COMMISSIE
COGEM
GENETISCHE
MODIFICATIE

ONDERZOEKSRAPPORT
2026-03

RECOMBINATION BETWEEN HIV AND LENTIVIRAL VECTORS

AN EXPERIMENTAL STUDY ON RECOMBINATION BETWEEN HIV AND HIV-DERIVED LENTIVIRAL VECTORS

ONAFHANKELIJKE WETENSCHAPPELIJKE ADVISERING EN SIGNALERING VAN DE RIJKSOVERHEID

Recombination between HIV and Lentiviral Vectors

An experimental study on recombination between HIV and HIV-derived lentiviral vectors

COGEM Report

November 13, 2025

Atze T. Das, Bep Klaver, Ben Berkhout

Amsterdam UMC, University of Amsterdam, Laboratory of Experimental Virology, Medical Microbiology and Infection prevention, Meibergdreef 9, Amsterdam, Netherlands

Project Guidance Committee
Rob Hoeben (COGEM - chair)
Nancy Beerens (COGEM)
Danny Goovaerts (COGEM)
Barbara te Riet-Schulte (Dutch GMO office; BGGO)
Irene van Grinsven (COGEM secretariat)
Clara Posthuma (COGEM secretariat)

Cover illustration
Abstract representation of the genome organization of HIV-1. Created by Irene van Grinsven.

Disclaimer
This report was commissioned by COGEM. The content of this publication is the sole responsibility of the authors and does not necessarily reflect the views of COGEM.

Dit rapport is in opdracht van de Commissie Genetische Modificatie (COGEM) samengesteld. De mening die in het rapport wordt weergegeven is die van de auteurs en weerspiegelt niet noodzakelijkerwijs de mening van de COGEM.

Table of Contents

1. VOORWOORD

Lentivirus-afgeleide vectoren (LVs) worden veel gebruikt voor de permanente genetische modificatie van zoogdiercellen in laboratoriumonderzoek. Daarnaast zijn ze een belangrijk hulpmiddel in gentherapie: ze maken het mogelijk om genen te introduceren in cellen van patiënten, waarna deze genetisch aangepaste cellen weer teruggeplaatst worden bij de patiënt. Met dergelijke gentherapie-procedures hoopt men de schadelijke effecten van de genetische aandoeningen teniet te doen.

Meestal zijn deze vectoren afgeleid van het humane immunodeficiëntie virus type 1 (HIV-1). Alle virusgenen die voor eiwitten coderen zijn uit deze vectoren verwijderd, waardoor de LV's zich niet zelfstandig kunnen repliceren of verspreiden na integratie in het DNA van een gastheercel. Vaak wordt gebruik gemaakt van LV's waarin ook een groot deel van de 'enhancer' en promotor is verwijderd. Deze zogenaamde 'self-inactivating' lentivectoren (SIN-LV's) schakelen zichzelf uit na integratie in het gastheer-DNA, zodat alleen het aangebrachte transgen in de vector met eigen promotor tot expressie kan komen. LV's bevatten vaak ook 'posttranscriptional regulatory elements' (PRE's), die het aflezen en verwerken van genen efficiënter laten verlopen en zo de expressie van het transgen verhogen.

Deze SIN-LV's goed zijn ingeperkt en kunnen niet zelfstandig repliceren. Er bestaat echter een kans dat een met een LV gemodificeerde cel wordt geïnfecteerd door een wildtype HIV, waarvan de eiwitten de replicatie van de LV enigszins kunnen faciliteren ('mobilisatie'). Dit risico is vooral aanwezig voor vectoren zonder deletie van de HIV-1 'enhancer' en promotor (de niet-SIN-LV's); bij SIN-LV's treedt deze mobilisatie nauwelijks op. Wanneer HIV-1 en een LV in tegelijk in één cel aanwezig zijn, kan hiertussen theoretisch ook recombinatie optreden. Over de aard en frequentie hiervan is in de wetenschappelijke literatuur erg weinig informatie te vinden. Om deze kennislacune te verkleinen heeft de COGEM laboratoriumonderzoek laten uitvoeren naar dit fenomeen.

Het onderzoek is uitgevoerd door Atze T. Das, Bep Klaver en Ben Berkhout van de afdeling Experimentele Virologie van het Amsterdam UMC. In hun rapport beschrijven zij het experimenteel onderzoek dat gericht was op het vaststellen van de aard en de frequentie van dergelijke recombinatie tussen HIV en de vector. Hierbij werden condities gebruikt die de frequentie van het optreden van dergelijke recombinaties zouden moeten vergroten. Hun resultaten laten zien dat zelfs na langdurig kweken geen recombinatie werd gevonden tussen SIN-LV's en HIV1.

Bij HIV-infectie van cellen met niet-SIN-LV's werden wel recombinanten gevonden; hierbij werden in alle gevallen alleen homologe gebieden uitgewisseld. Er werden geen recombinanten aangetroffen waarbij (delen van) het transgen, de heterologe promotor, of de PRE-sequenties uit de vector in het replicerende HIV-1-genoom werden opgenomen. Deze waardevolle inzichten helpen bij het uitvoeren van gefundeerde risico-evaluaties bij het gebruik van LV's in klinische studies en laboratoria, en onderstrepen het belang van het maken van onderscheid tussen SIN-LV's en niet-SIN-LV's hierbij.

De uitvoerders van het projectzijn bijzonder zorgvuldig te werk gegaan gaan en hebben – met veel toewijding – aanvullende analyses gedaan, om een zo volledig mogelijk beeld te geven van de potentiële risico's en mechanismen. Deze grondige en brede aanpak heeft geleid tot een rapport dat van grote waarde is voor het veld.

Prof. Dr. Rob Hoeben, voorzitter van de begeleidingscommissie

2. SAMENVATTING

Lentivirale vectoren (LV's), afgeleid van het humaan immunodeficiëntievirus (HIV), worden veel gebruikt voor het genetisch modificeren van cellen, zowel in laboratoriumonderzoek als in klinische toepassingen. Net als HIV-virusdeeltjes bevatten LV-deeltjes een RNA-genoom dat, eenmaal de cel binnengedrongen, wordt omgezet in DNA. Dit DNA wordt vervolgens geïntegreerd in het erfelijk materiaal van de cel, wat leidt tot een permanente genetische verandering.

Het LV-RNA bevat alle regulatie-elementen afkomstig van HIV die nodig zijn voor dit transductieproces. De HIV-genen die coderen voor viruseiwitten zijn echter grotendeels verwijderd en vervangen door een zogenoemd 'gene-of-interest' (GOI), meestal een gen dat codeert voor een eiwit dat men in de cel wil laten produceren. Dit GOI wordt vaak gecombineerd met een promotor en andere regulatie-elementen, niet afkomstig van HIV, die zorgen voor optimale expressie. Hieronder valt ook een posttranscriptioneel regulatie-element (PRE) afkomstig van een hepatitisvirus. In klinische toepassingen codeert het GOI meestal voor een therapeutisch eiwit dat – zolang de LV-behandelde cellen in het lichaam aanwezig zijn – helpt bij het bestrijden of genezen van een ziekte.

Voor de productie van LV-deeltjes wordt het construct dat codeert voor het LV-RNA tijdelijk gecombineerd met zogenoemde 'verpakkingsconstructen' die coderen voor de benodigde virale eiwitten. Omdat de vector zelf deze eiwitten niet kan aanmaken, kunnen LV's niet zelfstandig repliceren of zich verspreiden naar andere cellen.

Het gebruik van LV's in laboratoriumonderzoek en klinische toepassingen valt onder de regelgeving voor genetisch gemodificeerde organismen (GGO's) en vereist een zorgvuldige risicobeoordeling. Een potentieel risico dat hierbij wordt meegenomen, is de mogelijke vorming van replicatie-competente virussen door recombinatie tussen het LV-RNA-construct en de verpakkingsconstructen tijdens productie van de LV-deeltjes. Moderne verpakkingsconstructen zijn echter zo ontworpen dat de kans hierop minimaal is. Daarnaast worden tegenwoordig vaak zogenoemde zelf-inactiverende ('self-inactivating'; SIN) LV's gebruikt, waarin het HIV promotorfragment - dat essentieel is voor de productie van HIV-RNAs in de viruscontext - grotendeels is verwijderd. Deze SIN deletie verkleint de kans op de ongewenste vorming van een replicerend virus verder.

Een andere overweging in de risicobeoordeling is de mogelijke aanwezigheid van het HIV-virus in de te modificeren cellen, bijvoorbeeld wanneer de cellen afkomstig zijn van HIV-geïnfecteerde personen. De door HIV geproduceerde eiwitten maken de vorming van virusdeeltjes met een LV-RNA-genoom mogelijk, waardoor LV's zich - samen met HIV - kunnen verspreiden naar andere cellen. SIN-vectoren produceren echter aanzienlijk minder LV-RNA, hetgeen mobilisatie en verspreiding van de LV sterk reduceert. Bij aanwezigheid van HIV kan er mogelijk ook recombinatie tussen HIV en de vector plaatsvinden. Hoewel dit nog nooit is beschreven, zou dit kunnen leiden tot nieuwe HIV-varianten die genetische informatie (d.w.z. nucleotidesequenties) van LV's bevatten en welke zich mogelijk anders gedragen, bijvoorbeeld wat betreft pathogeniteit of gastheerspecificiteit.

Om de kans op het ontstaan van dergelijke recombinante HIV-varianten, en het bijbehorende risico, beter in te kunnen schatten, hebben wij in dit experimentele onderzoek getest of recombinatie tussen HIV en LV's kan leiden tot replicatie-competente HIV-varianten die LV-nucleotidesequenties bevatten. Hiertoe hebben we verschillende HIV-varianten langdurig gekweekt in humane cellen die waren behandeld met een SIN- of niet-SIN-vector, waarna we het virale genoom hebben geanalyseerd. Bij gebruik van niet-SIN-vectoren zagen

we frequent recombinatie tussen HIV en de vector, waarbij HIV sequenties werden vervangen door corresponderende sequenties uit de LV. Deze recombinatie vond vooral plaats in de HIV regulatie-elementen, maar ook in een eiwit-coderende sequentie. Er werd geen recombinatie gevonden waarbij LV-specifieke onderdelen, zoals het PRE-element, in HIV werden opgenomen. Bij gebruik van SIN-vectoren werd helemaal geen recombinatie waargenomen.

Onze bevindingen, verkregen onder experimentele omstandigheden die erop gericht waren de kans op recombinatie te maximaliseren, geven aan dat recombinatie tussen HIV en LV's uitsluitend optreedt bij het gebruik van niet-SIN-vectoren. Daarbij dient opgemerkt te worden dat de recombinatie die we hebben gedetecteerd in de viruskweekexperimenten uitsluitend uitwisseling betrof van vergelijkbare sequenties tussen HIV en LV's. Theoretisch kan HIV-LV-recombinatie ook leiden tot de insertie van LV-specifieke elementen zoals de PRE, GOI of GOI-promotor. Door ons geconstrueerde HIV-varianten met een PRE-element lieten echter een verminderde replicatie zien. Daarnaast zullen de meeste GOI's in LV's, zeker in klinische toepassingen, niet coderen voor factoren die HIV-replicatie bevorderen en is recombinatie tussen de HIV-promotor en GOI-promotoren veelal niet waarschijnlijk vanwege een beperkte nucleotidesequentie-overeenkomst.

Hoewel de kans klein lijkt dat HIV-LV recombinatie resulteert in een meer pathogeen virus of een virus met ander gastheerbereik, kan dit – bij gebruik van niet-SIN LV's – echter niet op voorhand worden uitgesloten. Daarom is voor de risicobeoordeling van laboratoriumonderzoek en klinische toepassingen waarbij cellen worden gemodificeerd met behulp van niet-SIN-LV's, de potentiële aanwezigheid van HIV in de cellen niet alleen relevant vanwege de mogelijke mobilisatie en verspreiding van de vector, maar ook vanwege de kans op HIV-LV recombinatie. Aangezien de SIN-aanpassing mobilisatie en verspreiding van de LV sterk reduceert en HIV-LV recombinatie voorkomt, lijkt de potentiële aanwezigheid van HIV bij het gebruik van SIN-LV's minder relevant voor de risicobeoordeling.

Tot slot blijft is de aanwezigheid van HIV altijd relevant voor de Arboveiligheid van werkzaamheden, aangezien HIV een pathogeen virus is dat AIDS veroorzaakt indien het niet voldoende wordt onderdrukt door middel van antiretrovirale therapie.

Belang

Lentivirale vectoren afgeleid van het humaan immunodeficiëntievirus (HIV) worden veelvuldig gebruikt voor de genetische modificatie van cellen in zowel laboratoriumonderzoek als klinische toepassingen. Hun nauwe verwantschap met HIV roept echter vragen op over de bioveiligheid, waaronder de mogelijkheid van recombinatie tussen het HIV en deze vectoren. Hoewel dit nog nooit is beschreven, zou dergelijke recombinatie kunnen leiden tot de vorming van virusvarianten met ongewenste eigenschappen. Deze experimentele studie laat zien dat HIV kan recombineren met vectoren die een volledige LTR bevatten, maar niet met vectoren die een SIN-deletie in de LTR hebben. Deze bevindingen zijn van belang voor de risicobeoordeling van laboratoriumonderzoek en klinische toepassingen waarbij lentivirale vectoren worden gebruikt in situaties waarin het HIV aanwezig kan zijn.

3. SUMMARY

Lentiviral vectors (LVs) derived from human immunodeficiency virus type 1 (HIV-1) are widely used to genetically modify cells in both laboratory research and clinical applications. Similar to HIV-1 virions, LV particles carry an RNA genome that is reverse transcribed into DNA upon infection. This DNA integrates into the host cell's chromosomes, where it is stably maintained, making LVs a preferred tool for permanent genetic modification of cells.

The LV RNA contains all the HIV regulatory sequences required for this transduction process. Most HIV protein-coding regions are removed and replaced with a 'gene-of-interest' (GOI), along with promoter and other sequence elements that drive its expression, typically including a posttranscriptional regulatory element (PRE). In gene therapy applications, the GOI may encode a therapeutic protein, offering sustained effects as long as the LV-modified cells persist or proliferate in the patient. For the production of LV particles, a construct encoding the LV RNA is transiently combined with 'packaging' constructs that provide the structural and enzymatic proteins required for particle assembly and transduction. Because the LV RNA itself does not encode these proteins, the vector cannot replicate or spread beyond the initially transduced cells.

Use of LVs in laboratory research and clinical applications requires thorough risk assessment under GMO regulations. A major concern is the potential generation of a replication-competent virus through recombination between the production constructs and the vector. However, modern packaging systems are designed to minimize this risk. Moreover, the use of 'self-inactivating' (SIN) lentiviral vectors - in which the HIV-derived sequences that drive viral RNA production in the HIV context are largely deleted - further reduces the likelihood of replication-competent virus formation.

Another key consideration is the presence of HIV in transduced cells. HIV proteins can enable mobilization of the LV, potentially allowing its unintended spread. The SIN deletion, which markedly reduces vector RNA production, inhibits such mobilization. Simultaneous presence of HIV and the LV also raises the possibility of recombination, potentially generating HIV variants containing LV sequences. Such variants might exhibit altered replication, pathogenicity, or host range. However, experimental or clinical data either demonstrating such recombination or refuting its possibility are lacking.

In this experimental study, we investigated whether recombination between HIV-derived LVs and HIV can generate replication-competent HIV variants containing LV-derived sequences. We cultured a highly replicative HIV strain and attenuated derivatives for up to three months on T-cells transduced with SIN and non-SIN LVs. HIV genome analysis revealed frequent HIV-LV recombination when non-SIN vectors were used, predominantly involving exchange of shared regulatory sequences, but recombination involving shared protein-coding sequences was also observed. We did not detect any recombination leading to the insertion of the PRE or other LV-specific fragments into the HIV genome. In contrast to the frequent recombination observed with non-SIN LVs, no recombination events were detected when the cells were transduced with a SIN vector.

