. However, they were taken from patients splenectomized for HIV-related disease, suggesting a poor virological control and an active viral replication. In order to assess pre-integrative latency in patients, it could be important to include patients on long-term antiretroviral therapy.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Human participants and samples
  - Cell lines and primary cell cultures
- METHOD DETAILS
  - Cell isolation and differentiation
  - Viral productions
  - *In vitro* cell infection and treatment
  - Flow cytometry staining
  - Immunofluorescence staining
  - DNA extraction
  - Adapter hybridization
  - Ligation and repair
  - Nested qPCR – Quantification of ULD and 2LTR circles by qPCR
  - Nested PCR
  - Library pool preparation and sequencing
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.crmeth.2023.100443.

ACKNOWLEDGMENTS

H.M.R. was the recipient of Université de Paris' doctoral fellowship. This work was carried out with the support of the Institut de Médecine et d'Epidémiologie Appliquée (IMEA), Viiv Healthcare France, and the French Agence Nationale de recherche sur le SIDA et les hépatites (ANRS, France). The authors express their gratitude to GENOM'IC, CYBIO, and HISTIM core facilities of the Institut Cochin, INSERM U1016, for their expert and helpful collaboration. The authors would like to thank both healthy volunteers and patients for their participation in the study. The Galaxy server that was used for some calculations is in part funded by Collaborative Research Centre 992 Medical Epigenetics (DFG grant SFB 992/1 2012) and German Federal Ministry of Education and Research (BMBF grants 031 A538A/A538C RBC, 031L0101B/031L0101C de.NBI-epi, and 031L0106 de.STAIR (de.NBI)).[38] The following reagent was obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: HIV-1 NL4-3 infectious molecular clone (pNL4-3) from Dr. Malcolm Martin (cat# 114). The following reagent was obtained through the NIH HIV Reagent Program, Division of AIDS, NIAID, NIH: HIV-1 AD8 infectious molecular clone (pNL(AD8)), ARP-11346, contributed by Dr. Eric O. Freed.

AUTHOR CONTRIBUTIONS

F.C., A.H., and J.D. developed the proof of concept. H.M.R., F.C., R.C., and J.D. designed experiments. H.M.R. and J.D. wrote the paper. H.M.R., S.F., J.M., L.S., L.A., and J.D. performed the experiments. H.M.R., J.H., M.S., L.S., and J.D. analyzed the results.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: March 4, 2022
Revised: January 28, 2023
Accepted: March 10, 2023
Published: April 5, 2023

