**Citation: Molecular Therapy—Nucleic Acids (2014) 3, e213;** doi:10.1038/mtna.2014.65

www.nature.com/mtna

# Comparison Between Several Integrase-defective Lentiviral Vectors Reveals Increased Integration of an HIV Vector Bearing a D167H Mutant

**Muhammad Qamar Saeed[1,2,3], Noelle Dufour[1,4], Cynthia Bartholmae[5], Urzula Sieranska[6], Malaika Knopf[5], Eloïse Thierry[7], Sylvain Thierry[7], Olivier Delelis[7], Nicolas Grandchamp[8], Héloïse Pilet[8], Philippe Ravassard[9], Julie Massonneau[1,2], Françoise Pflumio[6], Christof von Kalle[5], François Lachapelle[1,2], Alexis-Pierre Bemelmans[1,4], Manfred Schmidt[5] and Ché Serguera[1,2]**

**HIV-1 derived vectors are among the most efficient for gene transduction in mammalian tissues. As the parent virus, they carry out vector genome insertion into the host cell chromatin. Consequently, their preferential integration in transcribed genes raises several conceptual and safety issues. To address part of these questions, HIV-derived vectors have been engineered to be nonintegrating. This was mainly achieved by mutating HIV-1 integrase at functional hotspots of the enzyme enabling the development of streamlined nuclear DNA circles functional for transgene expression. Few integrase mutant vectors have been successfully tested so far for gene transfer. They are cleared with time in mitotic cells, but stable within nondividing retina cells or neurons. Here, we compared six HIV vectors carrying different integrases, either wild type or with different mutations (D64V, D167H, Q168A, K186Q+Q214L+Q216L, and RRK262-264AAH) shown to modify integrase enzymatic activity, oligomerization, or interaction with key cellular cofactor of HIV DNA integration as LEDGF/p75 or TNPO3. We show that these mutations differently affect the transduction efficiency as well as rates and patterns of integration of HIV-derived vectors suggesting their different processing in the nucleus. Surprisingly and most interestingly, we report that an integrase carrying the D167H substitution improves vector transduction efficiency and integration in both HEK-293T and primary CD34+ cells.**

*Molecular Therapy—Nucleic Acids* (2014) **3**, e213; doi:10.1038/mtna.2014.65; published online 2 December 2014
**Subject Category:** Gene insertion, deletion & modification Gene vectors

## Introduction

Among a wide variety of viral vectors developed for gene transfer, those based on HIV-1 are highly prized for their efficiency and compliance. Extensive amount of work on HIV favored the early development of nonreplicating vectors, which are able to transduce cells at any stage of the cell cycle.[1] Reverse transcription of the viral RNA genome into a double-strand DNA and its integration into cell chromatin are characteristic events of all retroviruses including HIV-1 and their derived vectors. These events are mediated by the viral enzymes reverse transcriptase (RT) and integrase (IN) through coordinated interactions with cellular factors.[2] Within capsids, as part of the reverse-transcription complex and then of the preintegration complex (PIC), ability of IN to interact with RT, or other viral and cellular proteins, and binding to the viral DNA is central for effective infection.[2,3]

Enzymatic activity of IN consists in two reactions modifying the viral DNA, the 3′ processing of long terminal repeat (LTR) extremities and the pairing of the processed LTR termini to chromosomal DNA.[4] In addition, ability of IN to sequentially multimerize is mandatory for proper interaction with viral and cellular factors involved in viral genome processing and integration.[3] The proteins Lens Epithelium-Derived Growth Factor (LEDGF/p75) and the karyopherin transportin 3 (TNPO3) are two important cellular factors interacting with IN that facilitate integration of the viral DNA.[3,5–7]

This sequence of events ensures viral DNA integration in cellular transcription hotspots but may as well cause deleterious genes disruption.[3] This causes setbacks for gene therapy applications and has stimulated research either to target vector integration toward desired loci[8] or to prevent integration through the development of integrase-defective lentiviral vectors, also called nonintegrating lentiviral vectors (NILV).[9,10]

During HIV infection, only about half of nuclear HIV genomes integrate, while the rest remains as linear or circular DNA.[11] The rate of circle formation is increased when integration is impaired by modifying IN activity, its substrate the viral LTRs, or reducing the availability of its cellular cofactor LEDGF/p75.[10,12,13] Among these strategies, using IN mutants is the most efficient to prevent HIV DNA integration.[14,15] Few mutations on this protein preclude integration only (class I

The first two authors contributed equally to this work.