These results - generated under conditions designed to maximize the likelihood of recombination - demonstrate that HIV-LV recombination occurs exclusively with non-SIN vectors. The observed recombination events involved only the exchange of similar sequences between HIV and LVs. Theoretically, HIV-LV recombination could also lead to the insertion of LV-specific elements, such as the PRE, GOI, or the GOI promoter. However, HIV variants we constructed with a PRE element showed reduced replication. Additionally, most GOIs in LVs -

particularly in clinical applications - are unlikely to encode factors that promote HIV replication, and recombination between the HIV promoter and the GOI promoter is generally improbable due to limited nucleotide sequence similarity.

Although the probability seems low that HIV-LV recombination would result in a more pathogenic virus or one with a broader host range, this cannot be completely ruled out when using non-SIN LVs. Therefore, for the risk assessment of both laboratory research and clinical applications involving the use of non-SIN LVs, the presence of HIV is relevant not only because of the potential mobilization and spread of the vector, but also due to the possibility of HIV-LV recombination. Since the SIN modification significantly reduces mobilization and spread of the vector, and prevents HIV-LV recombination, the presence of HIV seems less relevant for the risk assessment when using SIN-LVs.

It should be noted that the presence of HIV is always a critical consideration for occupational health and safety, as HIV is a pathogenic virus that can lead to AIDS if not effectively controlled through antiretroviral therapy.

Importance

Lentiviral vectors derived from the human immunodeficiency virus (HIV) are frequently used for the genetic modification of cells in both laboratory research and clinical applications. However, their close relationship with HIV raises biosafety concerns, including the potential for recombination between the HIV virus and these vectors, which could lead to the formation of viral variants with undesirable properties. This study demonstrates that HIV can recombine with vectors containing a complete LTR, but not with vectors that have a SIN deletion in the LTR. These findings are relevant for the risk assessment of laboratory research and clinical applications involving lentiviral vectors in settings where HIV may be present.

4. INTRODUCTION

Lentiviral vectors (LVs) derived from human immunodeficiency virus type-1 (HIV) are widely employed to genetically modify cells in both laboratory research and clinical applications. Several LV-based gene therapy products have already received regulatory approval for human use. These therapies primarily involve genetically modified T cells (e.g., CAR T cells) and hematopoietic stem cells, which are transduced with the LV *ex vivo*, following cell isolation, and subsequently (re)introduced into the patient.

A unique and potentially concerning scenario could arise when HIV-infected cells are transduced with an LV, or conversely, when LV-transduced cells are later infected with HIV. In such cases, recombination between the vector and the virus might occur, which could theoretically lead to the generation of HIV variants with altered phenotypic characteristics, such as changes in pathogenicity or tropism.

The aim of this research project is to investigate, through *in vitro* experimentation, whether recombination between an HIV-derived LV and HIV itself is feasible, and whether such recombination could give rise to replication-competent HIV variants that incorporate LV-derived genetic elements.

Lentiviral Vectors and HIV

Lentiviral vectors (LVs) are predominantly derived from human immunodeficiency virus type-1 (HIV-1), the etiological agent of acquired immunodeficiency syndrome (AIDS). HIV-1 is a retrovirus belonging to the genus Lentivirus and is characterized by the presence of two copies of a ~9-kilobase single-stranded RNA genome within its virion. Upon infection, this RNA genome is reverse transcribed into double-stranded DNA, which is subsequently integrated into a host cell chromosome, where it remains stably maintained. Transcription of this integrated viral DNA, the provirus, leads to the production of new viral particles capable of infecting additional host cells. HIV primarily replicates in human $CD4^+$ T lymphocytes, and transmission occurs via unprotected sexual contact, vertical transmission (mother-to-child), blood transfusion, and organ transplantation.

Although combination antiretroviral therapy (ART) - which involves the concurrent administration of several antiviral agents - can effectively suppress HIV replication and prevent progression to AIDS, it does not eliminate the virus. HIV persists as integrated proviral DNA within latently infected cells, forming a long-lived viral reservoir. If ART is discontinued, viral replication resumes, potentially resulting in disease progression. As a result, individuals living with HIV require lifelong ART to maintain viral suppression.

The HIV proviral DNA (Figure 1) encodes three essential (gag, pol and env), two regulatory (tat and rev) and four accessory (vif, vpr, vpu, nef) protein-coding genes. This proviral genome is flanked by long terminal repeat (LTR) sequences, each comprising three domains: U3 (unique to the 3′ end of the viral RNA), R (a repeated region present at both ends), and U5 (unique to the 5′ end of the viral RNA). The LTRs contain critical regulatory elements for viral transcription and RNA processing. For example, the U3 region includes binding sites for transcription factors such as NF-κB and Sp1, while the R region contains the TAR element and the polyadenylation signal. The 5'LTR functions as promoter for viral RNA transcription, which is initiated at the beginning of the R region. This transcription is significantly

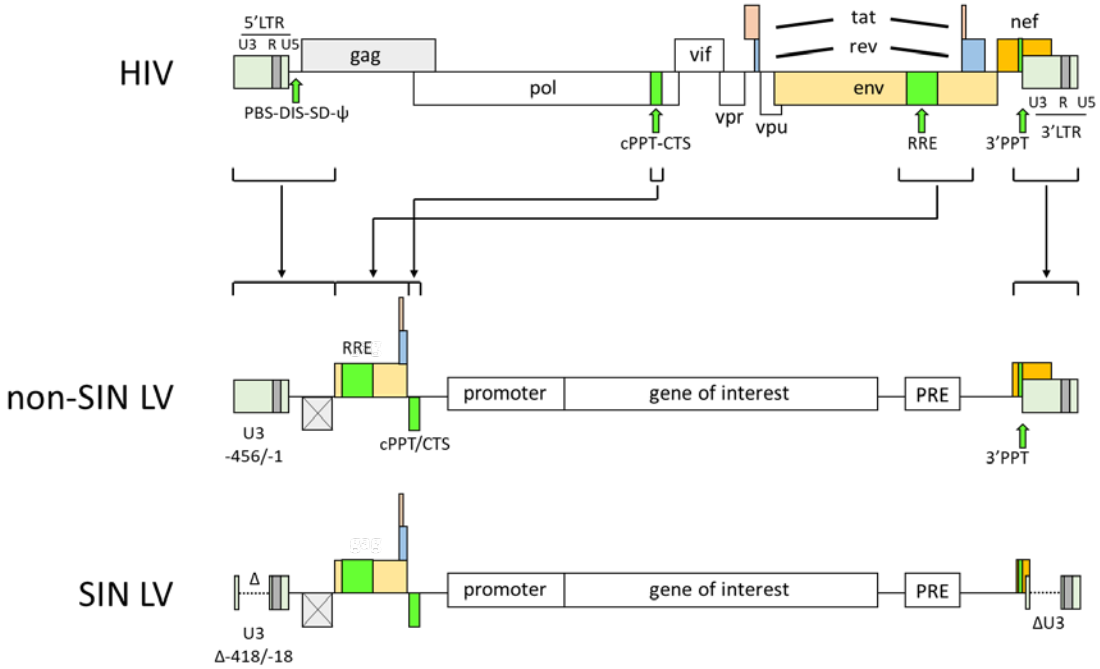


Figure 1. Schematic of HIV-1 and lentiviral vector (LV) proviral DNAs. The structures shown represent the proviral DNAs upon completion of reverse transcription and integration into the host genome. The 5′ and 3′ long terminal repeat (LTR) elements with U3, R and U5 domains are indicated. The origin of the HIV genomic segments retained in the LVs is shown. In non-self-inactivating (non-SIN) LVs, the 5′ and 3′ LTR are intact. The self-inactivating (SIN) LVs contain a deletion in the U3 region, spanning nucleotides −418 to −18 relative to the transcription start site. PBS, primer binding site; DIS, dimerization initiation site; SD, major splice donor site; ψ, packaging signal; cPPT, central polypurine tract; CTS, central termination sequence; 3'PPT, 3' polypurine tract; RRE, Rev-response element; LTR, long terminal repeat; U3, unique to the 3′ end of the viral RNA; R, repeat region; U5, unique to the 5′ end of the viral RNA; PRE, post-transcriptional regulatory element. The gag sequence present in the LVs contains a G to GCG frameshift mutation, which causes premature termination of translation (crossed gray box).

enhanced by the binding of the viral Tat protein to the TAR hairpin element in the nascent RNA transcripts. The transcripts are polyadenylated at the end of the R region in the 3'LTR.

HIV-derived LVs (Figure 1) contain the 5' and 3' LTR sequences, but the HIV sequences between the LTRs have been largely removed, except for (1) the primer-binding site (PBS) downstream of the 5'LTR, which forms the binding site for the cellular $tRNA^{lys3}$ that functions as primer for reverse transcription, (2) the RNA dimerization, packaging and splicing signals in the leader region, (3) an env fragment including the Rev-responsive element, which allows Rev-mediated export of unspliced and incompletely spliced viral RNAs from the nucleus to the cytoplasm, (4) the central and 3' polypurine tracts (cPPT and 3'PPT), and the central termination site (CTS), which have important priming and termination functions in the reverse transcription process. LVs have either a truncated U3 region, in which transcription-factor binding sites have been largely deleted to inactivate the LTR promoter activity (self-inactivating [SIN] LVs)(5), or a complete U3 region (non-SIN LVs). In these non-SIN vectors, like in HIV, the 5'LTR has transcriptional promoter activity, which can be enhanced by the viral Tat protein (if supplied in trans).

In addition to the minimized HIV sequences, lentiviral transfer vectors typically include a therapeutic gene-of-interest (GOI), driven by a non-HIV promoter, and a post-transcriptional regulatory element (PRE), frequently derived from woodchuck hepatitis virus (WPRE) or human hepatitis B virus (HPRE), which is included to increase expression of the GOI. An

essential feature of lentiviral vectors is their capacity to integrate into the genomic DNA of target cells, thereby ensuring their persistent presence. When the LV-transduced cells exhibit prolonged survival or retain proliferative potential following administration in humans, the intended gene therapy may yield sustained therapeutic effects.

Risk Assessment of Lentiviral Vector Use

The use of HIV-derived LVs in clinical and research settings necessitates a thorough risk assessment. Key considerations in such assessments include:
(1) generation of replication-competent lentivirus (RCL) through recombination of vector components during the production process, which could pose a biosafety risk,
(2) mobilization of the LV upon infection with HIV, potentially leading to unintended dissemination and spread of the vector within the host or to others, and
(3) recombination between HIV and the LV, which may result in the emergence of recombinant viruses with altered biological properties, such as modified tropism or pathogenicity.

RCL formation. LV's based on HIV can be categorized based on the used packaging system and the transfer vector. Modern LV systems typically employ second- or third-generation packaging constructs in combination with a SIN transfer vector, or a translentiviral packaging system in combination with a non-SIN transfer vector. These configurations are designed to minimize the risk of RCL formation, by removing non-essential viral sequences, splitting essential packaging components across multiple plasmids, and reducing the sequence similarity between these plasmids. Extensive empirical data from clinical studies have consistently shown no evidence of RCL emergence when third-generation packaging constructs are used in combination with a SIN vector (6, 7).

Mobilization. Mobilization of the LV transfer vector may occur when transduced cells become infected with HIV (8). Upon infection, HIV drives the production of all viral RNAs and proteins required for the assembly of new infectious particles, including the viral Tat protein, which strongly enhances transcription from the HIV 5'LTR promoter. In the case of non-SIN LVs, Tat can also activate transcription from the vector's 5'LTR, leading to the production of full-length LV RNAs. Although SIN LVs contain a deletion in the U3 region of the LTR that significantly reduces transcriptional activity (5), low-level LV RNA production cannot be entirely excluded (9, 10).

Because the LV RNAs contain all necessary sequence motifs for dimerization and packaging, newly assembled virions may contain either an HIV RNA homodimer, an LV RNA homodimer, or an HIV-LV RNA heterodimer. These virions can mediate the spread of both HIV and the LV to other cells. Moreover, as discussed in more detail below, infection with virions containing an HIV-LV heterodimer provides an opportunity for recombination between the HIV genome and the LV vector.

Recombination. Phylogenetic analyses of circulating HIV variants indicate frequent recombination during virus evolution. This phenomenon has also been extensively observed in *in vitro* cell culture experiments (4, 11, 12). When cells are co-infected with two different HIV variants, recombination can occur if their RNA genomes form heterodimers and are co-packaged into the same viral particles. Upon infection of a new target cell, both RNA genomes serve as template for reverse transcription. During this process, template switching—facilitated by homologous sequences shared between the two RNA templates—results in recombination.

Schlub et al. estimated an average of approximately 12.5 recombination events per genome per replication cycle (13).

When cells are transduced with an LV and infected with HIV, recombination between the virus and the vector could occur if LV-HIV RNAs heterodimers are formed and packaged into viral particles. Upon subsequent infection, both HIV and LV RNAs will be used as template for reverse transcription, and template switching at homologous sequences may lead to recombination. Because HIV RT lacks proofreading activity, recombination could (theoretically) also result in a virus variant incorporating other LV sequences, such as GOI or PRE sequences. If this occurs, HIV-LV recombination could generate HIV variants with distinct phenotypic characteristics, such as increased pathogenicity or altered tropism.

Goal of this study

In this study, we aimed at experimentally investigating whether HIV-LV recombination is possible when cells are both transduced with an LV and infected with HIV. We combined both SIN and non-SIN LVs with a highly replicative HIV-1 strain, as well as with genetically modified, attenuated variants for which we hypothesized that HIV-LV recombination might enhance viral replication. Furthermore, to assess whether PRE sequences can improve HIV replication, we constructed HIV variants containing different, commonly used PREs and analyzed their replication capacity and genetic stability.

5. RESULTS

Recombination between HIV and LV

To investigate whether HIV can recombine with an LV, SupT1 T cells - which support efficient replication of HIV-1 and are commonly used in HIV culture studies - were transduced with either SIN or non-SIN LVs, and subsequently infected with different HIV-1 variants. We included several fully replication-competent HIV-1 variants, for which complete ‘molecular clone’ plasmids are available (Figure 2). **HIV-LAI** is an efficiently replicating HIV-1 subtype B variant, based on the LAI isolate, which expresses all HIV proteins and is able to induce large syncytia in SupT1 T cells (1). **HIV-HBT** is a recently developed HIV reporter virus based on HIV-LAI, in which the short HiBiT luciferase tag is fused to integrase (3). **HIV-CD** is an HIV-LAI variant with several nucleotide substitutions in the polyA hairpin sequence in the R-U5 region (14). These mutations change the nucleotide sequence in the stem region of the RNA hairpin structure, but do not significantly affect its thermodynamic stability. Replication of HIV-HBT and HIV-CD is slightly attenuated when compared to the wild-type HIV-LAI. **HIV-rtTA** is a doxycycline (dox)-controlled HIV-LAI derivate (15, 16). In this virus, the Tat-TAR transcription activation mechanism is inactivated, and functionally replaced by the Tet-On system for dox-inducible gene expression. For the construction of this virus, inactivating mutations were introduced in both the Tat gene and the TAR sequence in the R region, the gene encoding the dox-inducible ‘reverse tetracycline transcriptional activator’ protein (rtTA) was inserted at the site of the nef gene, and tet operator (tetO) rtTA-binding sites were inserted in the U3 promoter region. This virus replicates exclusively in the presence of dox. The HIV-rtTA variant used in this study, HIV-rtTA$^{G19F\ E37L\ P56K}$, has several mutations in the rtTA gene, which prevent the loss of dox control in long term cultures, but also slightly attenuate virus replication (17).