REFERENCES

1. Sloan, R.D., and Wainberg, M.A. (2011). The role of unintegrated DNA in HIV infection. Retrovirology *8*, 52. https://doi.org/10.1186/1742-4690-8-52.
2. Wu, Y., and Marsh, J.W. (2001). Selective transcription and modulation of resting T cell activity by preintegrated HIV DNA. Science *293*, 1503–1506. https://doi.org/10.1126/science.1061548.
3. Wu, Y., and Marsh, J.W. (2003). Early transcription from nonintegrated DNA in human immunodeficiency virus infection. J. Virol. *77*, 10376–10382. https://doi.org/10.1128/jvi.77.19.10376-10382.2003.
4. Gelderblom, H.C., Vatakis, D.N., Burke, S.A., Lawrie, S.D., Bristol, G.C., and Levy, D.N. (2008). Viral complementation allows HIV-1 replication without integration. Retrovirology *5*, 60. https://doi.org/10.1186/1742-4690-5-60.
5. Esposito, D., and Craigie, R. (1998). Sequence specificity of viral end DNA binding by HIV-1 integrase reveals critical regions for protein-DNA interaction. EMBO J. *17*, 5832–5843. https://doi.org/10.1093/emboj/17.19.5832.
6. Maertens, G.N., Hare, S., and Cherepanov, P. (2010). The mechanism of retroviral integration from X-ray structures of its key intermediates. Nature *468*, 326–329. https://doi.org/10.1038/nature09517.
7. Yoder, K., Sarasin, A., Kraemer, K., McIlhatton, M., Bushman, F., and Fishel, R. (2006). The DNA repair genes XPB and XPD defend cells from retroviral infection. Proc. Natl. Acad. Sci. USA *103*, 4622–4627. https://doi.org/10.1073/pnas.0509828103.
8. Perelson, A.S., Essunger, P., Cao, Y., Vesanen, M., Hurley, A., Saksela, K., Markowitz, M., and Ho, D.D. (1997). Decay characteristics of HIV-1-infected compartments during combination therapy. Nature *387*, 188–191. https://doi.org/10.1038/387188a0.
9. Murray, J.M., Emery, S., Kelleher, A.D., Law, M., Chen, J., Hazuda, D.J., Nguyen, B.Y.T., Teppler, H., and Cooper, D.A. (2007). Antiretroviral therapy with the integrase inhibitor raltegravir alters decay kinetics of HIV, significantly reducing the second phase. AIDS *21*, 2315–2321. https://doi.org/10.1097/QAD.0b013e3282f12377.
10. Cardozo, E.F., Andrade, A., Mellors, J.W., Kuritzkes, D.R., Perelson, A.S., and Ribeiro, R.M. (2017). Treatment with integrase inhibitor suggests a new interpretation of HIV RNA decay curves that reveals a subset of cells with slow integration. PLoS Pathog. *13*, e1006478. https://doi.org/10.1371/journal.ppat.1006478.
11. Zack, J.A., Arrigo, S.J., Weitsman, S.R., Go, A.S., Haislip, A., and Chen, I.S. (1990). HIV-1 entry into quiescent primary lymphocytes: molecular analysis reveals a labile, latent viral structure. Cell *61*, 213–222. https://doi.org/10.1016/0092-8674(90)90802-l.
12. Zack, J.A., Haislip, A.M., Krogstad, P., and Chen, I.S. (1992). Incompletely reverse-transcribed human immunodeficiency virus type 1 genomes in quiescent cells can function as intermediates in the retroviral life cycle. J. Virol. *66*, 1717–1725.
13. Pierson, T.C., Zhou, Y., Kieffer, T.L., Ruff, C.T., Buck, C., and Siliciano, R.F. (2002). Molecular characterization of preintegration latency in human immunodeficiency virus type 1 infection. J. Virol. *76*, 8518–8531.
14. Stevenson, M., Stanwick, T.L., Dempsey, M.P., and Lamonica, C.A. (1990). HIV-1 replication is controlled at the level of T cell activation and proviral integration. The EMBO Journal *9*, 1551–1560.
15. Donahue, D.A., and Wainberg, M.A. (2013). Cellular and molecular mechanisms involved in the establishment of HIV-1 latency. Retrovirology *10*, 11. https://doi.org/10.1186/1742-4690-10-11.
16. Zhou, Y., Zhang, H., Siliciano, J.D., and Siliciano, R.F. (2005). Kinetics of human immunodeficiency virus type 1 decay following entry into resting

CD4+ T cells. J. Virol. *79*, 2199–2210. https://doi.org/10.1128/JVI.79.4.2199-2210.2005.