[1]CEA, DSV, I[2]BM, Molecular Imaging Research Center (MIRCen), Fontenay-aux-Roses, France; [2]INSERM, UMS27, Fontenay-aux-Roses, France; [3]Institute of Pure and Applied Biology, Microbiology Division, Bahauddin Zakariya University Multan, Pakistan; [4]Laboratoire des maladies neurodégénératives CNRS, CEA URA 2210; [5]Department of Translasional Oncology, National Center for Tumor Diseases (NCT) and German Cancer Research Center (DKFZ), Heidelberg, Germany; [6]CEA, Laboratoire des Cellules Souches Hématopoïétiques et Leucémiques, (IRCM), Fontenay-aux-Roses, France; [7]EBIM, Ecole Normale Supérieure de Cachan, Cachan, Paris, France; [8]NewVectys, Villebon-sur-Yvette, France; [9]ICM-BBL, CNRS UMR 7225, Hôpital Pitié-Salpêtrière, Paris, France Correspondence: This work was performed in Fontenay-aux-Roses, France, and Heidelberg, Germany. Che Serguera, CEA/INSERM UMS27, 18 route du Panorama, Fontenay-aux-Roses, 92265, France. E-mail: che.serguera@cea.fr

**Keywords:** CD34+ progenitors; integrase; integration sites; lentiviral vectors; residual integration; transduction
Received 13 February 2014; accepted 9 October 2014; published online 2 December 2014. doi:10.1038/mtna.2014.65

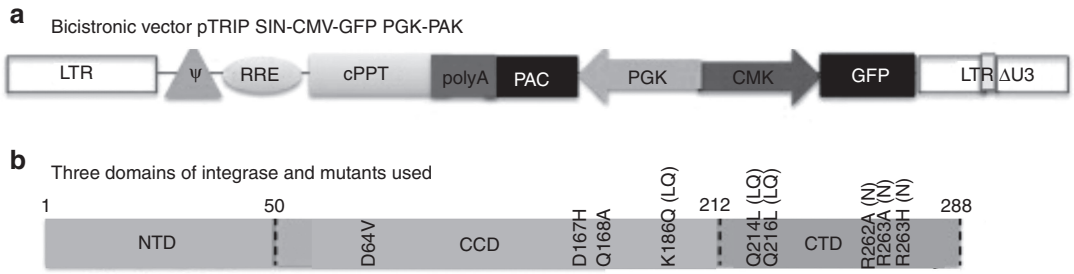


**Figure 1 Schematic representation of the vector and different integrases used in our experiments.** (**a**) We used a bicistronic self-inactivating (SIN) pTrip vector expressing the tracer GFP and the resistance gene puromycin N-acetyl transferase (pac) under control of promoters CMV and PGK, respectively (pTrip PGK-Pac/CMV-Gfp). Only *cis*-acting sequences of HIV-1 are present in the vector, *i.e.,* the long terminal repeat (LTR), the Rev responsive element (RRE), the encapsidation signal Ψ, and the central poly-purine tract (cPPT). *Enhancer* from U3 of 3′LTRs is deleted, creating a self inactivating vector (SIN) upon second strand synthesis. (**b**) Representation of the different integrase domains: N-terminal domain (NTD) spans the first 49 amino acids; amino acids 50–212 form the catalytic core domain (CCD); and the C-terminal domain (CTD) extends from amino acid 213 to 288. Mutations introduced for our study are highlighted. Aspartate-64, implicated in catalytic activity, is replaced with valine (D64V). Aspartate-167 and glutamine-168 involved in interaction with LEDGF are respectively replaced with histidine (D167H) and alanine (Q168A) in two different constructs. In IN-LQ, lysine-186 is replaced with glutamine (K186Q) in L-region and glutamine-214 and 216 are substituted with two leucines (Q214L and Q216L). N region is mutated by replacing 262-arginine-arginine-lysine-264 with two alanines and one histidine (RRK262-4AAH). (Reprinted from ref. 9).

mutations), while others also affect additional steps of HIV processing (class II mutations).[16]