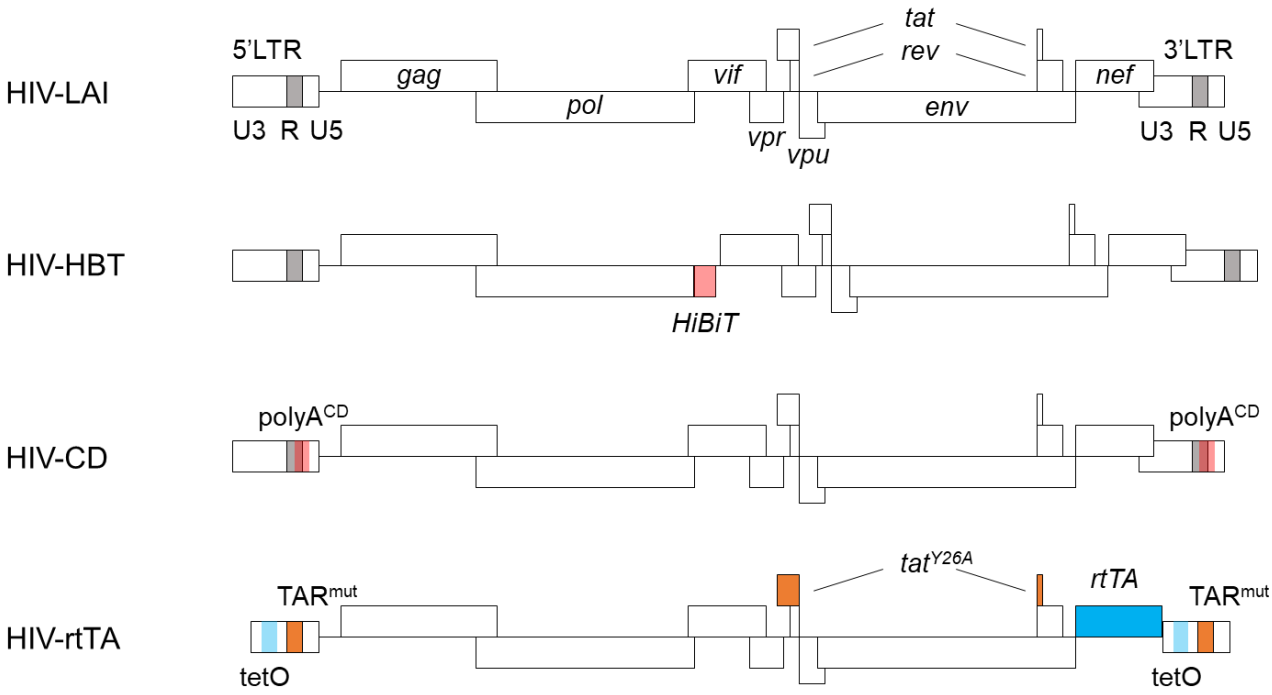


Figure 2. HIV variants. The structures shown represent the proviral DNAs upon completion of reverse transcription and integration into the host genome. All variants are based on the HIV-LAI molecular clone (1). In HIV-HBT, a HiBiT tag is fused to the integrase domain in pol (3). In HIV-CD, the polyA hairpin sequence in the R-U5 domains of the 5' and 3' LTR is mutated. In HIV-rtTA, the Tat-TAR axis of transcription activation is functionally replaced by the Tet-On system for doxycycline-inducible gene expression by [1] a Y26A amino acid substitution in Tat, [2] several nucleotide substitutions in the TAR sequence in the R domain of the 5' and 3' LTR, [3] insertion of the reverse tetracycline controlled transcriptional activator (rtTA) gene at the site of nef, and [4] insertion of rtTA-binding tet operator (tetO) elements in the U3 domain of the 5' and 3' LTR.

We selected several LVs for transduction of the SupT1 cells (LV1 to LV5 in Figure 3A). LV1 is a non-SIN LV with a complete U3 promoter region, whereas LV2, LV3, LV4 and LV5 are SIN LVs in which the U3 region was largely deleted. Except for this SIN/non-SIN variation, the LVs contain nearly identical HIV-derived sequences (Figure 3B), which are similarly configurated. The LVs have different promoter/GOI cassettes and PREs, but all encode a fluorescent reporter protein (either EGFP or mCherry). LV1-LV4 are frequently requested LVs from Addgene. LV5 contains a dox-inducible gene expression cassette, with the rtTA gene encoding the optimized rtTA-V16 protein that is more active and more dox-sensitive than the rtTA variant present in HIV-rtTA (18).

Sequence alignment (Figure 3B) revealed a very high sequence similarity between HIV-LAI and the HIV-derived parts in the LVs, but several nucleotide differences are present. Sequence alignment with the HIV-HXB2 molecular clone (not shown) indicates that the LV sequences likely originate from this HIV variant, which was generated upon *in vitro* culturing of the LAI isolate. Importantly, the high sequence similarity between HIV-LAI and the LVs will be optimal for homology-driven recombination, while the nucleotide differences will allow the identification of HIV-LV recombination events.

LV particles produced in 293T cells by co-transfection of the LV transfer vector and packaging plasmids were used to transduce SupT1 cells. Expression of the fluorescent reporter proteins in these cells was analyzed at 8 days after transduction by flow cytometry

A

	lentiviral vector	LTR	GOI configuration	PRE		
				origin	size (nt)	X orf
LV1	pLVX-EGFP-IRES-puro	non-SIN	CMV-EGFP-IRES-Puro-WPRE	Woodchuck hepatitis virus	589	wt
LV2	pCDH-CMV-mCherry-T2A-Puro	SIN	CMV-mCherry-T2A-Puro-WPRE	Woodchuck hepatitis virus	589	mut
LV3	pRRL-SIN-cPPT-PGK-GFP-WPRE	SIN	PGK-EGFP-WPRE	Woodchuck hepatitis virus	589	wt
LV4	pLenti-CMV-GFP-Puro	SIN	CMV-EGFP-WPRE-PGK-Puro	Woodchuck hepatitis virus	589	wt
LV5	TRE-CMV-rtTA^{V16}	SIN	Ptet-EGFP-CMV-rtTA^{V16}-HPRE	Human hepatitis B virus	1137	wt

B

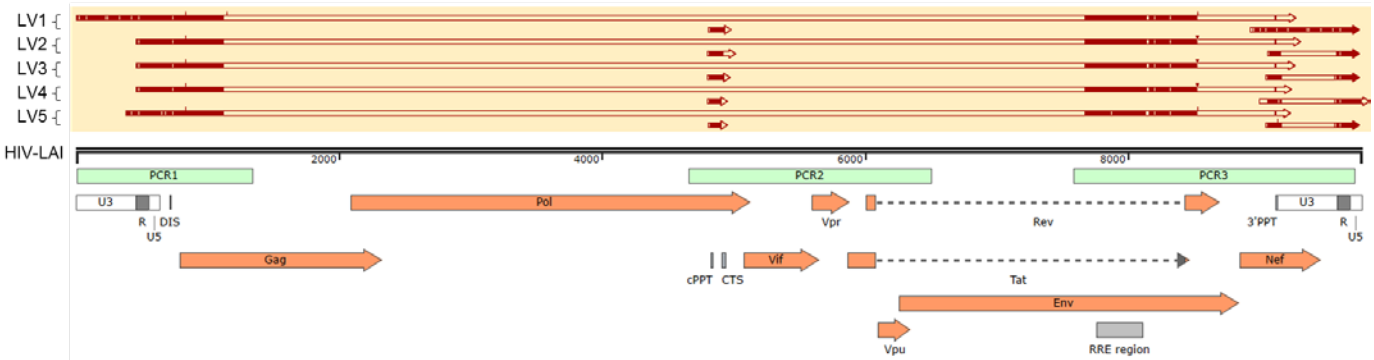


Figure 3. Lentiviral vectors. (A) Configuration of the LTR, GOI cassette and PRE element in the LVs used in this study. The wild-type PRE fragment of woodchuck hepatitis virus (WHV), WPRE^{wt}, contains a truncated protein X open reading frame, corresponding to the 60 N-terminal amino acids of the 142 amino acid WHV X protein. In the WPRE^{mut} element, the protein X translation start site is inactivated through several nucleotide substitutions (2). The PRE fragment of human hepatitis B virus (HBV), HPRE, encodes the nearly complete HBV protein X (147 of the 154 amino acids). CMV, cytomegalovirus promoter; PGK, phosphoglycerate kinase 1 promoter; Ptet, tetO promoter. EGFP, mCherry, d2EGFP: fluorescent proteins; rtTA^{V16}, optimized rtTA-V16; Puro, puromycin resistance; T2A, "self-cleaving" peptide (ribosomal skipping site); IRES, internal ribosome entry site. **(B)** The HIV-derived sequences in the LVs are aligned to the HIV-LAI genome. The positions of the genes and regulatory elements in HIV-LAI, and the regions analyzed by PCR and sequencing after culturing of HIV in LV-transduced cells (Figures 5 to 10), are indicated.

analysis. This analysis revealed that transduction with the undiluted LV particles resulted in 87 to 99% fluorescent protein-positive cells, whereas this level gradually decreased upon serial dilution of the stocks (Figure 4A). Similar analysis of the transduced cells at later times, revealed a similar transduction level, indicating that the transduced cells were not selectively lost upon prolonged culturing (Figure 4B).

The LV-transduced cells with the highest level of transduction, produced with the undiluted vector stocks, were infected with the different HIV variants (experimental set up shown in Figure 4C). For every HIV-LV combination, 3 cultures were started and maintained for approximately 3 months. Dox was continuously added to the HIV-rtTA infected cultures to activate replication of this dox-dependent virus. When large HIV-induced syncytia were observed, approximately once a week, the virus was passaged to fresh LV-transduced cells to prolong virus replication and the opportunity for recombination. To analyze the different virus cultures for HIV-LV recombination events, the virus was passaged twice to regular, non-transduced SupT1 cells, followed by the isolation of the intracellular DNA, which contains the integrated proviral DNA. This DNA was used as template for PCR amplification of 3 different proviral DNA regions (PCR1-3 in Figure 3). The PCR products were analyzed by agarose gel electrophoresis and Sanger sequencing. The PCR-amplified regions included all HIV sequences with similarity to the LVs, as these sequences are considered to be prone to homology-driven recombination. Furthermore, the accessory genes, which are essential for replication *in vivo* but less important for replication in T cell lines *in vitro*, were covered, as the virus may -theoretically - tolerate mutations in these regions. Our analysis did not include the

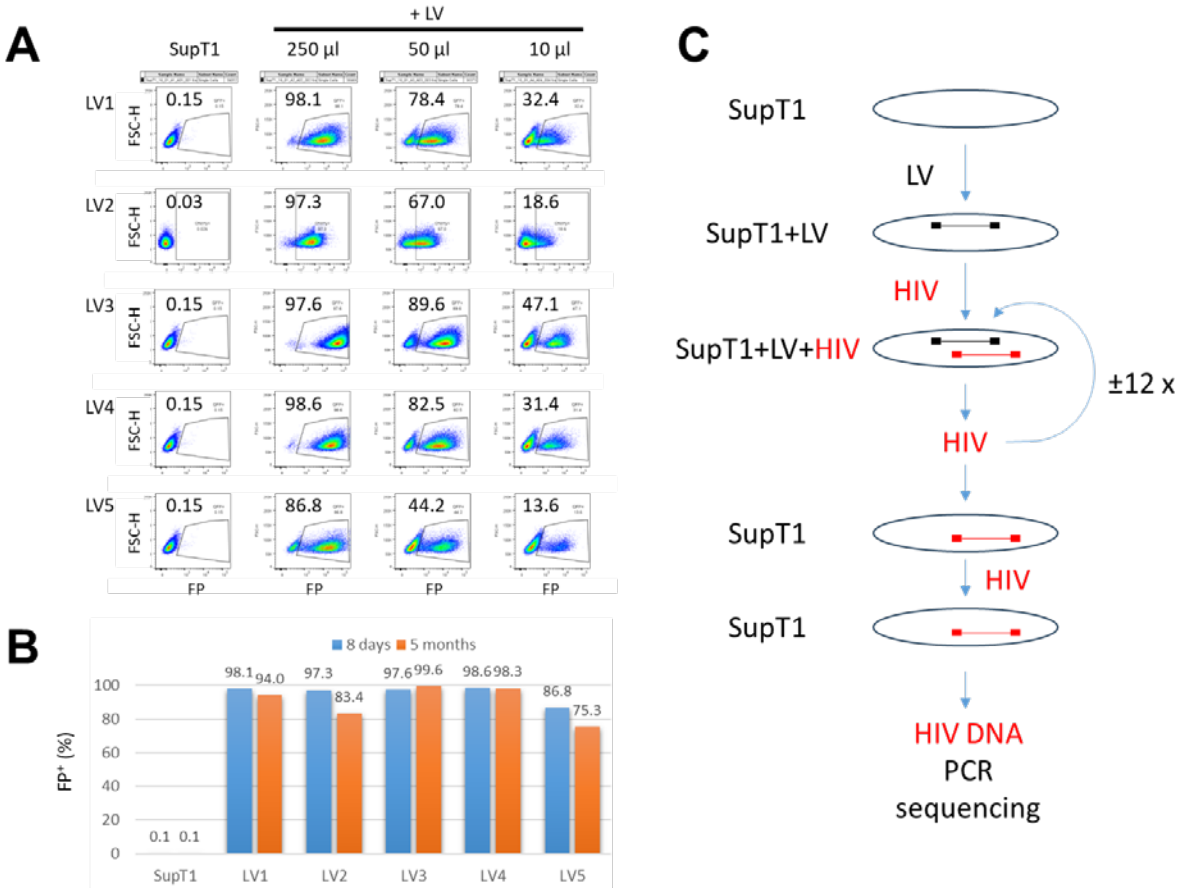


Figure 4. Transduction of SupT1 cells. (A) SupT1 cells (250 µl) were transduced with 10, 50 or 250 µl of the LV stocks. GFP (LV1, LV3, LV4, and LV5) or mCherry (LV2) expression was analyzed by flow cytometry at 8 days after transduction. The percentage of fluorescent protein-positive cells (FP^+) is indicated. **(B)** FP expression in SupT1 cells transduced with 250 µl LV at 8 days and 5 months after transduction. **(C)** Experimental set up. See text for details (LV DNA in black; HIV DNA in red).

central parts of the gag-pol and env genes, because the LVs lack similar sequences, and any insertion of LV sequences in these essential domains would likely inactivate, or at least significantly attenuate, virus replication.

Analysis of the PCR products by gel electrophoresis demonstrated that the size of nearly all PCR fragments was as expected for the different HIV variants (data not shown), which excluded the insertion of large LV-derived sequences. Several PCR3 products were perceptibly shorter than expected, and - as described below - sequence analysis of these fragments demonstrated that this was due to nucleotide deletions in the nef region, upstream of the 3'PPT and U3 domains. Deletions in the accessory nef gene are frequently detected upon long-term *in vitro* culturing of HIV in T-cell lines like SupT1, and likely result from slippage during the reverse transcription process, whereby the reverse transcriptase and newly formed DNA dissociate from the RNA template and re-anneal at an upstream site.