17. Petitjean, G., Al Tabaa, Y., Tuaillon, E., Mettling, C., Baillat, V., Reynes, J., Segondy, M., and Vendrell, J.P. (2007). Unintegrated HIV-1 provides an inducible and functional reservoir in untreated and highly active antiretroviral therapy-treated patients. Retrovirology *4*, 60. https://doi.org/10.1186/1742-4690-4-60.
18. Zamborlini, A., Lehmann-Che, J., Clave, E., Giron, M.L., Tobaly-Tapiero, J., Roingeard, P., Emiliani, S., Toubert, A., de Thé, H., and Saïb, A. (2007). Centrosomal pre-integration latency of HIV-1 in quiescent cells. Retrovirology *4*, 63. https://doi.org/10.1186/1742-4690-4-63.
19. Chun, T.W., Carruth, L., Finzi, D., Shen, X., DiGiuseppe, J.A., Taylor, H., Hermankova, M., Chadwick, K., Margolick, J., Quinn, T.C., et al. (1997). Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. Nature *387*, 183–188. https://doi.org/10.1038/387183a0.
20. Munir, S., Thierry, S., Subra, F., Deprez, E., and Delelis, O. (2013). Quantitative analysis of the time-course of viral DNA forms during the HIV-1 life cycle. Retrovirology *10*, 87. https://doi.org/10.1186/1742-4690-10-87.
21. Chaillon, A., Essat, A., Frange, P., Smith, D.M., Delaugerre, C., Barin, F., Ghosn, J., Pialoux, G., Robineau, O., Rouzioux, C., et al. (2017). Spatiotemporal dynamics of HIV-1 transmission in France (1999-2014) and impact of targeted prevention strategies. Retrovirology *14*, 15. https://doi.org/10.1186/s12977-017-0339-4.
22. Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics *34*, i884–i890. https://doi.org/10.1093/bioinformatics/bty560.
23. Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25. https://doi.org/10.1186/gb-2009-10-3-r25.
24. Barnett, D.W., Garrison, E.K., Quinlan, A.R., Strömberg, M.P., and Marth, G.T. (2011). BamTools: a C++ API and toolkit for analyzing and managing BAM files. Bioinformatics *27*, 1691–1692. https://doi.org/10.1093/bioinformatics/btr174.
25. Smith, T., Heger, A., and Sudbery, I. (2017). UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. Genome Res. *27*, 491–499. https://doi.org/10.1101/gr.209601.116.
26. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMtools. Bioinformatics *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.
27. Clement, K., Farouni, R., Bauer, D.E., and Pinello, L. (2018). AmpUMI: design and analysis of unique molecular identifiers for deep amplicon sequencing. Bioinformatics *34*, i202–i210. https://doi.org/10.1093/bioinformatics/bty264.
28. Kou, R., Lam, H., Duan, H., Ye, L., Jongkam, N., Chen, W., Zhang, S., and Li, S. (2016). Benefits and challenges with applying unique molecular identifiers in next generation sequencing to detect low frequency mutations. PLoS One *11*, e0146638. https://doi.org/10.1371/journal.pone.0146638.
29. Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004). WebLogo: a sequence logo generator. Genome Res. *14*, 1188–1190. https://doi.org/10.1101/gr.849004.
30. Berger, G., Durand, S., Goujon, C., Nguyen, X.N., Cordeil, S., Darlix, J.L., and Cimarelli, A. (2011). A simple, versatile and efficient method to genetically modify human monocyte-derived dendritic cells with HIV-1-derived lentiviral vectors. Nat. Protoc. *6*, 806–816. https://doi.org/10.1038/nprot.2011.327.
31. Coiras, M., Bermejo, M., Descours, B., Mateos, E., García-Pérez, J., López-Huertas, M.R., Lederman, M.M., Benkirane, M., and Alcamí, J. (2016). IL-7 induces SAMHD1 phosphorylation in CD4+ T lymphocytes, improving early steps of HIV-1 life cycle. Cell Rep. *14*, 2100–2107. https://doi.org/10.1016/j.celrep.2016.02.022.
32. Orlandi, C., Canovari, B., Bozzano, F., Marras, F., Pasquini, Z., Barchiesi, F., De Maria, A., Magnani, M., and Casabianca, A. (2020). A comparative analysis of unintegrated HIV-1 DNA measurement as a potential biomarker of the cellular reservoir in the blood of patients controlling and non-controlling viral replication. J. Transl. Med. *18*, 204. https://doi.org/10.1186/s12967-020-02368-y.
33. Igarashi, T., Brown, C.R., Endo, Y., Buckler-White, A., Plishka, R., Bischofberger, N., Hirsch, V., and Martin, M.A. (2001). Macrophage are the principal reservoir and sustain high virus loads in rhesus macaques after the depletion of CD4+ T cells by a highly pathogenic simian immunodeficiency virus/HIV type 1 chimera (SHIV): implications for HIV-1 infections of humans. Proc. Natl. Acad. Sci. USA *98*, 658–663. https://doi.org/10.1073/pnas.021551798.