The three functional domains of HIV-1 IN are, a zinc binding N-terminal domain, a $Mg^{2+}$-dependent central catalytic core domain, and a C-terminal domain that binds DNA (**Figure 1b**). Typical class I mutations are those modifying any of the three amino acids DDE (D64, D116, and E152) of the catalytic triad.[16] Of these mutations, the D64V substitution is the most used in studies with NILV.[14,15] The other mutations studied here have been less documented or are studied for the first time. However, they are all expected to induce a class II phenotype of variable intensity but share the characteristic to potentially abrogate IN interactions with cellular factors necessary for PIC nuclear entry and viral integration. It is the case of the D167 and the Q168 residues, embedded within a stretch of residues (161–173) with nuclear localization signal properties[17] and important for interaction with LEDGF/p75 (ref. 18) and TNPO3 (ref. 19). It is also the case of the triple substitutions LQ (K186Q, Q214L, and Q216L) or N (R262A, R263A, and K264H), either impairing IN binding to importin-α1 (ref. 20) or TNPO3 (refs. 5,6) and resulting in lack of HIV DNA integration.[13] We documented the respective impact of these mutations on the integration rate of vector genomes and analyzed their association to a shift in integration site (IS) preference. Moreover, the surprising hyperintegrative phenotype of a D167H mutant IN is further studied in more detail.

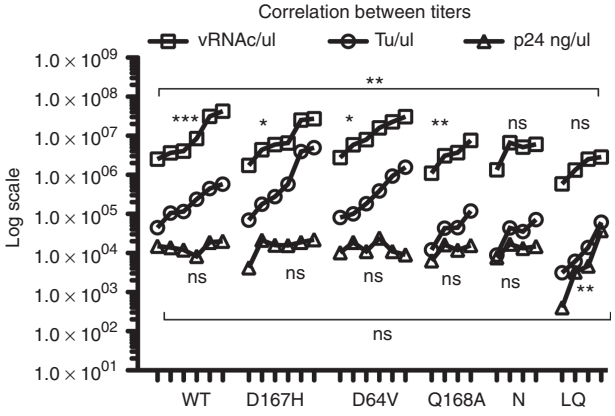


**Figure 2** Tittering methods are compared to define which of viral RNA copies/µl or ng of p24/µl measurements are better correlated to transducing units (TU/µl) values using different stocks of a pTrip PGK-Pac/CMV-Gfp vector carrying wild type (WT) or mutated (D167H, D64V, Q168A, N, and LQ) integrases. Computed data of all vector stocks indicate that only vRNAc/µl are statistically correlated to TU/µl ($P = 0.006$) while ng of p24/µl is not ($P = 0.29$) as assessed through Pearson's correlation. Independent comparison of these variables in each group of vector indicates that viral RNA variation is correlated to that of TUs for most types of vectors but two (N and LQ), while that of p24/µl reaches significance only for vectors LQ as assessed through Pearson's correlation.

## RESULTS

### Vectors tittering

We checked whether IN mutations D64V, D167H, Q168A, LQ, or N affect functional (transducing units (TU)) or physical (capsid protein p24 or vector RNA genome (vRNAg)) properties in HIV-1 vector particles. We compared 30 preparations of particles carrying the vector pTrip PGK-Pac/CMV-Gfp (**Figure 1a**) with wild-type (WT) or mutant IN proteins (**Figure 1b**). Concentration of p24 was equivalent for all vectors indicating that these mutations do not affect capsids formation (**Supplementary Figure S1b**). The amount of vRNAc/ml tended to be reduced in vectors Q168A, N, and LQ as compared with WT, D167H, or D64V, but this difference was not statistically significant (**Supplementary Figure S1b**). Then, although TU/ml was higher for vector with D167H IN as compared with all other vectors, this difference was not significant (**Supplementary Figure S1c**). We then compared if the ratios (TU/p24) and (TU/vRNAc) were different between vectors, and only TU/vRNAc was significantly higher for D167H vector, than that of all other vector types (D167H versus: WT, D64V, Q168A, N, or LQ; $P < 0.05$; **Supplementary Figure S1d,e**). This indicates that the compared mutations do not fundamentally affect particle production but point, as previously shown,[21] to a better correlation between TU/ml and