Sequence analysis of the HIV-LAI cultures (Figure 5) revealed that in 2 of the 3 LV1-transduced cultures the virus had acquired a T+23C mutation in the TAR motif, which is located in the R domain of the LTR. In the non-SIN vector LV1, a C nucleotide is present at this position, and the T+23C mutation in HIV-LAI therefore suggests that HIV-LV recombination occurred in this R region. Although a C nucleotide is also present at this TAR position in the SIN LVs, the virus did not acquire this mutation in any of the LV2 to LV5-transduced cell cultures. We

	PCR 1						PCR 2		PCR 3							
		U3[b]		TAR	leader	gag		CTS		RRE	env					nef-U3
	position[a]	-34	-8	+23	+281	+379		+4469		+7416	+7684	+7690	+7694	+7746	+7867	
	HIV-LAI	G	C	T	A	G	HIV-LAI	A	HIV-LAI	G	T	A	T	T	A	
	LV1	C	C	C	C	GCG[c]	LV1/3/5	A	LV1-4	G	T	C	C	C	T	several nt[d]
	LV2-5	Δ	T	C	C	GCG[c]	LV2/4	Δ	LV5	C	Δ	C	C	C	T	
LV1	1	G	C	T	A	G		A		G	T	A	T	T	A	
	2	G	C	C	A	G		A		G	T	A	T	T	A	
	3	G	C	C/T[e]	A	G		A		G	T	A	T	T	A	
LV2	1	G	C	T	A	G		A		G	T	A	T	T	A	
	2	G	C	T	A	G		A		G	T	A	T	T	A	
LV3	1	G	C	T	A	G		A		G	T	A	T	T	A	
	2	G	C	T	A	G		A		G	T	A	T	T	A	
LV4	1	G	C	T	A	G		A		G	T	A	T	T	A	
	2	G	C	T	A	G		A		G	T	A	T	T	A	
	3	G	C	T	A	G		A		G	T	A	T	T	A	
LV5	1	G	C	T	A	G		A		G	T	A	T	T	A	
	2	G	C	T	A	G		A		G	T	A	T	T	A	
	3	G	C	T	A	G		A		G	T	A	T	T	A	

Figure 5. Recombination of HIV-LAI with LVs. After culturing HIV-LAI for 3 months on LV-transduced cells (LV1-LV5; 2 or 3 cultures for each LV), the proviral DNA was amplified by PCR (PCR 1-3, as indicated in Figure 3) and the PCR product was directly used as template for Sanger sequencing. For every PCR-amplified region, sequence differences between HIV-LAI and the LVs, and the LV-derived mutations detected in the virus upon culturing, are indicated (orange boxes). Δ, deletion. **a**, position in HIV-LAI RNA (+1 at position 502 in pLAI); **b**, only differences at position -34 and -8 are shown; other differences in U3, including the SIN deletion in LV2-LV5, are not shown. The virus did not acquire any U3 mutation corresponding to the LV sequences; **c**: gcg frameshift mutation in gag; **d**: variable nucleotides in different LVs, never picked up by HIV; **e**: mixed sequence.

detected several other mutations, but these did not correspond to sequence differences in the LVs. Their presence and position varied for the different virus cultures, indicating that they likely resulted from random reverse transcriptase errors during virus replication (data not shown).

Analysis of the HIV-HBT cultures (Figure 6) revealed that this virus had acquired the T+23C mutation in all LV1-transduced cell cultures, but not in the cells transduced with any of the SIN LVs. Moreover, in one of these LV1-transduced cell cultures, the virus had acquired 3 single-nucleotide mutations in env. These mutations correspond with nucleotide differences in the LV1 env sequence. We did not detect any mutations that corresponded to other, upstream or downstream, nucleotide differences in LV1. These results indicate that HIV-LV recombination of both R and env sequences had occurred in this culture. Upon culturing in LV2 to LV5-transduced cells, none of these mutations, nor any other mutation that corresponded to nucleotide differences in these SIN LVs, were detected. Several random mutations were detected, which likely resulted from reverse transcriptase errors, as described above for HIV-LAI. Moreover, we detected a 15-bp nef deletion in the LV3-transduced cell culture 3, and a 237-bp nef deletion in the LV4-transduced cell culture 1.

Sequence analysis of the different HIV-CD cultures, revealed that in all LV1-transduced cell cultures the virus had acquired several mutations in both the TAR and polyA hairpin sequences in the R-U5 region. These mutations correspond to nucleotide differences in this region of LV1 (Figure 7), indicating that HIV-LV recombination had occurred. While the R-U5 sequences are identical in all LVs, the virus did not acquire these TAR and polyA mutations in any of the LV2 to LV5-transduced cell cultures. We did also not detect any other mutations corresponding to sequence differences in the LVs. As described for HIV-LAI and HIV-HBT,

	PCR 1						PCR 2		PCR 3							
		U3[b]		TAR	leader	gag		CTS		RRE	env					nef-U3
	position[a]	-34	-8	+23	+281	+379		+4469		+7416	+7684	+7690	+7694	+7746	+7867	
	HIV-HBT	G	C	T	A	G	HIV-HBT	A	HIV-HBT	G	T	A	T	T	A	
	LV1	C	C	C	C	GCG[c]	LV1/3/5	A	LV1-4	G	T	C	C	C	T	several nt[d]
	LV2-5	Δ	T	C	C	GCG[c]	LV2/4	Δ	LV5	C	Δ	C	C	C	T	
LV1	1	G	C	C	A	G		A		G	T	C	C	C	A	
	2	G	C	C	A	G		A		G	T	A	T	T	A	
	3	G	C	C	A	G		A		G	T	A	T	T	A	
LV2	1	G	C	T	A	G		A		G	T	A	T	T	A	
	2	G	C	T	A	G		A		G	T	A	T	T	A	
	3	G	C	T	A	G		A		G	T	A	T	T	A	
LV3	1	G	C	T	A	G		A		G	T	A	T	T	A	
	2	G	C	T	A	G		A		G	T	A	T	T	A	
	3	G	C	T	A	G		A		G	T	A	T	T	A	Δ15 bp nef
LV4	1	G	C	T	A	G		A		G	T	A	T	T	A	Δ237 bp nef
	2	G	C	T	A	G		A		G	T	A	T	T	A	
	3	G	C	T	A	G		A		G	T	A	T	T	A	
LV5	1	G	C	T	A	G		A		G	T	A	T	A/T[e]	A	
	2	G	C	T	A	nd		A		G	T	A	T	T	A	
	3	G	C	T	A	nd		A		G	T	A	T	T	A	

Figure 6. Recombination of HIV-HBT with LVs. After culturing HIV-HBT for 3 months on LV-transduced cells, the proviral DNA was amplified by PCR (PCR 1-3) and sequenced, as described in Figure 5. Sequence differences between HIV-HBT and the LVs are indicated (orange boxes). **a**, original position in HIV-LAI RNA reference sequence **b-d**: as described for Figure 5; **e**: mixed sequence (TTC-AGA mutation). Δ, deletion; nd, not determined; nt, nucleotide.

	PCR 1							PCR 2		PCR 3							
		U3[b]		TAR	polyA		leader	gag		CTS		RRE	env				nef-U3
	position[a]	-34	-8	+23	+59	+99	+281	+379		+4469		+7416	+7684	+7690	+7694	+7746	+7867	
	HIV-CD	G	C	T	GATCG	CGATC	A	G	HIV-CD	A	HIV-CD	G	T	A	T	T	A	
	LV1	C	C	C	ACTGC	GTAGT	C	GCG[c]	LV1/3/5	A	LV1-4	G	T	C	C	C	T	several nt[d]
	LV2-5	Δ	T	C	ACTGC	GTAGT	C	GCG[c]	LV2/4	Δ	LV5	C	Δ	C	C	C	T	
LV1	1	G	C	C	ACTGC	GTAGT	A	G		A		G	T	A	T	T	A	
	2	G	C	C	ACTGC	GTAGT	A	G		A		G	T	A	T	T	A	
	3	G	C	C	ACTGC	GTAGT	A	G		A		G	T	A	T	T	A	
LV2	1	G	C	T	GATCG	CGATC	A	G		A		G	T	A	T	T	A	
	2	G	C	T	GATCG	CGATC	A	G		A		G	T	A	T	T	A	
	3	G	C	T	GATCG	CGATC	A	G		A		G	T	A	T	T	A	
LV3	1	G	C	T	GATCG	CGATC	A	G		A		G	T	A	T	T	A	
	2	G	C	T	GATCG	CGATC	A	G		A		G	T	A	T	T	A	
	3	G	C	T	GATCG	CGATC	A	G		A		G	T	A	T	T	A	
LV4	1	G	C	T	GATCA	CGATC	A	G		A		G	T	A	T	T	A	
	2	G	C	T	GATCG	CGATC	A	G		A		G	T	A	T	T	A	
	3	G	C	T	GATCG	CGATC	A	G		A		G	T	A	T	T	A	
LV5	1	G	C	T	GATCG	CGATC	A	G		A		G	T	A	T	T	A	
	2	G	C	T	GATCG	CGATC	A	G		A		G	T	A	T	T	A	Δ117 nt nef
	3	G	C	T	GATCG	CGATC	A	G		A		G	T	A	T	T	A	

Figure 7. Recombination of HIV-CD with LVs. After culturing HIV-CD for 3 months on LV-transduced cells, the proviral DNA was amplified by PCR (PCR 1-3) and sequenced, as described in Figure 5. Sequence differences between HIV-CD and the LVs are indicated (orange boxes). **a**, original position in HIV-LAI RNA reference sequence; **b-d**: as described for Figure 5. Δ, deletion; nd, not determined; nt, nucleotide. Acquired nucleotide mutation in polyA hairpin sequence, not due to recombination, indicated in red and underlined.

	PCR 1							PCR 2		PCR 3																				
		U3[b]		TAR		leader	gag		CTS		RRE		Env					rtTA[e]												nef-U3
	position[a]	-34	-8	+22	+29	+281	+379		+4469		+7416	+7560	+7684	+7690	+7694	+7746	+7867	9	19	37	56	67	171	other						
	HIV-rtTA	G	C	AAG	ATT	A	G	HIV-rtTA	A	HIV-rtTA	G	T[f]	T	A	T	T	A	V	F	L	K	F	R							
	LV1	C	C	TCT	CTG	C	GCG[c]	LV1/3/5	A	LV1-3	G	G	T	C	C	C	T	na												several nt[d]
	LV2/3/5	Δ	T	TCT	CTG	C	GCG[c]	LV2	Δ	LV5	C	G	Δ	C	C	C	T	I	G	E	P	S	K							
LV1	1[g]	G	C	TCT	CTG	A	G		A		G	T	T	A	T	T	A	V	F	L	K	F	R	K33K aag-aaa						
	2	G	C	TCT	CTG	A	G		A		G	T	T	A	T	T	A	V	F	L	K	F	R	P161P cct-cca	D214N gat-aat					ΔtetO
LV2	1	G	C	AAG	ATT	A	G		A		G	T	T	A	T	T	A	V	F	L	K	F	R	T163A acc-gcc						
	2	G	C	AAG	ATT	A	G		A		G	T	T	A	T	T	A	V	F	L	K	F	R	F67F ttc-ttt						
	3	G	C	AAG	ATT	A	G		A		G	T	T	A	T	T	A	V	F	L	K	F	R							
LV3	1	G	C	AAG	ATT	A	G		A		G	T	T	A	T	T	A	V	F	L	K	F	R							
	2	G	C	AAG	ATT	A	G		A		G	T	T	A	T	T	A	V	F	L	K	F	R							
	3	G	C	AAG	ATT	A	G		A		G	T	T	A	T	T	A	V	F	L	K	F	R	L113L ctg-cta						
LV5	1	G	C	AAG	ATT	A	G		A		G	T	T	A	T	T	A	V	F	L	K	F	R							
	2	G	C	AAG	ATT	A	G		A		G	T	T	A	T	T	A	V	F	L	K	F	R							
	3	G	C	AAG	ATT	A	G		A		G	T	T	A	T	T	A	V	F	L	K	F	R							

Figure 8. Recombination of HIV-rtTA with LVs. After culturing HIV-rtTA for 3 months on LV-transduced cells, the proviral DNA was amplified by PCR (PCR 1-3) and sequenced, as described in Figure 5. Sequence differences between HIV-rtTA and the LVs are indicated (orange boxes). **a**, original position in HIV-LAI RNA reference sequence; **b-d**: as described for Figure 5; **e**, amino acid differences between rtTA^{safety} in HIV-rtTA and rtTA^{V16} in LV5 are indicated; **f**, G-to-T mutation at RRE/env position +7560 in HIV-rtTA (L586L in Env); **g**, analyzed after 2 months culturing. Δ, deletion; na, not applicable (LV1-3 do not carry the rtTA gene); nt, nucleotide.

several random mutations were detected, which were likely due to reverse transcriptase error. Moreover, we detected a 117-bp nef deletion in the LV5-transduced cell culture 2.

Analysis of HIV-rtTA upon culturing in LV1-transduced cells (Figure 8) revealed that - in both analyzed cultures - the virus had acquired several mutations in TAR, which corresponded with nucleotide differences in LV1, indicating HIV-LV recombination in the R region. The virus

A

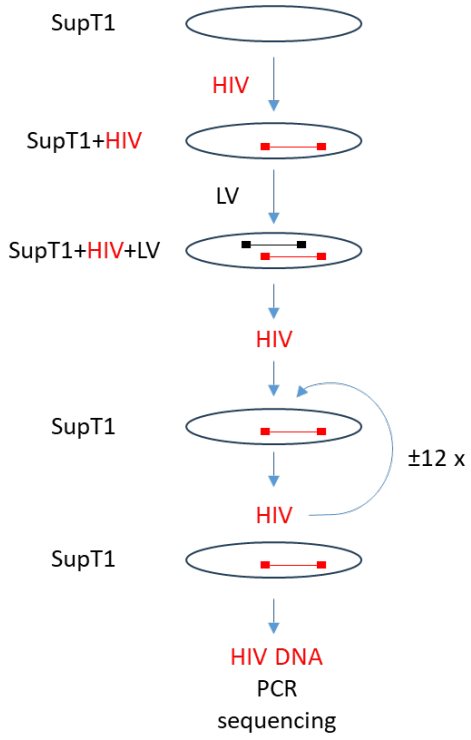
HIV-rtTA	culturing (3 months)	analysis (2 weeks)	rtTA sequence
	SupT1-LV5	SupT1	
	dox (ng/ml)	dox (ng/ml)	
1	1000	10	F67L (ttc-ctc) L70M (ctg-atg)
2	1000	100	no mutations
3	1000	100	D236D (gac-gat)

B

HIV-rtTA	culturing (3 months)	analysis (2 weeks)	rtTA sequence
	10 ng/ml dox	10 ng/ml dox	
1	SupT1-LV5	SupT1-LV5	no mutations
		SupT1	no replication
2	SupT1-LV5	SupT1-LV5	no mutations
		SupT1	no replication
3	SupT1-LV5	SupT1-LV5	no mutations
		SupT1	no replication

Figure 9. No recombination between HIV-rtTA and rtTA gene in LV5. (A) After culturing HIV-rtTA for 3 months on LV5-transduced SupT1 cells in the presence of 1000 ng/ml dox, the virus was passaged twice to regular SupT1 cells and cultured with 10 or 100 ng/ml dox. For cultures 2 and 3, replication was observed only at 100 ng/ml dox, whereas for culture 1 replication was also observed at 10 ng/ml dox. The rtTA region in the proviral DNA was amplified by PCR, and the PCR product was analyzed by Sanger sequencing. Mutations observed in the rtTA region are indicated. **(B)** After culturing HIV-rtTA for 3 months on LV5-transduced SupT1 in the presence of 10 ng/ml dox, the virus was passaged twice to LV5-transduced SupT1 or to regular SupT1 cells and cultured with 10 ng/ml dox. Replication was only observed on LV5-transduced cells. The rtTA region in the proviral DNA was amplified by PCR, and the PCR product was analyzed by Sanger sequencing. This analysis did not identify any mutations in the rtTA gene.