34. Zalar, A., Figueroa, M.I., Ruibal-Ares, B., Baré, P., Cahn, P., de Bracco, M.M.d.E., and Belmonte, L. (2010). Macrophage HIV-1 infection in duodenal tissue of patients on long term HAART. Antiviral Res. *87*, 269–271. https://doi.org/10.1016/j.antiviral.2010.05.005.
35. Cribbs, S.K., Lennox, J., Caliendo, A.M., Brown, L.A., and Guidot, D.M. (2015). Healthy HIV-1-infected individuals on highly active antiretroviral therapy harbor HIV-1 in their alveolar macrophages. AIDS Res. Hum. Retroviruses *31*, 64–70. https://doi.org/10.1089/AID.2014.0133.
36. Ko, A., Kang, G., Hattler, J.B., Galadima, H.I., Zhang, J., Li, Q., and Kim, W.K. (2019). Macrophages but not astrocytes harbor HIV DNA in the brains of HIV-1-Infected aviremic individuals on suppressive antiretroviral therapy. J. Neuroimmune Pharmacol. *14*, 110–119. https://doi.org/10.1007/s11481-018-9809-2.
37. Ganor, Y., Real, F., Sennepin, A., Dutertre, C.A., Prevedel, L., Xu, L., Tudor, D., Charmeteau, B., Couedel-Courteille, A., Marion, S., et al. (2019). HIV-1 reservoirs in urethral macrophages of patients under suppressive antiretroviral therapy. Nat. Microbiol. *4*, 633–644. https://doi.org/10.1038/s41564-018-0335-z.
38. Afgan, E., Baker, D., Batut, B., van den Beek, M., Bouvier, D., Cech, M., Chilton, J., Clements, D., Coraor, N., Grüning, B.A., et al. (2018). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Res. *46*, W537–W544. https://doi.org/10.1093/nar/gky379.
39. Amara, A., Vidy, A., Boulla, G., Mollier, K., Garcia-Perez, J., Alcamí, J., Blanpain, C., Parmentier, M., Virelizier, J.L., Charneau, P., and Arenzana-Seisdedos, F. (2003). G protein-dependent CCR5 signaling is not required for efficient infection of primary T lymphocytes and macrophages by R5 human immunodeficiency virus type 1 isolates. J. Virol. *77*, 2550–2558. https://doi.org/10.1128/jvi.77.4.2550-2558.2003.
40. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nature methods *9*, 676–682. https://doi.org/10.1038/nmeth.2019.
41. Perrin, S., Firmo, C., Lemoine, S., Le Crom, S., and Jourdren, L. (2017). Aozan: an automated post-sequencing data-processing pipeline. Bioinformatics *33*, 2212–2213. https://doi.org/10.1093/bioinformatics/btx154.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-CD4 PE-Vio770	Miltenyi Biotec	Cat# 130-113-227; RRID: AB_2726038
Anti-CD25-APC	Miltenyi Biotec	Cat# 130-115-535; RRID: AB_2727076
Anti-CD69-APC-Vio770	Miltenyi Biotec	Cat# 130-112-616; RRID: AB_2659077
Anti-HLA-DR-VioBlue	Miltenyi Biotec	Cat# 130-111-794; RRID: AB_2652162
Anti-MAC387	ABD SEROTEC	Cat# MCA874G; RRID: AB_321963
Anti-CD14	Millipore	Cat# CBL453; RRID: AB_93594
Bacterial and virus strains		
pNL4-3	NIH AIDS reagent program	Cat# ARP-114
pNL(AD8)	NIH AIDS reagent program	Cat# ARP-11346
pVSV-G	Addgene	Cat# 138479
pSIV3+	Berger, G. et al. Nature Protocols 2011[30]	Provided by Dr A. Cimarelli, ENS Lyon, France
Biological samples		
Healthy donors blood	Etablissement Français du Sang	https://dondesang.efs.sante.fr/
HIV-1 chronically infected PBMCs and SMCs	Hôpital Saint Louis, Paris, France (1992–1996)	N/A
Chemicals, peptides, and recombinant proteins		
Human IL-7 Premium grade	Miltenyi Biotec	Cat# 130-095-362
Human M-CSF Premium grade	Miltenyi Biotec	Cat# 130-096-489
Benzonase Nuclease	Merck Millipore	Cat# 70746
Dolutegravir	Selleckchem	Cat# S2667
Nevirapine	Selleckchem	Cat# S1742
Proteinase K	Qiagen	Cat# 19131
RNase A	Roche	Cat# 10109142001
T4 DNA ligase	Themo Scientific	Cat# EL0012
Klenow Fragment	Thermo Scientific	Cat# EP0052
Phusion High-Fidelity PCR Master Mix with HF Buffer	Thermo Scientific	Cat# F531L
SPRIselect	Beckman Coulter	Cat# B23318
PEG NaCl Solution	Swift Biosciences	Cat# SW90196
Low EDTA TE Solution	Swift Biosciences	Cat# SW90296
Critical commercial assays		
Lenti-X p24 Rapid Titer kit	Takara Bio	Cat# 632200
β-galactosidase enzyme assay system	Promega	Cat# E2000
EasySep Human Resting $CD4^+$ T-cell Isolation Kit	Stemcell	Cat# 17962
Kapa Library quantification Kit	Roche	Cat# 07960298001
Bioanalyzer High Sensitivity DNA	Agilent	Cat# 5067-4626
MiSeq Reagent Kit v2 (300 cycles)	Illumina	Cat# MS-102-2002
Deposited data		
HIV-1 NL4-3 reference sequence	Los Alamos national laboratory HIV Sequence Database	https://www.hiv.lanl.gov/content/index
HIV-1 B.US.1985 reference sequence	Los Alamos national laboratory HIV Sequence Database	https://www.hiv.lanl.gov/content/index