**Table 1** Biochemical effects of some of the compared IN mutations

| Δaa | RT | LEDGF binding | TNPO3 binding | Oligom. | 3′ p | DNA st | Nuc. tr. | Integrat. | Infect | C I/II |
|---|---|---|---|---|---|---|---|---|---|---|
| D64N | ↓[35] | →[18] | ND | →[35] | ↓↓↓[35] | ↓↓↓[35] | ND | ND | ↓↓↓[18] | I |
| D64E | →[46] | | | | | | →[47] | ↓↓[47] | ↓↓↓[46] | I |
| D64V | →[48] | | | | | | ND | ↓↓↓[46] | ↓↓[48] | I |
| | | | | | | | →[48] | ↓↓↓[48] | | |
| D167K | ↓[35] | ↓[18] | ND | →[35] | ↓[35] | ↑[35] | ND | ND | ↑[18] | II |
| D167A | | ↓[18] | ND | ND | ND | ND | ND | ↓[49] | ↓↓[18] | II |
| Q168A | ↓[35] | ↓[18] | ND | ↓ ↓[35] | →[35] | ↑[35] | →[36] | ↓↓[36] | ↓↓[18,36] | II |
| | | ↓↓[36] | | | | | | | | |
| K186Q | ↓[35] | ↓[35] | ND | ↓ ↓[35] | ↓[35] | ↓↓[35] | ↓[13] | ↓↓[13] | ↓↓[13] | II |
| | | | | ↓[13] | | | | | | |
| Q214/216L | ↓[35] | ↓↓↓[35] | ND | →[35] | ↓↓[35] | ↓↓[35] | ↓[13] | ↓↓[13] | ↓↓[13] | II |
| | | | | ↓[13] | | | | | | |
| K186Q/Q214/216L | ND | ND | ND | ↓↓[13] | ND | ND | ↓[13] | ↓↓[13] | ↓↓[13] | II |
| RRK262/64AAH | ↓[37] | →[5,6] | ↓↓[5,6] | →[5,13] | ND | ND | →[13] | ↓↓[13] | ↓↓[13] | II |

Note that the type of amino acid substitution may have different effect on IN kinetic.
Δaa, amino acid substitution; CI/II, class I or II mutation of IN; DNA st, DNA strand transfer; Infect, infectiosity; Integrat, integration; ND, not determined; Nuc. tr., nuclear translocation; Oligom., oligomerization of IN; 3′ p, 3′-processing of HIV proviral extremities; RT, reverse transcription; →, normal; ↑, increased; ↓, decreased ≤50%; ↓↓, decreased ≥50%; ↓↓↓, abolished.