A



B

	PCR (3' LTR)					
		U3[b]		TAR	polyA	
	position[a]	-34	-8	+23	+59	+99
	HIV-CD	G	C	T	GATCG	CGATC
	LV1	C	C	C	ACTGC	GTAGT
	LV2-5	Δ	T	C	ACTGC	GTAGT
LV1	1	G	C	C	ACTGC	GTAGT
	2	G	C	C	ACTGC	GTAGT
	3	G	C	C	ACTGC	GTAGT
LV2	1	G	C	T	GATCG	CGATC
	2	G	C	T	GATCG	CGATC
LV3	1	G	C	T	GATCG	CGATC
	2	G	C	T	GATCG	CGATC
	3	G	C	T	GATCG	CGATC
LV4	1	G	C	T	GATCG	CGATC
	2	G	C	T	GATCG	CGATC
	3[c]	G	C	T	GATCG	CGATC
LV5	1	G	C	T	GGTCG	CGATC
	2	G	C	T	GATCG	CGATC
	3	G	C	T	GATCG	CGATC
control	1	G	C	T	GATCG	CGATC
	2	G	C	T	GATCG	CGATC
	3	G	C	T	GATCG	CGATC

Figure 10. Recombination of HIV-CD with LVs during transient replication in LV-transduced cells. (A) Experimental set up. SupT1 cells were infected with HIV-CD, and transduced with the LVs at three days post-infection (50 µl LV1-LV5 stock per 100 µl infected cells; expected to yield >67% LV1 to LV4-transduced cells and >44% LV5-transduced cells, based on Figure 4). As a control, cells were not transduced. At seven days post-infection, massive virus-induced syncytia were observed in all cultures, and the virus was passaged to regular, non-transduced SupT1 cells. After culturing the virus for 3 months on regular SupT1 cells, the proviral 3' LTR region was amplified by PCR and the PCR product was directly used as template for Sanger sequencing (HIV DNA in red; LV DNA in black). **(B)** Sequence analysis after 3 months of culturing on regular SupT1 cells. Sequence differences between HIV-CD and the LVs are indicated (orange boxes). **a**, original position in HIV-LAI RNA reference sequence; **b**: as described for Figure 5; **c**, analyzed after culturing for 2 months on regular SupT1 cells. Acquired nucleotide mutation in polyA hairpin sequence, not due to recombination, indicated in red and underlined. Δ, deletion.

did not acquire these mutations when cultured in cells transduced with one of the SIN LVs (LV2, LV3 and LV5; LV4, which contains the same WPRE sequence as LV3, was not tested). We did not detect any other mutation corresponding to sequence differences in the LVs. Also now we detected several random mutations that likely resulted from reverse transcriptase errors, including mutations in the rtTA gene (Figure 8). When HIV-rtTA, which encodes the suboptimal $rtTA^{safety}$ variant, was cultured on cells transduced with LV5, which contains the optimized $rtTA^{V16}$ gene, the virus did not acquire any mutation corresponding to sequence differences in $rtTA^{V16}$. Thus, we did not detect any recombination of rtTA sequences, despite a very high sequence similarity and the higher transcriptional activity of the LV5-encoded $rtTA^{V16}$ variant. Similar results were obtained in an independent experiment in which HIV-rtTA was cultured on LV5-transduced cells. Also now, we detected only mutations in rtTA that did not correspond to nucleotide differences in $rtTA^{V16}$ (Figure 9A). The virus was also cultured at a reduced dox level that is not sufficient to activate HIV-rtTA replication in regular, non-

experiment	HIV	LV	Recombination			
			R-U5	cPPT/CTS	RRE/Env	rtTA
1	HIV-LAI	non-SIN	2/3	0/3	0/3	
		SIN	0/10	0/10	0/10	
	HIV-HBT	non-SIN	3/3	0/3	1/3	
		SIN	0/12	0/12	0/12	
	HIV-CD	non-SIN	3/3	0/3	0/3	
		SIN	0/12	0/12	0/12	
	HIV-rtTA	non-SIN	2/2	0/2	0/2	
		SIN	0/9	0/9	0/9	0/9
2	HIV-CD	non-SIN	3/3			
		SIN	0/10			
TOTAL		non-SIN	13/14	0/11	1/11	
		SIN	0/53	0/43	0/43	0/9

Figure 11. HIV-LV recombination. Recombination detected upon continuous replication (experiment 1; experimental setup as shown in Figure 4C) and transient replication (experiment 2; experimental setup as shown in Figure 10A) of HIV variants on LV-transduced SupT1 cells. The frequency of recombination in the different regions is based on the data shown in Figures 5-9 (experiment 1) and Figure 10 (experiment 2). **a**, rtTA gene exclusively present in LV5; **b**, based on HIV-rtTA + LV5 data shown in Figures 8 and 9.

transduced SupT1 cells, but is sufficient for replication in LV5-transduced cells that express the optimized rtTA^{V16} variant. At this reduced dox level, acquisition of the rtTA^{V16} sequence was expected to significantly improve HIV-rtTA replication, in particular in regular SupT1 cells (15). However, after 3 months of culturing, we did not detect any mutation in the viral rtTA gene, and virus replication on regular SupT1 cells had not improved (Figure 9B), indicating that also this low-dox condition did not result in HIV-LV recombination in the rtTA region.

HIV-LV recombination in the R-U5 region was frequently observed when HIV had been cultured continuously in LV1-transduced cells for 3 months. An additional experiment was performed in which HIV replicated only transiently in SupT1 cells transduced with LVs, followed by long-term culturing of the virus on regular, non-transduced, cells (Figure 10A). We limited this experiment to the HIV-CD variant, which has multiple attenuating mutations in the polyA hairpin sequence in the R-U5 region. HIV-LV recombination of R-U5 sequences will restore the wild-type polyA hairpin sequence, which will improve viral replication.

SupT1 cells were infected with HIV-CD, and at three days post-infection - when the presence of small virus-induced syncytia indicated ongoing virus replication - the cells were transduced with the different LVs. At seven days post-infection, massive virus-induced syncytia were observed in all cultures, and the virus was passaged to regular, non-transduced SupT1 cells. After culturing the virus for 3 months, with weekly passaging to fresh non-transduced SupT1 cells, the viral LTR sequence was analyzed for each culture (Figure 10B). In the cultures where the virus had transiently replicated in LV1-transduced cells, the virus had acquired the TAR and polyA mutations corresponding to nucleotide differences in LV1, indicating HIV-LV recombination in the R-U5 region. These mutations were not detected when the virus had transiently replicated in cells transduced with a SIN LV. These results demonstrate that short-term replication of HIV in cells transduced with a non-SIN LV can be sufficient for frequent HIV-LV recombination in the R-U5 region.

These experiments, summarized in Figure 11, demonstrate frequent HIV-LV recombination in the R-U5 region when a non-SIN LV is used. Moreover, we detected a single recombination event in the env sequences when a non-SIV vector was used. We did not detect any recombination when a SIN LV was used.

Construction and replication of HIV-variants with PRE elements

All LVs contained a PRE element, but we did not detect any HIV-LV recombination that resulted in the insertion of this element into the viral genome. To investigate whether the PRE elements can improve HIV replication, we constructed HIV variants in which different PRE elements commonly present in LVs were inserted (Figure 12). We tested the 589-bp PRE fragment of woodchuck hepatitis virus (WHV), WPRE, which contains a truncated WHV protein X open reading frame (corresponding to the 60 N-terminal amino acids of the 142 amino acid viral protein), and a mutated fragment in which the protein X translation site had been inactivated (2). Furthermore, we tested the 1137-bp PRE fragment of human hepatitis B virus (HBV), HPRE, which encodes the nearly complete HBV protein X (147 of the 154 amino acids). The PRE fragments were inserted between the env and nef gene in the HIV-LAI genome, because insertion at a different position is more likely to affect essential viral regulatory motifs or viral protein functions. Inserted at this position, the PRE sequences will be present in all spliced and unspliced HIV RNAs encoding the different viral proteins. This insertion resulted in HIV-LAI variants with the wild-type WPRE (HIV-WPRE^{wt}), mutant WPRE (HIV-WPRE^{mut}), and HPRE (HIV-HPRE).

To compare the virus production of wild-type LAI and the different PRE variants, 293T cells were transfected with the virus-encoding plasmids, and the CA-p24 protein level in the

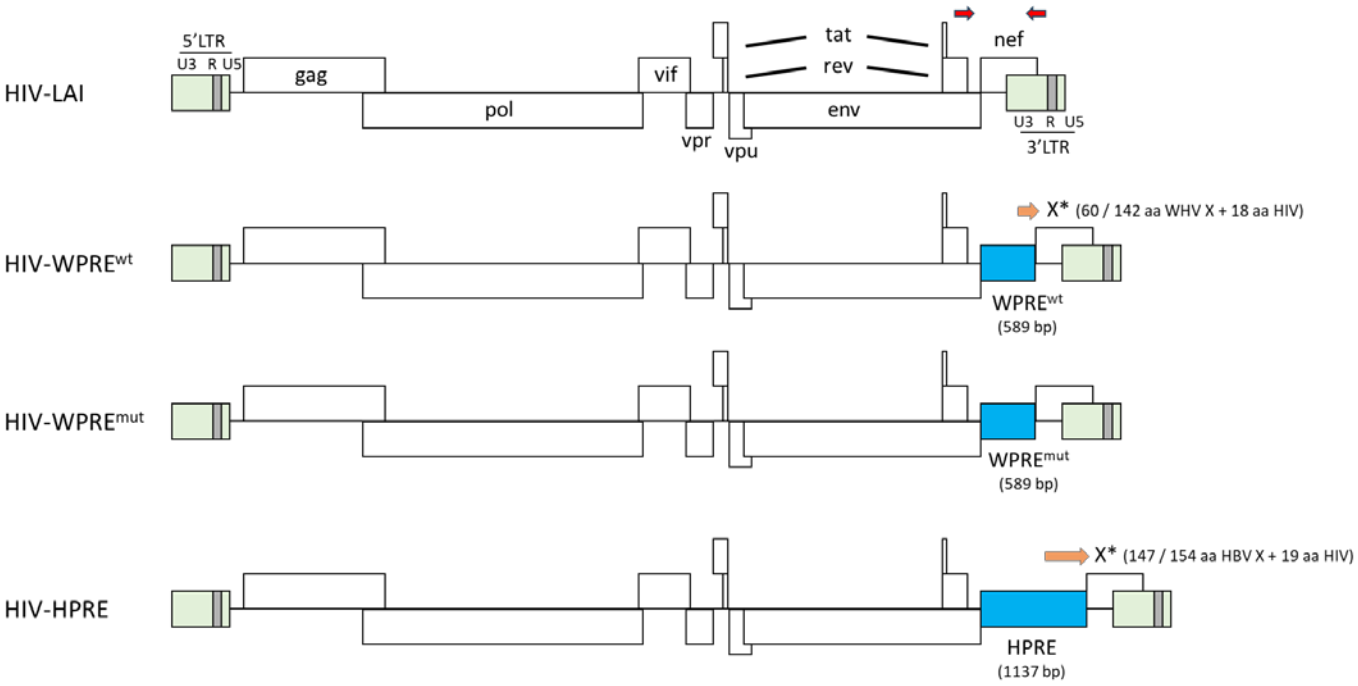


Figure 12. HIV-PRE variants. Different PRE fragments commonly used in LVs were inserted between the env and nef genes of HIV-LAI. The WHV PRE fragment in HIV-WPRE^{wt} includes a truncated open reading frame of the WHV protein X, encoding the N-terminal 60 amino acids of the 142-amino acid protein. If translation initiates at the protein X translation start site, it will terminate at a stop codon within the nef region, producing an X* protein composed of 60 amino acids from WHV X and 18 additional amino acids. In the WPRE^{mut} fragment used in HIV-WPRE^{mut}, the start site in the protein X ORF is inactivated through multiple nucleotide substitutions, preventing translation initiation. The HPRE insert in HIV-HPRE encodes nearly the full-length HBV protein X (147 of the 154 amino acids). Translation initiating at the protein X translation start site will terminate at a stop codon within nef, generating an X* protein consisting of 147 amino acids from HBV X and 19 additional amino acids. The red arrows indicate the position of the primers used in the PCR analyses shown in Figures 14 and 15.

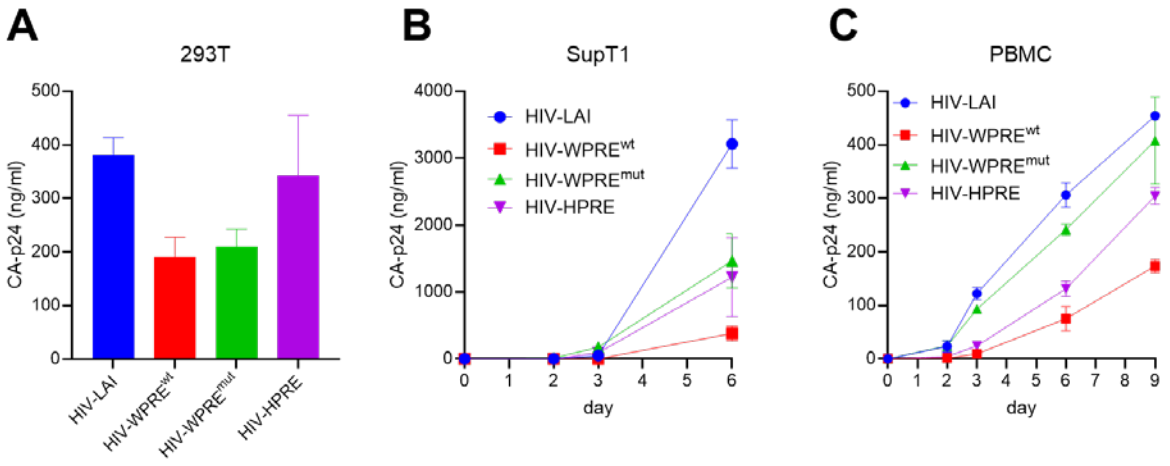


Figure 13. Viral production and replication of HIV-PRE variants. (A) HEK293T cells were transfected with the HIV-LAI construct and different PRE variants. The level of CA-p24 in the culture supernatant was measured at 2 days post-transfection. Data represent the mean ± SD from three independent experiments. (B-C). SupT1 cells (B) and PBMC (C) were infected with HEK293T-produced virus stocks (SupT1 inoculum corresponding to 0.11 ng CA-p24; PBMC inoculum corresponding to 0.55 ng CA-p24). Viral replication was monitored by measuring the CA-p24 level in the culture supernatant at several times post-infection (n=2, mean ± SD are shown).

culture supernatant was measured at 2 days after transfection (Figure 13A). This analysis revealed that the HIV-HPRE variant produced a similar level of CA-p24 as HIV-LAI, whereas the HIV-WPRE^{wt} and HIV-WPRE^{mut} variants showed a two-fold decreased level. To analyze the replication capacity of the different variants, equal amounts of the 293T-produced viruses (based on the level of the structural CA-p24 protein) were used to infect SupT1 T cells (Figure 13B) and primary peripheral blood mononuclear cells (PBMCs)(Figure 13C). Upon infection, virus replication resulted in a gradual increase in the CA-p24 level in all cultures. However, at 6 days after infection of the SupT1 cells, and at 9 days after infection of the PBMC, the CA-p24 levels detected for the wild-type HIV-LAI were higher than those for the PRE variants, indicating a reduced replication capacity of these variants. Overall, the replication capacity of the different PRE variants varied from highest for HIV-WPRE^{mut}, intermediate for HIV-HPRE and lowest for HIV-WPRE^{wt}.