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental models: Cell lines		
MT4-R5	Amara A., et al., Journal of Virology, 2003[39]	N/A
HEK-293T	ATCC	Cat# CRL-3216
HeLa TzmBL	NIH AIDS Reagent Program	Cat# ARP-11795
Oligonucleotides		
See Table S1	This paper	N/A
Software and algorithms		
Roche LightCycler480 Software	Roche	Cat# 04994884001
FlowJo V10	Becton Dickinson	https://www.flowjo.com/solutions/flowjo
ImageJ (FiJi 1.51W)	Schindelin J., et al., Nature Methods, 2012[40]	https://fiji.sc/
Fastp	Chen S., et al., Bioinformatics, 2018[22]	https://usegalaxy.eu/
Bowtie2	Langmead B., et al., Nature Methods, 2009[23]	https://usegalaxy.eu/
BAMtools	Barnett D. W., et al., Bioinformatics, 2011[24]	https://usegalaxy.eu/
Samtools	Li H., et al., Bioinformatics, 2009[26]	https://usegalaxy.eu/
UMI-tools deduplicate	Smith T., et al., Genome Research, 2017[25]	https://usegalaxy.eu/
Graphpad Prism 6	Graphpad	https://www.graphpad.com

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Jacques Dutrieux (jacques.dutrieux@inserm.fr).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Data availability: data reported in this paper will be shared by the lead contact upon request.
- Code availability: This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human participants and samples

Healthy blood was obtained from two donors from the Etablissement Français du Sang under convention number 18/EFS/30.

Six french male patients (age unknown) chronically infected with HIV-1, were included in this study. They were sampled at the Hôpital Saint-Louis (Paris France) between 1992 and 1996. They were untreated and their thrombocytopenia required a splenectomy. At sampling time, the French legislation neither required ethical approval nor inform consent for sample considered as surgical waste.

Cell lines and primary cell cultures

MT4-R5[39] cells, resting LT-CD4$^+$ and MDM were cultivated in Roswell Park Memorial Institute medium (RPMI, Gibco) containing GlutaMAX supplemented with 10% Fetal Bovine Serum (FBS, Gibco), penicillin (100U/mL, Gibco) and streptomycin (100μg/mL, Gibco), 20 ng/mL human IL-7 (for resting LT-CD4$^+$ cells, Miltenyi Biotec) and 50 ng/mL M-CSF (for MDM, Miltenyi Biotech).

HEK-293T (ATCC) and HeLa TzmBL (NIH AIDS Reagent) cells were cultivated in Dulbecco's Modified Eagle Medium (DMEM, Gibco) containing GlutaMAX supplemented with 10% Fetal Bovine Serum (FBS, Gibco), penicillin (100U/mL) and streptomycin (100μg/mL).

METHOD DETAILS

Cell isolation and differentiation

Peripheral blood mononuclear cells were purified by centrifugation at 1000*g* during 20 min at room temperature on Ficoll Paque Plus (GE Healthcare). The PBMC layer was collected and underwent 3 washes in PBS (5min centrifugations at 600 g at room temperature).