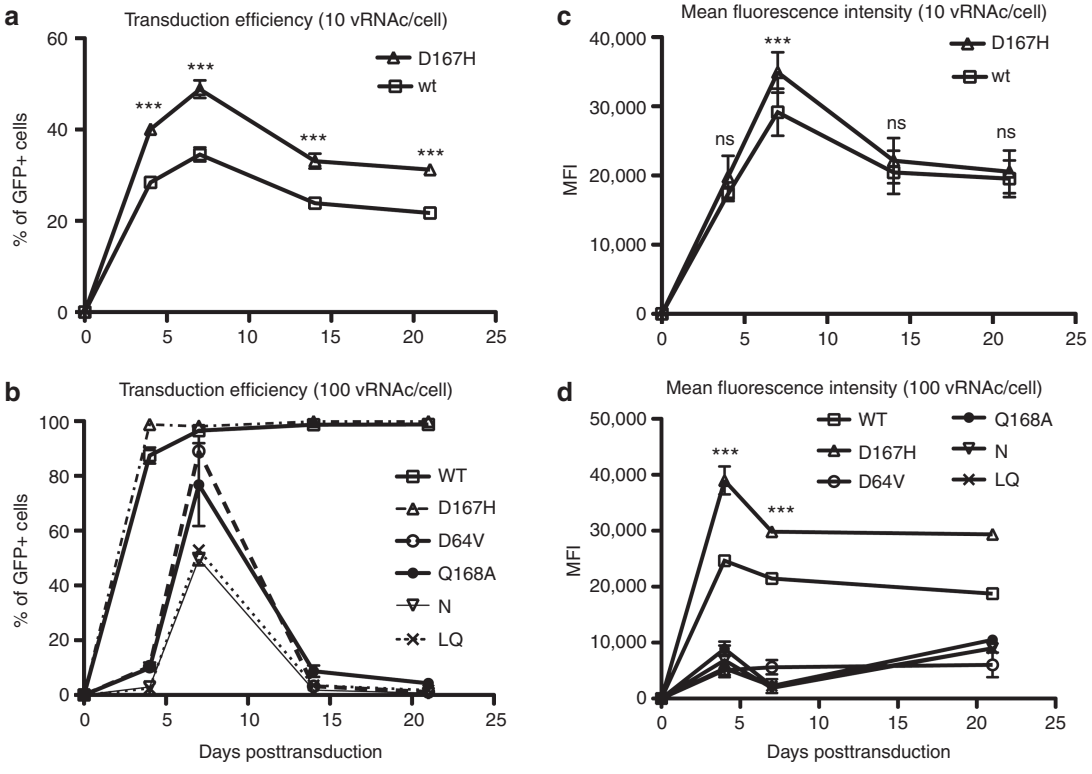


**Figure 3 Comparison of GFP expression in HEK-293T cells transduced with the same MOI of pTrip PGK-Pac/CMV-Gfp vectors carrying different IN mutations.** (**a**) Using 10 vRNAc per cell the mutant D167H allows to transduce more cells compared with WT vector. The equivalent decrease of GFP expression observed 1 week after transduction for both vectors is interpreted as the clearance of circles across cells divisions. (**b**) Mean fluorescence intensity (MFI) is higher in cells transduced with D167H as compared with that transduced with WT only at day 7 when the rate of transduced cells is above 30%, when the probability of multiple copies per cell increases. (**c**) At a higher dose of 100 vRNAc per cell, transduction efficiencies increased with all vectors. (**d**) Although both WT and D167H can transduce nearly 100% of the cells, MFI is statistically higher with D167H at all time points. For vectors D64V, Q168A, N, and LQ, MFI remains within equivalent values. Higher values of MFI at later time points, although not significant, possibly reflect increased GFP expression from integrated copies of vectors. Experiments were done at least three times in duplicate with different stocks of vectors. Error bars: mean + SEM. Statistics: One-way ANOVA and Tuckey's *post hoc* test; ns = $P > 0.05$; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

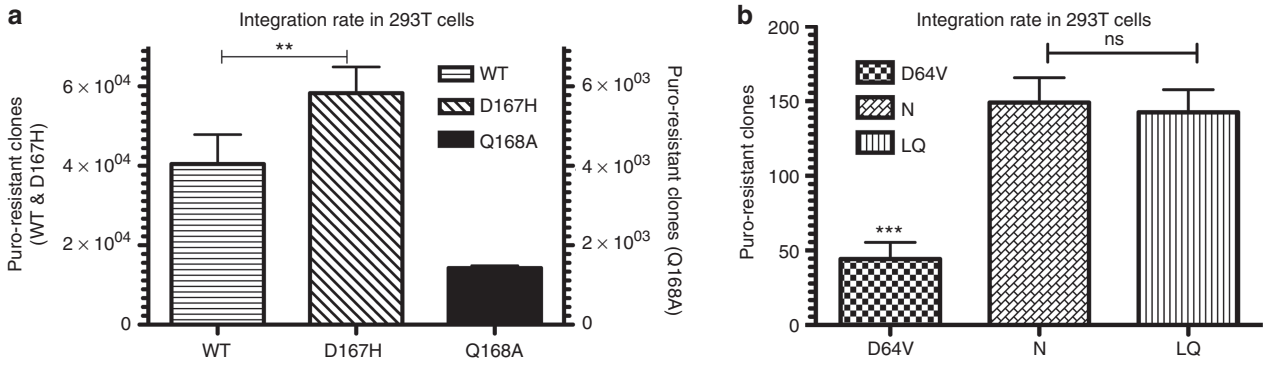


**Figure 4 Integration rates of HIV vectors pTrip PGK-Pac/CMV-Gfp carrying IN with different mutations in HEK-293T cells. Cells were incubated with an amount of each vector allowing 30% cell transduction. After 4 days, cells were cultured in presence of puromycin to select integration events.** (**a**) D167H mutant vector allowed more integrating events than WT ($P < 0.01$) and both WT and D167H led to more colonies than vectors with Q168A IN ($P < 0.001$). Note that the right scale bar is 10 times less than the left one and reflects the number of colonies obtained after transduction with the vector with Q168A IN. (**b**) Among vectors with a nonintegrating phenotype, D64V mutant was the least integrating ($P < 0.001$). All experiments were done four times in duplicate and with different stocks. Nontransduced cells all died in puromycin medium, and all resistant clones were also positive for GFP expression. Error bars: mean + SEM. Statistics: one-way ANOVA and Tuckey's multiple comparison test; ns = $P > 0.05$; *$P < 0.005$; **$P < 0.01$; ***$P < 0.001$.

vRNAc/ml, than between TU/ml and p24/ml (**Figure 2**). Consequently, in the following experiments, we used vRNAc/ml to normalize the amount of particles for cell transduction and comparisons of vector efficiencies for gene transfer. These data also indicate that after normalization of TU/ml with vRNAc/ml, D167H mutant vectors present a significant advantage for cell transduction over the other vectors including WT.