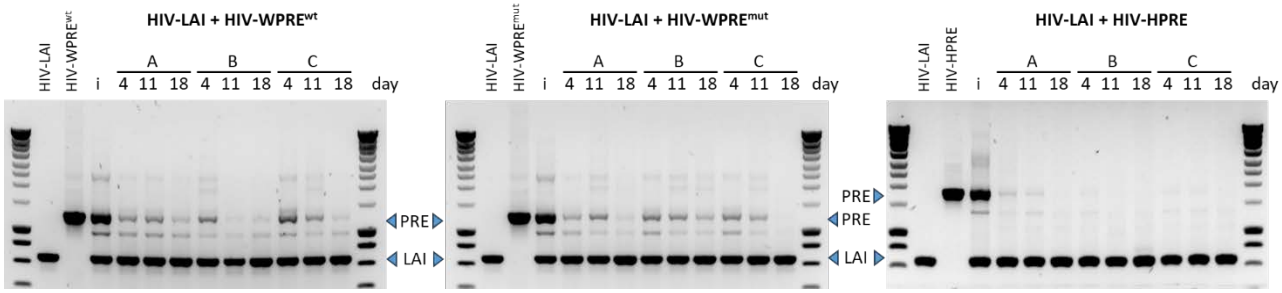


Figure 14. Virus competition assay. SupT1 cells were infected with a mixed virus stock, generated by simultaneous transfection of 293T cells with plasmids encoding HIV-LAI and one of the PRE variants. The input DNA mixture (i) and the integrated proviral DNA at 4, 8 and 11 days post-infection was analyzed by PCR using primers targeting the env and U3 regions flanking the PRE insertion site (indicated in Figure 12). This PCR produced differently sized products for HIV-LAI and the PRE variants, which were visualized by agarose gel electrophoresis. The individual HIV plasmids were included in the PCR analysis to confirm the identity of the bands corresponding to HIV-LAI and the PRE variant. The infections and subsequent analyses were performed in triplo (indicated with A, B and C).

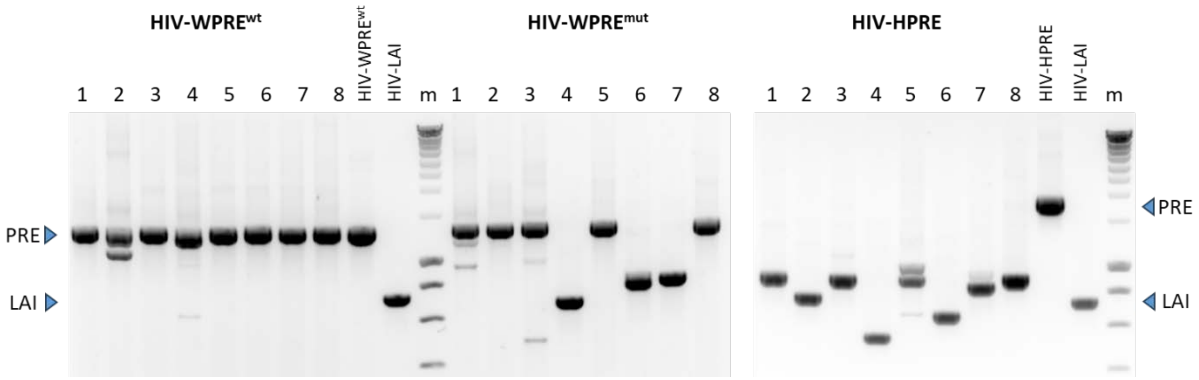


Figure 15. Genetic stability of HIV-WPRE^{wt}, HIV-WPRE^{mut} and HIV-HPRE. For each HIV-PRE variant, 8 independent cultures were initiated by infecting SupT1 cells. The virus was passaged weekly to fresh cells. After 87 days of culturing, integrated proviral DNA was analyzed by PCR using primers targeting the env and U3 regions flanking the PRE insertion site. The PCR products were visualized by agarose gel electrophoresis. HIV-PRE and HIV-LAI plasmids were included in the PCR analysis as references.

We also compared the replication capacity of wild-type HIV-LAI and the PRE variants in a virus competition assay. Virus stocks were produced by simultaneous transfection of 293T cells with the plasmids encoding wild-type HIV-LAI and one of the PRE variants. These stocks were used to infect SupT1 cells, and when massive virus-induced syncytia were detected, the viruses were passaged to fresh cells to prolong the cultures. At different times, the intracellular DNA - containing the integrated proviral DNA - was isolated and analyzed by PCR and agarose gel electrophoresis to detect both viruses. In this PCR, primers annealing upstream and downstream of the PRE insert were used, which resulted in differently sized PCR products for the two competing viruses, with the smaller product corresponding to wild-type HIV-LAI and the larger fragment corresponding to the PRE variant (Figure 14). For all virus combinations (HIV-LAI+HIV-WPRE^{wt}, HIV-LAI+HIV-WPRE^{mut}, and HIV-LAI+HIV-HPRE) 3 cultures were analyzed (Figure 14). As a control, the input DNA used to transfect the 293T cells was also analyzed by PCR, which yielded similar amounts of the two differently sized PCR products, reflecting an equal input of the wild-type HIV-LAI and PRE virus (input lanes in Figure 14). Analysis of the viral DNA isolated during culturing of the viruses revealed a rapid decrease in the larger PCR product for all cultures, indicating that wild-type HIV-LAI rapidly outcompeted the PRE variants. This result is in agreement with the superior replication of the wildtype HIV-LAI virus when compared to the PRE variants in SupT1 cells (Figure 13B).

To determine the genetic stability of the PRE-containing viruses upon prolonged culturing, we started 8 independent cultures for every variant in SupT1 cells. The virus was passaged weekly, when massive virus-induced syncytia were observed, to fresh cells. After culturing for 87 days, the intracellular DNA was isolated, and the integrated proviral DNA was analyzed by PCR amplification of the PRE region, followed by agarose gel electrophoresis (Figures 15) and sequencing of the PCR products (Figure 16).

Analysis of the HIV-WPRE^{wt} cultures revealed that the virus stably maintained the PRE element in 6 cultures (cultures W1, 3, 5, 6, 7, 8). In culture W4, the virus had acquired an 84-bp deletion in the 5' part of the WPRE region. PCR amplification of the proviral DNA from culture W2 yielded two, differently sized PCR products (Figure 15, lane W2). Sequence analysis of the smaller fragment revealed that the virus had acquired a 234-bp deletion in the PRE insert, upstream of the protein X translation start site. In culture W7, a 17-bp deletion was

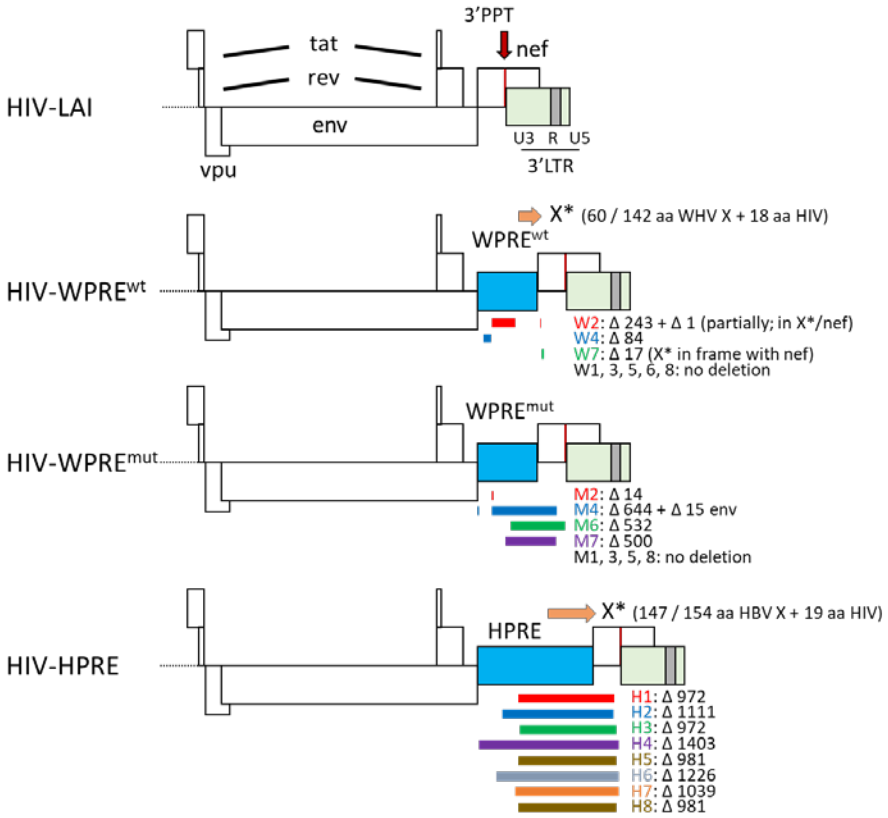


Figure 16. Genetic stability of HIV-PRE variants. After culturing HIV-WPRE^{wt}, HIV-WPRE^{mut} and HIV-HPRE for 87 days, integrated proviral DNA was analyzed by PCR using primers targeting the env and U3 regions flanking the PRE insertion site. The PCR products (Figure 15) were directly used as template for Sanger sequencing. The size and position of nucleotide deletions observed in and around the PREs are depicted in the schematics of the 3' half of the proviral genomes. For the HIV-WPRE^{wt} culture W2, the smaller PCR product was purified from gel and sequenced. In addition to the deletions, several nucleotide substitutions were detected in the analyzed fragments, but the occurrence and position of these mutations varied for every culture.

observed in nef, downstream of the X coding sequence, which placed the downstream nef sequence in frame with the upstream X sequences. If translations starts at the X start codon in this W7 variant, this will result in a X-Nef fusion protein, including 187 of the 206 Nef amino acids, but translation of the X open reading frame in the PRE variants was not analyzed.

Similar analysis of the HIV-WPRE^{mut} cultures revealed that this virus stably maintained the PRE insert in 4 cultures (M1, 3, 5, 8). A short 14-bp deletion in PRE was detected in culture M2, whereas larger deletions, varying in size from 500 to 644 bp, that included not only PRE but also adjacent nef sequences, were detected in 3 cultures (M4, 6 and 7). Notably, the deletions did not affect the 3'PPT site, which is in agreement with the essential function of this motif in reverse transcription. In the M4 culture, we detected an additional 15-bp deletion at the 3'end of the env gene, which affected the 3 C-terminal amino acids of the Env protein, but larger env deletions were not detected, which is in agreement with the essential function of this protein for infection.

Analysis of the HIV-HPRE cultures showed that the virus had acquired a large deletion in all cultures. These deletions varied in size from 972 to 1403 bp and included PRE and adjacent nef sequences. These deletions did not include upstream env or downstream 3'PPT sequences.

The frequent deletion of PRE sequences, as observed for HIV-WPRE^{mut} and HIV-HPRE, indicates that these sequences are not beneficial for virus replication, and that their removal enhances the *in vitro* replication efficiency. These findings are consistent with the attenuated phenotypes observed in both replication and competition assays.

Notably, the complete PRE insert was stably maintained in six HIV-WPRE^{wt} cultures. In the remaining two cultures, the deletions did not affect the region encoding 60 of the 142 amino acids of the WHV protein X. In contrast, the protein X-coding sequences in the HIV-WPRE^{mut} variant, in which the protein X translation start site had been inactivated, were deleted in three

cultures. Furthermore, the PRE sequences corresponding to the nearly complete HBV protein X protein in HIV-HPRE (encoding 147 of the 154 amino acids) were deleted in all cultures. These observations raise the possibility that the HIV-WPRE^{wt} variant may produce, and derive benefit from, a truncated X protein. However, further investigation is required to explore this speculative hypothesis. To exclude the possibility that large deletions - including the protein X region - were missed in the HIV-WPRE^{wt} cultures due to the loss of env or U3 sequences required for PCR primer binding, we conducted an additional PCR analysis using primers targeting more upstream, essential tat/rev sequences and more downstream, essential R sequences. It is highly unlikely that the virus would delete these primer-binding regions, as such deletions would compromise critical viral functions. This analysis did not reveal any deletions beyond those previously identified, thereby confirming the stable maintenance of the X coding region.

6. DISCUSSION

We experimentally investigated recombination between different HIV and LVs. Upon long-term culturing of HIV in T cells transduced with a non-SIN LV, we detected frequent recombination of R-U5 sequences, but also a single recombination event involving env sequences. We did not detect any recombination event when a SIN LV is used. Moreover, we never observed insertion of other LV sequences into the HIV genomes.

HIV-LV recombination requires co-packaging of the HIV RNA and the LV RNA within virions. Upon infection of a target cell, both RNAs are reverse transcribed, resulting in the formation of double-stranded HIV and LV DNAs. Frequent HIV-LV recombination in the R-U5 region, as observed in cells transduced with a non-SIN LV, likely arises from an intermolecular first strand-transfer event during reverse transcription, as illustrated in Figure 17. Additionally, recombination in the env region was detected, which likely results from template switching during minus-strand DNA synthesis following the first strand-transfer step, as depicted in Figure 18.

HIV-LV recombination was exclusively detected in cells transduced with a non-SIN LV, and not in those transduced with a SIN LV. The intact, Tat-responsive LTR promoter present in integrated non-SIN LV DNA enables LV RNA production upon HIV infection, when Tat is expressed. In contrast, the U3 deletion in SIN LV DNA inactivates the LTR promoter, resulting in a marked reduction of LV RNA production in transduced cells, even in the presence of Tat. These findings indicate that the SIN deletion suppresses LV RNA production sufficiently to prevent HIV-LV recombination under the conditions used in our experiments.