Resting $CD4^+$ T cell ($CD4^+CD69^-CD25^-HLA\text{-}DR^-$) were isolated by magnetic bead separation (EasySep Human Resting $CD4^+$ T cell Isolation Kit, Stemcell), according to the manufacturer protocol.

For MDM differentiation, PBMC were seeded at a concentration of 1.10^7cells/mL and incubated at 37°C during 3h. Adherent cells were recovered and seeded at a concentration of 2.10^5cells/mL in presence of 50 ng/mL of M-CSF during 7 days.

Viral productions

5.10^6 HEK 293T cells were seeded 24h before transfection in a 10cm diameter Petri dish with, in 10mL of complete DMEM. 2 h before transfection, the medium was replaced.

500μL of a mix containing 30μg plasmids mix (pNL4-3 plasmid (NIH AIDS Reagent Program); 30μg pNL-(AD8) plasmid (NIH AIDS Reagent); 20μg of pSIV3+ plasmid (provided by Dr A. Cimarelli) and 10μg of pVSV-G plasmid (addgene)) and 250mM of $CaCl_2$ was added slowly and under agitation to 500μL of 2X HEPES Buffer Saline (HBS, Sigma Aldrich). After a 30min incubation at room temperature, 1mL of the mix was added drop by drop to the Petri dish, which was then placed 24 h at 37°C.

The medium was removed and the excess was eliminated by dilution with 1X Phosphate Buffer Saline (D-PBS, Gibco). D-PBS was removed and replaced with 5mL of medium without FBS and cells were placed 24 h at 37°C.

The supernatant was collected and centrifuged at 1000g for 5min. It was then filtered at 0.45μm and treated with Benzonase (250U/mL, Merck Millipore) for 30 min at 37°C. The viral production was aliquoted and stored at −80°C for NL4.3 and NL-AD8. For Vpx containing VLP, viral particles were concentrated by ultracentrifugation on a 20% sucrose cushion at 82000g for 2h at 4°C. The viral pellet was resuspended in D-PBS and stored at −80°C.

Viral concentration was estimated using Enzyme-Linked Immunosorbent Assay (ELISA) of p24 protein (Lenti-X p24 Rapid Titer kit, Takara Bio), following the manufacturer's protocol. The infectious dose 50 (ID_{50}) of each virus was determined by serial dilution of the viruses on reporter cell line HeLa TzmBL and β-galactosidase activity was determined 24h post-infection using β-galactosidase enzyme assay (Promega) according to the manufacturer protocol.

In vitro cell infection and treatment

Cells were pre-treated 2h prior infection with 5μM Dolutegravir (DTG, Selleckchem) and, for MDM with $200ng_{p27}$/mL of Vpx containing VLP. MT4-R5 cells were infected with 1 ID_{50}/mL of the NL-4.3 viral strain, resting LT-$CD4^+$ cells with 5 ID_{50}/mL of the NL-4.3 viral strain and MDM with 1 ID_{50}/mL of the NL-AD8 viral strain. 10μM Nevirapine (NVP, Selleckchem) was added 1 day (for MT4R5 and MDM) or 5 days (for resting LT-$CD4^+$). Medium was replaced every day of the experiment.

Flow cytometry staining

Resting LT-$CD4^+$ were stained with CellTrace CFSE (Invitrogen) according to the manufacturer's protocol and cultured in presence of human IL-7 (20 ng/mL, Miltenyi Biotec) or activated with anti-CD3/CD28 beads (Thermo Scientific). Cells were sampled at different time points post-culture, washed in D-PBS by 5min centrifugation at 4 °C at 600g and stained for viability using the Fixable Live/Dead Aqua reagent (Thermo Scientific) according to the manufacturer protocol. Cells were then washed and stained with anti-CD4-PE-Vio770, anti-CD25-APC, anti-CD69-APC-Vio770 and anti-HLA-DR-VioBlue (Miltenyi Biotec) according to the manufacturer protocol. Acquisition was performed using a BD LSRII flow cytometer (Becton Dickinson). Results were analyzed using the FlowJo v10 Software (BD Biosciences).