Effect of LV-derived mutations on HIV replication

In our experiments, HIV variants generated through HIV-LV recombination with mutations that enhance the viral replication capacity, will outcompete the parental virus and become detectable after prolonged culturing. Recombination in the R-U5 region consistently introduced a T+23C mutation in the HIV TAR element. This mutation alters the bulge sequence in the TAR RNA hairpin from UUU to UCU. The bulge, together with the apical loop of TAR, is critical for binding Tat and cellular cofactors that enhance transcription from the U3 promoter (19). The frequent outgrowth of HIV-LAI and HIV-HBT variants carrying the T+23C mutation suggests

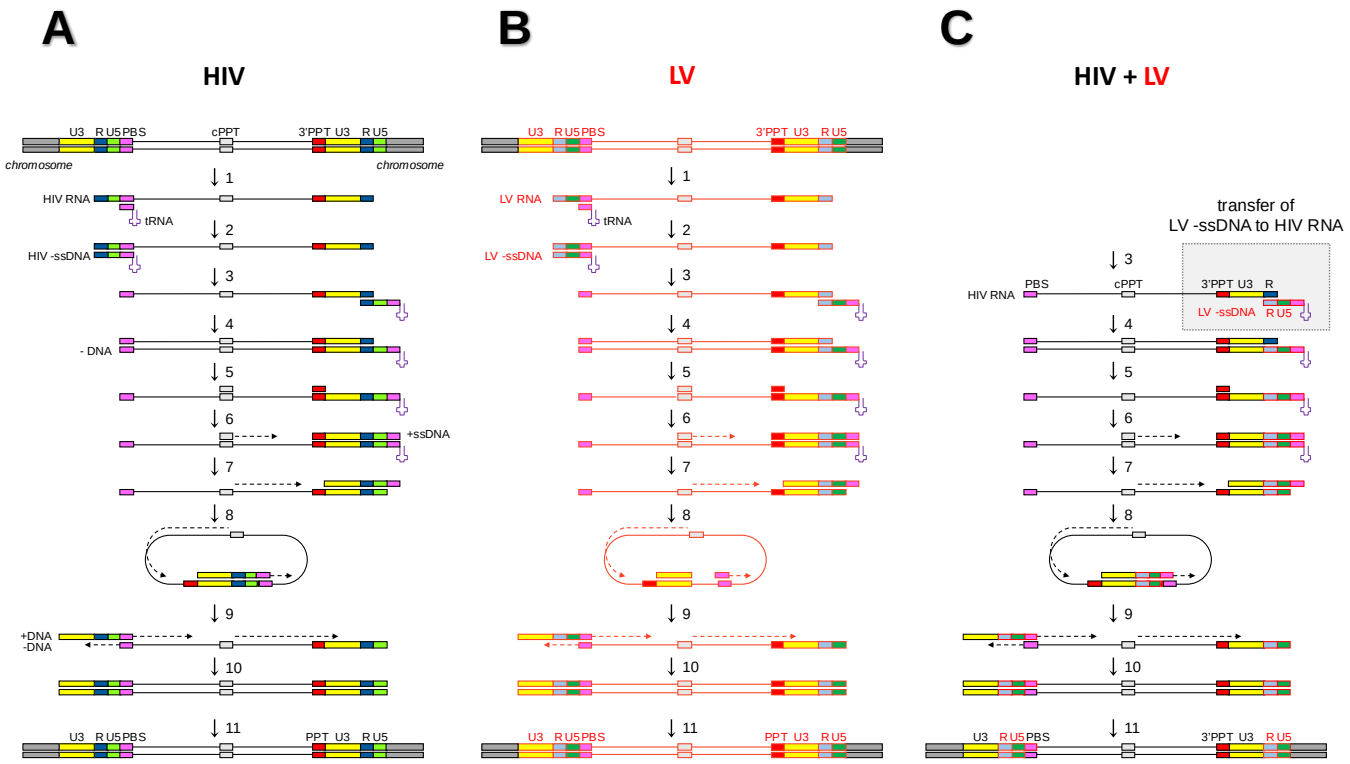


Figure 17. HIV-LV recombination in the R-U5 region resulting from an intermolecular transfer of minus-strand strong-stop DNA. (A) Transcription initiation at the 5' LTR of the HIV proviral DNA, and subsequent polyadenylation at the 3' LTR, produces full-length HIV RNA genomes. Two of these RNAs form a dimer, which is packaged into virions (step 1). Upon infection of the cell, the HIV RNA is copied into double-stranded DNA (dsDNA) by the viral reverse transcriptase (RT; steps 2-10). The cellular $tRNA^{Lys3}$ binding to the primer binding site (PBS) near the 5' end of the HIV RNA serves as primer for reverse transcription of the R-U5 region (step 2). This leads to the production of a minus-strand strong-stop DNA (-ssDNA) and degradation of the reverse transcribed RNA fragment. The -ssDNA is transferred to the 3' end of the HIV RNA, which is facilitated by the sequence complementarity between the R regions (step 3). This strand transfer allows upstream RNA sequences to be copied into the minus-strand DNA (-DNA; step 4). Concomitant with -DNA synthesis, the viral RNA template is degraded by the RNase H activity of RT, except for the central and 3' polypurine tracts (cPPT and 3'PPT), which resist cleavage and function as primer for plus-strand DNA (+DNA) synthesis (steps 5-7). Sequence complementarity between the ends of the partially double-stranded HIV DNA induces circularization (step 8). Subsequent steps, including continuation of +DNA synthesis, linearization (involving strand displacement of the LTR sequences), and completion of -DNA synthesis, result in a linear dsDNA molecule with complete 5' and 3' LTR motifs (step 9-10). This full-length dsDNA is integrated into the host cellular DNA by the viral integrase (step 11). **(B)** Similar to HIV RNA, transcription, dimerization and packaging of LV RNAs results in virus particles with two LV RNA molecules (step 1). Upon infection, these LV RNAs are reverse transcribed into double-stranded LV DNA (step 2-10), which is integrated into the cellular DNA (step 11). **(C)** Mechanism for HIV-LV recombination in the R-U5 region. If both HIV and LV RNAs are produced in the same cell, heterodimers can form, consisting of one HIV RNA and one LV RNA. Upon infection, both RNAs are reverse transcribed, producing HIV and LV -ssDNAs, which can be transferred to the 3' end of the same RNA template (intramolecular strand transfer; as shown in A and B) or to the co-packaged RNA template (intermolecular strand transfer). When the -ssDNA generated from the LV RNA is transferred to the 3' end of the HIV RNA (step 3), the LV-derived R-U5 sequences will be incorporated into both the 5' and 3' LTRs of the double-stranded HIV DNA during subsequent reverse transcription steps (step 4-10).

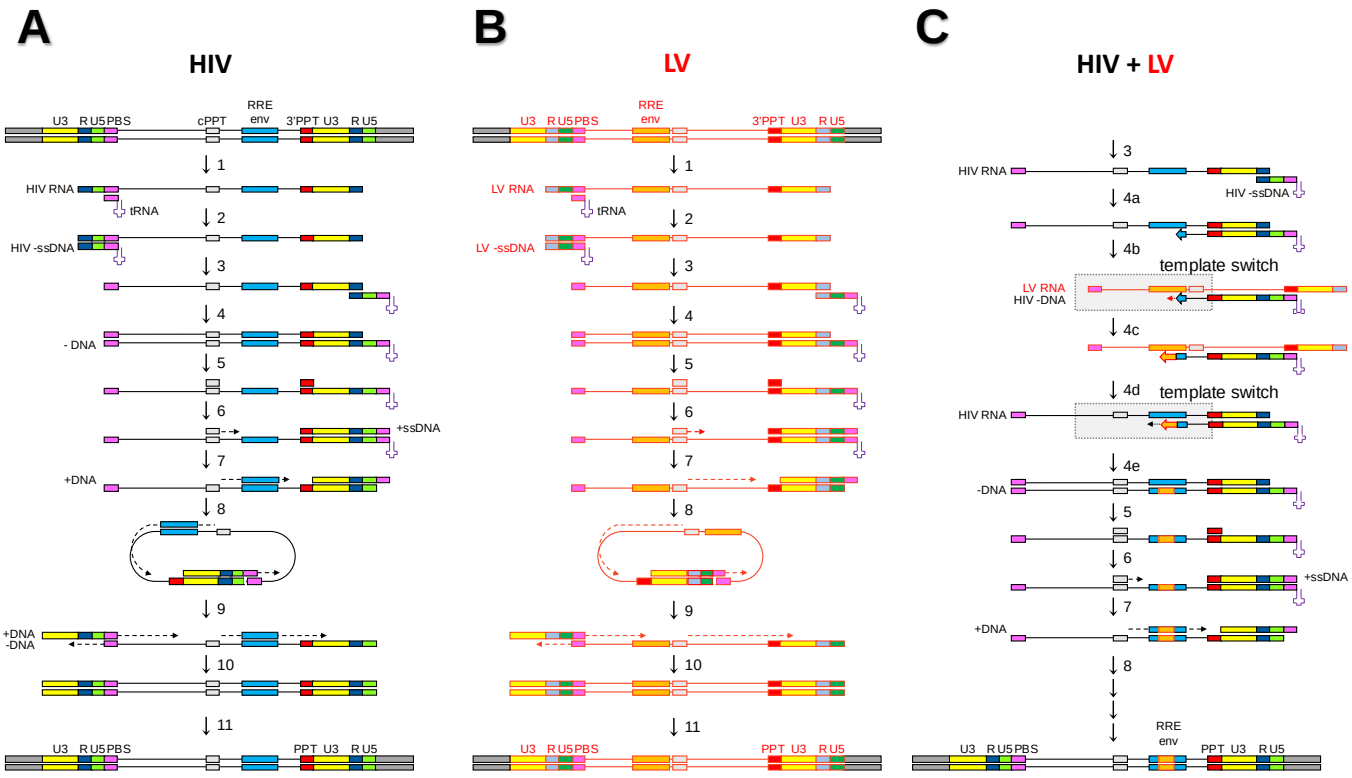


Figure 18. HIV-LV recombination in the env region resulting from template switching during minus-strand DNA synthesis. (A-B) HIV and LV transcription and reverse transcription as described for Figure 17A-B. The position of the env sequences (including the RRE domain) present in both genomes is depicted. **(C)** Mechanism for HIV-LV recombination in the env region. If both HIV and LV RNAs are produced in the same cell, heterodimers can form that consist of one HIV RNA and one LV RNA. Upon infection, both RNAs serve as template for reverse transcription. Following the intramolecular transfer of the -ssDNA synthesized from the HIV RNA to the 3' end of the RNA (step 3), upstream RNA sequences, including env sequences, can be copied into minus-strand DNA (-DNA; step 4a). Concurrent with DNA synthesis, the RNase H activity of reverse transcriptase degrades the HIV RNA template, exposing single-stranded regions of the nascent -DNA that can base pair with complementary sequences in the LV RNA, thereby facilitating a template switch (step 4b). Continued reverse transcription results in the incorporation of the LV env sequences into the -DNA (step 4c). A subsequent template switch will transfer the -DNA back to the HIV RNA template (step 4d), after which the upstream HIV sequences can be copied into the -DNA (step 4e). In the later stages of reverse transcription, the LV-derived env sequences present in the -DNA are copied to the plus-strand DNA (+DNA) (step 7) and a double-stranded HIV DNA will form with env mutations in both strands. This mechanism follows the 'dynamic copy choice' model of template switching (reviewed in (4)).

that this bulge mutation enhances viral replication in SupT1 cells. In HIV-CD, recombination in the R-U5 region not only introduced the T+23C mutation in TAR, but also restored the wild-type polyA hairpin sequence, which likely further enhanced replication. Similarly, R-U5 recombination in HIV-rtTA resulted not only in an UCU TAR bulge, but also restored the wild-type TAR loop sequence, which likely contributed to enhanced replication, as the mutated Tat protein encoded by HIV-rtTA can activate transcription from a wild-type TAR-containing LTR promoter, albeit at a reduced level compared to wild-type Tat (20). HIV-LV recombination in the env region of HIV-HBT introduced two non-synonymous mutations (A+7690C and T+7694C, resulting in N630H and M631T amino acid substitutions in Env) and one synonymous mutation (T+7746C, resulting in H648H in Env). The outgrowth of this variant suggests that these Env mutations, like the T+23C R-U5 mutation, enhance viral replication in SupT1 cells. The R-U5 and env sequences in the LVs correspond to the sequences in the HIV HXB2 molecular clone, which is a derivative of the LAI isolate, like the HIV-LAI molecular clone. Prior to the generation of the HXB2 molecular clone, however, the LAI isolate was extensively cultured on T cell lines (21). It is possible that *in vitro* evolution during this culturing process led to the acquisition of mutations that optimized replication, including the T+23C TAR mutation and the env mutations. Taken together, these findings suggest that the R-U5 and env mutations acquired by HIV through HIV-LV recombination enhance viral replication in SupT1 cell cultures.

When we cultured HIV-rtTA in cells transduced with LV5, which contains an optimized rtTA^{V16} gene, we did not detect any recombination involving rtTA sequences. This absence of recombination can be explained by the fact that LV5 carries a self-inactivating (SIN) LTR, which - as described above - prevents sufficient LV RNA production to enable HIV-LV recombination. The CMV promoter upstream of the rtTA^{V16} gene drives the production of rtTA^{V16} mRNAs. Theoretically, template switching between HIV-rtTA RNAs and rtTA^{V16} mRNAs could result in the insertion of the optimized rtTA sequence into HIV-rtTA. However, rtTA^{V16} mRNAs lack the HIV-specific signals required for dimerization and packaging, effectively excluding this recombination pathway.

Potential implications of the experimental findings for risk assessment

Our experimental study demonstrates that HIV-LV recombination can occur when cells are transduced with a non-SIN LV. It is important to emphasize that the experimental conditions were deliberately optimized to maximize the likelihood of recombination:

(1) High sequence similarity: The HIV variants and the HIV-derived LV components shared extensive sequence homology.
(2) High transduction levels: Cells were massively transduced, with some cells likely harboring multiple integrated LV DNA copies.
(3) Prolonged viral replication: Cultures were maintained over extended periods, allowing unrestricted HIV replication and resulting in high viral loads.

In contrast, in clinical and most research applications of LVs, the conditions will be less favorable for HIV-LV recombination:

(1) Lower sequence similarity: The sequence similarity between the LV and natural virus isolates will be lower, in particular in the env coding region, which is among the most variable regions of the HIV genome.
(2) Lower LV copy number: Multiple integration events are generally avoided (to reduce the risk of insertional oncogenesis), resulting in fewer LV DNA copies per cell.
(3) Limited HIV replication:

- In most laboratory research settings, HIV-induced cytopathogenic effects lead to cell death, curtailing replication.
- In clinical settings, individuals treated with LV's or LV-modified cells are typically closely monitored medically and will, upon HIV diagnosis, promptly initiate antiretroviral therapy (ART) to effectively suppress viral replication.

Please note that HIV-LV recombination is only relevant in risk assessments if the LV-transduced cells (or the cells that will be transduced) are, were, or could become susceptible to HIV infection. This includes (for example) human T cells ('are'), induced pluripotent stem cells (iPSCs) derived from human T cells ('were'), or progenitor cells that may differentiate into HIV-susceptible cell types ('could become').

The presence of HIV can be assessed either by:

- Direct testing of the patient, cell donor, or the cells themselves (though this poses challenges in terms of sensitivity and appropriate controls), or
- Reasoned exclusion, based on the origin and differentiation potential of the cells (i.e., whether it can be confidently ruled out that the cells are, were, or will become susceptible to HIV).

Laboratory research with LVs

In laboratory studies, cells are frequently transduced with LVs to introduce a GOI. HIV-LV recombination may occur either when HIV-infected cells are transduced with an LV or when LV-transduced cells are infected with HIV.

Additionally, LV-transduced cells can be introduced into laboratory animals, or the LV itself can be directly administered to animals. However, HIV does not infect commonly used laboratory animals such as mice and rats, with the exception of genetically modified animals that, for example, possess a humanized immune system derived from human cells.

- **Laboratory research with SIN-LVs**

Our experimental data indicate that HIV-LV recombination does not occur when SIN LVs are used.

- **Laboratory research with non-SIN LVs**

In contrast to SIN LVs, our data show that HIV-LV recombination is possible when non-SIN LVs are used. If HIV-infected cells are transduced with a non-SIN LV, recombination may occur frequently in the R-U5 region, and at lower frequency in other homologous regions shared between the virus and the LV. These recombination events may affect HIV replication, particularly when the input virus is attenuated by mutations in these regions (as exemplified by the HIV-CD and HIV-rtTA variants in our experiments), but it seems unlikely that they will alter the host range of the virus.

HIV-LV recombination could theoretically also result in the insertion of other LV components, such as the PRE, GOI, or GOI promoter, but the likelihood of such events occurring, and leading to a virus with increased pathogenicity or altered tropism, is low:

- **PRE**: Although engineered HIV-PRE variants exhibited attenuated replication, suggesting that PRE insertion cannot improve HIV replication, we cannot formally exclude that insertion of PRE sequences at another position - than the tested position between the env and nef genes - or insertion of smaller PRE fragments, can lead to

less attenuation or even improve virus replication. When HIV-LV recombination was experimentally tested, however, HIV did not acquire the PRE insert in cultures transduced with a non-SIN LV (LV1).

- **GOI**: In most cases, the GOI encodes a cellular factor that is unlikely to enhance HIV replication.
- **GOI Promoter**: Recombination between the HIV LTR and the GOI promoter is improbable when sequence similarity is low.

Additionally, the insertion of foreign sequences into the HIV genome typically attenuates replication by disrupting the complex RNA splicing processes required for the balanced production of over 40 alternatively spliced HIV transcripts encoding the various viral proteins.

Nevertheless, when non-SIN LVs are used, it cannot be excluded that HIV-LV recombination can lead to the insertion of LV sequences, which may enhance pathogenicity or alter host tropism, for example if the GOI encodes a factor that stimulates HIV replication.

Clinical applications of LVs in humans

In clinical settings, cells are typically isolated from the patient (autologous transduction) or from a donor (allogeneic transduction), transduced *ex vivo* with a lentiviral vector (LV), and subsequently (re)infused into the patient. If the patient or donor is HIV-infected prior to cell isolation, or if the patient acquires HIV infection after treatment, both HIV and the LV may be present in the same cell, creating a potential opportunity for recombination. Although donors are routinely screened for HIV - making it unlikely that donor-derived cells are HIV-positive - this scenario was included in our considerations for completeness.

- **Clinical applications of SIN LVs**

Our experimental data indicate that HIV-LV recombination does not occur when SIN LVs are used.

- **Clinical applications of non-SIN LVs**

As described above for the use of non-SIN LVs in laboratory research, when non-SIN LVs are used in clinical applications, it cannot be excluded that HIV-LV recombination could generate a virus variant with increased pathogenicity or altered tropism.

If the patient or donor is HIV-infected prior to cell isolation, HIV replication can be suppressed through antiretroviral therapy (ART), which reduces the viral load and thereby the chance of recombination upon LV transduction. However, during the transduction process, reverse transcriptase and integrase inhibitors—which are commonly used components of ART—cannot be applied, as they would interfere with the lentiviral transduction itself. This creates a temporary window during which HIV may replicate and potentially recombine with the LV. Following (re)infusion of the transduced cells, ART can be resumed, effectively suppressing HIV replication and preventing HIV-LV recombination. Moreover, any HIV-LV recombinant generated during the transduction period is expected to be sensitive to ART, as it is unlikely that acquisition of LV sequences would confer resistance to commonly used ART drugs.

Similarly, if the patient acquires HIV infection after receiving LV gene therapy, ART can be used to inhibit HIV replication, which will minimize the risk of recombination and suppress any HIV-LV recombinant that may have formed prior to ART initiation.

However, it is impossible to guarantee continuous effective ART, as side effects may lead to temporary discontinuation of ART, and the emergence of drug-resistant HIV variants cannot be ruled out, particularly in cases of suboptimal drug adherence. Such circumstances could create opportunities for HIV-LV recombination.

Finally, it should be noted that the presence of HIV is always a critical consideration for occupational health and safety, as HIV is a pathogenic virus that can cause AIDS. While ART effectively suppresses HIV replication and disease progression, it does not provide a permanent cure, as the virus persists in integrated DNA form within infected cells. As a consequence, discontinuation of ART leads to renewed viral replication and disease progression. Therefore, people living with HIV require lifelong ART, which carries risks of resistance development and adverse side effects.

7. METHODS

Cells. All virus and cell culturing was performed at the BSL-3 (ML-III) lab. Human embryonic kidney 293T cells (HEK293T, ATCC CRL-11268) were cultured in Dulbecco's Modified Eagle Medium (DMEM) (GIBCO, Life Technologies) supplemented with 10% fetal calf serum (FCS), non-essential amino acids (0.1 mM, Invitrogen), penicillin and streptomycin (both at 100 U/ml). The SupT1 T cell line (22) was obtained through the NIH HIV Reagent Program (AIDS Research and Reference Reagents Program, Division of AIDS, ARP-100). SupT1 cells were cultured in advanced RPMI1640 medium (GIBCO, Life Technologies) with 1% L-glutamine, FCS (1%), penicillin and streptomycin (both at 15 U/ml). When indicated, 10, 100 or 1000 ng/ml doxycycline (dox) was added to the culture medium before use. Human peripheral blood mononuclear cells (PBMC) were isolated from fresh buffy coats by standard Ficoll–Hypaque density centrifugation and cultured as previously described (23).

LV transduction. Plasmids pLVX-EGFP-IRES-puro (LV1) was a gift from Robert Sobol (Addgene 128652)(24). pCDH-CMV-mCherry-T2A- Puro (LV2) was a gift from Kazuhiro Oka (Addgene 72264). pRRL-SIN-cPPT-PGK-GFP-WPRE (LV3) was a gift from Didier Trono (Addgene 12252). pLenti-CMV-GFP-Puro (658-5)(LV4) was a gift from Eric Campeau & Paul Kaufman (Addgene 17448)(25). TRE-CMV-rtTA-V16 (LV5) was previously described (26, 27). LV particles were produced as described previously (28). Briefly, 293T cells were transfected with Lipofectamin 2000 (ThermoFisher) in Optimem and LV1 plus psPAX2 (Addgene 12260) and pVSV-G (Addgene 138479), and LV2 to LV5 plus pMDLg/pRRE (Addgene 12251), pRSV-Rev (Addgene 12253) and pVSV-G. The LV-containing culture supernatant was harvested at 2 days posttransfection and filtered (0.45 µm filter). SupT1 cells (1 x 10^5 cells) in 250 µl RPMI medium were mixed with 250 µl LV particles (undiluted, and 5- or 25-fold diluted in culture medium). After 6 hr, cells were pelleted by centrifugation at 180 x g for 5 min and resuspended in fresh culture medium. Cells were expanded and cryopreserved in aliquots. Fluorescent protein expression was assessed by flow cytometry at several times post-transduction.

Viruses. All HIV-1 constructs are based on the pLAI plasmid (1) encoding the full-length HIV-1 LAI genome. HIV-HBT-A is an optimized HiBiT-tagged variant described by Dekker et al. (3). HIV-CD has several nucleotide substitutions in the polyA hairpin sequence in the R-U5 region (14). The HIV-rtTA variant used in this study, HIV-rtTA$^{G19F\ E37L\ P56K}$, contains specific mutations in the rtTA gene that prevent the loss of dox control during long-term culture (17). The 422-bp BamHI-XhoI env-nef fragment in pLAI was replaced with a DNA fragment containing the WPRE^{wt} sequence, flanked by MluI and NotI sites, between the env and nef sequences, which resulted in pHIV-WPRE^{wt}. The MluI-NotI fragment in pHIV-WPRE^{wt} was replaced with DNA fragments containing the WPRE^{mut} and HPRE sequences, which resulted in pHIV-WPRE^{mut} and pHIV-HPRE.

Virus stocks were produced by transfection of the plasmids into 293T cells with Lipofectamin 2000 (29). HIV-rtTA transfected cells were cultured with 1 µg/ml doxycycline. The virus-containing culture supernatants were harvested at 2 days post-transfection and centrifuged at 180 x g for 5 min to remove cells.

HIV-LV recombination experiments. For the experiments shown in Figure 5 to 9), LV-transduced SupT1 cells (~2 x 10^5 cells) in 1 ml culture medium were infected with 1 µl virus stock. Cultures were maintained for 3 months. HIV-rtTA-infected cells were cultured with 1000 ng/ml dox (unless differently indicated in the legend). When extensive virus-induced syncytia appeared, i.e. approximately weekly, the virus-containing culture supernatant was passaged to fresh LV-transduced cells. Initially, 1 µl undiluted virus sample was used for passaging, but

the virus input was gradually reduced to 1 µl of a 40-fold diluted sample. After 3 months, the virus was similarly passaged twice to regular, non-transduced SupT1 cells. When massive syncytia were observed after the second passage cells, the proviral was analyzed as previously described (29). In brief, the intracellular DNA was solubilized by resuspending the cells in Tris-HCl (pH 8.0), 0.1 mM EDTA, and 0.5% Tween 20, followed by incubation with 200 mg/mL proteinase K at 56°C for 60 min and 95°C for 10 min, and used as template for PCR (PCR1-3 in Figure 3) with three primer sets: forward U3 primer 9 (agggctaattcactcccaacgaagac) plus reverse gag primer 2 (tcttgtggggtggctccttctga), forward pol primer 3 (gttggtgggcgggaatcaagca) plus reverse env primer 4 (ttgtgggttggggtctgtgggt), and forward env primer 15 (ctcccatcagcggacaaatta) plus reverse U5 primer 16 (agttaccagagtcacacaacag). The PCR products were analyzed by agarose gel electrophoresis and directly used as template for Sanger sequencing, which revealed the dominant proviral DNA sequence in the culture. When indicated the PCR product was purified from gel before sequencing.

For the experiment shown in Figure 10, SupT1 cells were infected with HIV-CD, and transduced with the LVs at three days post-infection (50 µl LV1-LV5 stock per 100 µl infected cells). At seven days post-infection, massive virus-induced syncytia were observed, and the virus was passaged to regular, non-transduced SupT1 cells. After culturing the virus for 3 months on regular SupT1 cells, the proviral 3' LTR region was amplified by PCR with forward env primer EndEnvS (tagaagaataagacagggcttgg) and reverse U5 primer CN1 (ggtctgagggatctctagttaccagagtc). The PCR product was directly used as template for Sanger sequencing.

8. REFERENCES

1. Peden K, Emerman M, Montagnier L. 1991. Changes in growth properties on passage in tissue culture of viruses derived from infectious molecular clones of HIV-1LAI, HIV-1MAL, and HIV-1ELI. Virology 185:661–72.
2. Zanta-Boussif MA, Charrier S, Brice-Ouzet A, Martin S, Opolon P, Thrasher AJ, Hope TJ, Galy A. 2009. Validation of a mutated PRE sequence allowing high and sustained transgene expression while abrogating WHV-X protein synthesis: application to the gene therapy of WAS. Gene Ther 16:605–19.
3. Dekker JG, Klaver B, Berkhout B, Das AT. manuscript in preparation. Construction and optimization of an efficiently replicating HiBiT-tagged HIV-1 reporter virus.
4. Smyth RP, Davenport MP, Mak J. 2012. The origin of genetic diversity in HIV-1. Virus Res 169:415–29.
5. Zufferey R, Dull T, Mandel RJ, Bukovsky A, Quiroz D, Naldini L, Trono D. 1998. Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. J Virol 72:9873–80.
6. Cornetta K, Yao J, Jasti A, Koop S, Douglas M, Hsu D, Couture LA, Hawkins T, Duffy L. 2011. Replication-competent lentivirus analysis of clinical grade vector products. Mol Ther 19:557–66.
7. Cornetta K, Duffy L, Turtle CJ, Jensen M, Forman S, Binder-Scholl G, Fry T, Chew A, Maloney DG, June CH. 2018. Absence of Replication-Competent Lentivirus in the Clinic: Analysis of Infused T Cell Products. Mol Ther 26:280–288.
8. Evans JT, Garcia JV. 2000. Lentivirus vector mobilization and spread by human immunodeficiency virus. Hum Gene Ther 11:2331–9.
9. Hanawa H, Persons DA, Nienhuis AW. 2005. Mobilization and mechanism of transcription of integrated self-inactivating lentiviral vectors. J Virol 79:8410–21.

10. Logan AC, Haas DL, Kafri T, Kohn DB. 2004. Integrated self-inactivating lentiviral vectors produce full-length genomic transcripts competent for encapsidation and integration. J Virol 78:8421–36.
11. Chen J, Powell D, Hu WS. 2006. High frequency of genetic recombination is a common feature of primate lentivirus replication. J Virol 80:9651–8.
12. Rawson JMO, Nikolaitchik OA, Keele BF, Pathak VK, Hu WS. 2018. Recombination is required for efficient HIV-1 replication and the maintenance of viral genome integrity. Nucleic Acids Res 46:10535–10545.
13. Schlub TE, Smyth RP, Grimm AJ, Mak J, Davenport MP. 2010. Accurately measuring recombination between closely related HIV-1 genomes. PLoS Comput Biol 6:e1000766.
14. Das AT, Klaver B, Klasens BI, van Wamel JL, Berkhout B. 1997. A conserved hairpin motif in the R-U5 region of the human immunodeficiency virus type 1 RNA genome is essential for replication. J Virol 71:2346–56.
15. Das AT, Berkhout B. 2016. Conditionally replicating HIV and SIV variants. Virus Res 216:66–75.
16. Verhoef K, Marzio G, Hillen W, Bujard H, Berkhout B. 2001. Strict control of human immunodeficiency virus type 1 replication by a genetic switch: Tet for Tat. J Virol 75:979–987.
17. Zhou X, Vink M, Berkhout B, Das AT. 2006. Modification of the Tet-On regulatory system prevents the conditional-live HIV-1 variant from losing doxycycline-control. Retrovirology 3:82.
18. Zhou X, Vink M, Klaver B, Berkhout B, Das AT. 2006. Optimization of the Tet-On system for regulated gene expression through viral evolution. Gene Ther 13:1382–1390.
19. Berkhout B, Silverman RH, Jeang KT. 1989. Tat trans-activates the human immunodeficiency virus through a nascent RNA target. Cell 59:273–82.
20. Das AT, Harwig A, Berkhout B. 2011. The HIV-1 Tat protein has a versatile role in activating viral transcription. J Virol 85:9506–16.
21. Su L, Kaneshima H, Bonyhadi ML, Lee R, Auten J, Wolf A, Du B, Rabin L, Hahn BH, Terwilliger E, McCune JM. 1997. Identification of HIV-1 determinants for replication in vivo. Virology 227:45–52.
22. Smith SD, Shatsky M, Cohen PS, Warnke R, Link MP, Glader BE. 1984. Monoclonal antibody and enzymatic profiles of human malignant T-lymphoid cells and derived cell lines. Cancer Research 44:5657–60.
23. van der Velden YU, Kleibeuker W, Harwig A, Klaver B, Siteur-van Rijnstra E, Frankin E, Berkhout B, Das AT. 2016. Construction of Nef-positive doxycycline-dependent HIV-1 variants using bicistronic expression elements. Virology 488:96–107.
24. Fang Q, Inanc B, Schamus S, Wang XH, Wei L, Brown AR, Svilar D, Sugrue KF, Goellner EM, Zeng X, Yates NA, Lan L, Vens C, Sobol RW. 2014. HSP90 regulates DNA repair via the interaction between XRCC1 and DNA polymerase ß. Nat Commun 5:5513.
25. Campeau E, Ruhl VE, Rodier F, Smith CL, Rahmberg BL, Fuss JO, Campisi J, Yaswen P, Cooper PK, Kaufman PD. 2009. A versatile viral system for expression and depletion of proteins in mammalian cells. PLoS One 4:e6529.
26. Markusic D, Oude-Elferink R, Das AT, Berkhout B, Seppen J. 2005. Comparison of single regulated lentiviral vectors with rtTA expression driven by an autoregulatory loop or a constitutive promoter. Nucleic Acids Res 33:e63.
27. Das AT, Zhou X, Metz SW, Vink MA, Berkhout B. 2016. Selecting the optimal Tet-On system for doxycycline-inducible gene expression in transiently transfected and stably transduced mammalian cells. Biotechnol J 11:71–9.
28. Liu Y, Binda CS, Berkhout B, Das AT. 2023. CRISPR-Cas attack of HIV-1 proviral DNA can cause unintended deletion of surrounding cellular DNA. J Virol 97:e01334–23.
29. Dekker JG, Klaver B, Berkhout B, Das AT. 2023. HIV-1 3'-Polypurine Tract Mutations Confer Dolutegravir Resistance by Switching to an Integration-Independent Replication Mechanism via 1-LTR Circles. J Virol 97:e0036123.

9. ACKNOWLEDGMENTS

We would like to express our sincere gratitude to the members of the project guidance committee for their invaluable expertise, insightful feedback, and constructive discussions, which helped us to shape and refine this study.