Immunofluorescence staining

MDMs were differentiated on a 18mm glass cover slide in presence of 50 ng/mL of M-CSF (Miltenyi Biotec) during 1 week. Cells were then fixated 5min using cold Methanol. Free aldehyde groups were quenched using 50mM NH_4Cl for 15min and slides were then saturated in PBS containing 5% BSA (Sigma Aldrich) and 10% Normal Goat Serum (Sigma Aldrich) for 30min. Primary antibodies were incubated overnight at 4°C (anti-CD14 (UCHM-1, Millipore), anti-MAC387 (ABD Serotec)). Slides were washed 5min 3 times in PBS containing 0.5% Tween 20 and secondary antibodies (anti-IgG2a Alexa Fluor 488, anti-IgG1 Alexa Fluor 546, Thermo Scientific) were added and incubated 2h at room temperature. Slides were washed and nucleus were stained by DAPI (Sigma Aldrich) before mounting using Fluoromont G (Southern Biotechnology). 88 fields were acquired using a Lamina slide scanner (Perkin Elmer) and were analyzed using the Fiji Software.

DNA extraction

Cells were lysed for 1h at 56°C in lysis buffer (Tris-HCl pH 8.0 10mM (Sigma Aldrich), EDTA 10mM (Sigma Aldrich); SDS 0.5% (Sigma Aldrich); Proteinase K 200μg/mL (Qiagen); RNase A 200μg/mL (Roche)). 1 volume of pH 8.6 equilibrated phenol (Sigma Aldrich) was added to the lysate, homogenized by agitation for 15s then centrifuged at 13000g for 20 min at 15°C. The upper aqueous phase was collected and 1/2 volume of pH 8.5 buffered phenol and 1/2 volume of chloroform containing isoamyl alcohol (1:24, Sigma Aldrich) were added, mixed and centrifuged as previously. The upper aqueous phase was collected and 1 volume of chloroform containing isoamyl alcohol (1:24, Sigma Aldrich) was added, mixed and centrifuged again. The upper aqueous phase was collected, $1/10^{th}$ volume of 3M acetate sodium and 2 volumes of ethanol were added. The mix was homogenized by inversion then centrifuged at 13000g for 40 min at 4°C. The supernatant was discarded and the DNA pellet was resuspended in 70% ethanol and centrifuged for 10 min at 13000 g at 4°C. The supernatant was discarded and the DNA pellet was dried then resuspended in 50μL of Low

EDTA TE (Tris Base 10mM, EDTA 0.1mM, Swift Biosciences). DNA was then quantified by either a Nanodrop or a Qubit 3.0 Fluorometer following the manufacturer protocol for Broad Range Double Stranded DNA quantification.

Adapter hybridization

Primers « Linker 1 » and « Linker 2 » (sequences in Table S1) were added in equal volumes, each at a concentration of 50μM, then hybridized according to the following program: 5 min at 95°C, 10 min at 25°C with a temperature decrease of 0.1 °C/s.

Ligation and repair

The adapter was added to 400ng of DNA sample, at a concentration of 50nM, in a total volume of 15μL. 5μL of ligation solution (2μL 10X T4 DNA Ligase Buffer, 2μL 50% PEG4000, 0.4U T4 DNA Ligase, H_2O to 5μL, Thermo Scientific) was added and ligation was performed according to the following program: 1h at 22°C, 10 min at 65°C.

Non ligated adapters and ligation buffer were eliminated by magnetic sorting: magnetic beads (SPRI Select, Beckman Coulter) were added to the samples with a ratio of 1.2. After a 5min incubation at room temperature, samples were placed on a magnetic rack for 5min and the supernatant was discarded. Samples underwent two washes of 30s in 80% ethanol freshly prepared. Samples were removed from the magnetic rack and pellets were resuspended in a repair solution (2μL 10X Klenow Buffer, 0.5μM dNTP, 5U Klenow Fragment, H2O to 20μL, Thermo Scientific). Repair was performed according to the following program: 15 min at 37°C, 10 min at 75°C.

Repair buffer was eliminated by magnetic sorting: PEG NaCl solution (Swift Biosciences) was added to the samples with a ratio of 1.2. After a 5min incubation at room temperature, samples were placed on a magnetic rack for 5min, then the supernatant was discarded. Samples underwent two washes: 30s in freshly prepared 80% ethanol. Sample pellets were then dried at room temperature, removed from the magnetic rack and resuspended in 20 μL TE Low EDTA (Tris Base 10mM, EDTA 0.1mM, Swift Biosciences). Samples were incubated 5 min at room temperature then placed on the magnetic rack for 5min. Supernatants were collected in new tubes.

Nested qPCR – Quantification of ULD and 2LTR circles by qPCR

5μL of repaired ligation product was added to 45μL of PCR mix containing 0,5μL of each primer (ULD OUT Rev/ULD OUT Fwd or 2LTR OUT Rev/2LTR OUT Fwd and CD3 OUT Rev/CD3 OUT Fwd; sequences in Table S1) at 100μM, and 2X Phusion HF Master Mix (Thermo Scientific), H_2O up to 45μL. The PCR was performed according to the following program: 30 s at 98°C then 21 cycles of 30 s at 98°C, 30 s at 58°C, and 30 s at 72°C.

PCR product was diluted (1:100), and 5μL of diluted PCR product was added to 15μL of qPCR mix containing 1μL of each primer (ULD IN Rev/ULD IN Fwd or 2LTR IN Rev/2LTR IN Fwd and CD3 IN Rev/CD3 IN Fwd) at 100μM, 0,2μL of each probe (ULD Probe or 2LTR probe and CD3 Probe) at 10μM, and 2X Probe Master Mix (Roche), H_2O up to 15μL. The qPCR was performed according to the following program: 10 min at 95°C, then 45 cycles of 15 s at 95°C and 1 min at 60°C. Fluorescence was read after each cycle. Plasmids harboring a single copy of CD3γ and ULD amplicons were used to generate standard curves.

Nested PCR

20μL of repaired ligation product was added to 30μL of PCR mix containing 0,5μL of each primer (ULD OUT Rev, ULD OUT Fwd; sequences in Table S1) at 100μM, and 2X Phusion HF Master Mix (Thermo Scientific), H_2O to 30μL. The PCR was performed according to the following program: 30 s at 98°C then 35 cycles of 30 s at 98°C, 30 s at 58°C and 30 s at 72°C. PCR buffer and excess primers were eliminated by magnetic sorting following the protocol explained in the "ligation" section using a beads ratio of 0.8.

PCR product (20μL) was added to 25μL of 2X Phusion Master Mix (Thermo Scientific). Each sample received its own pair of primer (2,5μL of D50x and 2,5μL D7xx; sequences in Table S1), at 2μM. The PCR was performed according to the following program: 30 s at 98°C then 16 cycles of 30 s at 98°C, 30 s at 58°C and 30 s at 72°C.

PCR buffer and excess primers were eliminated by magnetic sorting following the protocol explained in the "ligation" section using a beads ratio of 0.8.

Library pool preparation and sequencing

The concentration of each library sample was determined using the KAPA Library Quantification Kit (Roche Sequencing Solutions). Each library sample was diluted to a concentration of 2nM and 5μL of each dilution was added to the library pool. The pool underwent a quality check using a Bioanalyzer High Sensitivity DNA Analysis (Agilent Technologies).

Library pool was loaded on a MiSeq Reagent Kit v2 (300-cycles, Illumina) and sequenced in a double 150 cycle paired-end run. Collected reads were exported in Fastq format after processing with the Aozan software[41] and analyzed using the Galaxy Europe server.[38]

QUANTIFICATION AND STATISTICAL ANALYSIS

All data are presented as mean ± SD except on Figure 5 where they are presented as mean ± SEM of n replicated experiment as indicated in the figure legends. Unpaired two-tail t test, linear regression and one phase decay non-linear regression were performed

using Graphpad Prism 6 Software. Linear regression is represented as a dotted line with its 95% confidence interval; R^2 and p of the regression are indicated on the figure. For Unpaired two-tail t test p value is indicated on the figure. One phase decay non-linear regression are indicated as dotted line on the figure; half-life ($t_{1/2}$), goodness of the fit (r^2) and 95% confidence interval (IC95%) are indicated on the figure.