### Transduction efficiencies of IN mutant vectors

To study the impact of several IN mutations on HIV vector transdcution efficiency, we measured green fluorescent protein (GFP) expression in 293T cells transduced with a bicistronic pTrip PGK-Pac/CMV-Gfp vector (**Figure 1a**) shipped by six HIV-1 vector particles containing either WT or mutated IN (**Figure 1b**). Most of the IN mutations compared here have been shown to produce different effects on enzyme biochemistry and virus processing (**Table 1**) and should have variable effects on transduction efficiency of vectors.

Mutations D64V, N, LQ, and Q168A affect HIV integration[9,10,22,23] but have never been directly compared. The nonconservative D167H substitution is studied for the first time. It was chosen to alter D167-K360 interaction between IN and LEDGF/p75 (ref. 24) and compare its effects on vector transduction to that of mutant Q168A.

HEK-293T cells were incubated with lower (10 vRNAc/cell) or higher (100 vRNAc/cell) amounts of WT or mutant vectors. Cells were analyzed by flow cytometry at 4, 7, 14, and 21 days after transduction to estimate the percentage of GFP-expressing cells and the mean fluorescence intensity (MFI) as a relative index of vector copy number per cell. With 10 vRNAc/cell, WT and D167H vectors allowed stable GFP expression over time (**Figure 3a**), indicating that both integrate in cell chromatin. The four mutants (D64V, Q168A, N, and LQ) displayed a nonintegrating phenotype (NILV) with GFP clearance over time (**Supplementary Figure S2a**). Surprisingly, across several experiments using different stocks of vectors, we observed that D167H vector allowed a 1.5 times higher transduction efficiency than WT vector, a significant increase noted at all time points of analysis (**Figure 3a**). Among the four NILV, transduction efficiency with D64V was two to six times higher than that observed with other mutants. Mutant Q168A was the second most efficient, followed by N and LQ (**Supplementary Figure S2a**). Using 10 vRNAc/cell, MFI was slightly but significantly higher after 7 days in cells transduced with D167H as compared with WT (**Figure 3c**). However, at later time points, when the percentage of transduced cells fall to ≤30%, a situation where the mean amount of vector copy/transduced cell equals 1, MFI was equivalent for both vectors (**Figure 3c**). All NILV permitted equivalent MFI that was about five times lower than that of integrating vectors (**Figure 3d**) indicating a less efficient transcription by HIV-1 vector circles compared to integrated vectors.[14,25]

At the higher dose of 100 vRNAc/cell, WT and D167H vectors transduced 100% of the cells (**Figure 3b**), but D167H vector permitted a 1.5 times higher MFI (**Figure 3d**). Cells transduced with higher input of D167H vector (300 and 900 vRNAc/cell) also displayed an equivalent increased MFI as compared with WT vector, indicating a nonsaturating gain of function (**Supplementary Figure S2b**). At higher dose of 2,700 vRNAc, no difference was observed, possibly due to toxic effect of overtransduction. Transduction efficiency of nonintegrating mutants with 100 vRNAc/cell was much closer to that of WT and D167H (**Figure 3c**). At greater doses of 300 vRNAc/cell, the four NILV allowed transduction of about 100% of the cells between days 4 and 7 after transduction (**Supplementary Figure S2c**).

Thus, D167H mutant vector has an integrating phenotype and appears improved for transducing HEK-293T cells compared with WT. Of the four nonintegrating vectors, the one carrying the D64V mutation displays the best transduction efficiency.

### Residual integration of LV with different IN mutations

To further characterize the phenotype of these vectors, we measured their respective ability for integration within the cell chromatin. For this, we measured puromycin resistance as a proof of stable expression of the puromycin N-acetyl transferase (PAC) transgene.

To compare the rate of integration of these vectors, we worked with populations of cells containing the same amount of nuclear vector DNA. A statistical model based on Poisson distribution indicates that when at most 30% of a cell population is infected or transduced, each targeted cell contains one copy of the virus.[4] This is true in practice for lentiviral vectors.[26] Thus, for each stock of vector, we first defined the functional titer in TU/ml. We then used appropriate amount of particles to allow transduction of 30% of HEK-293T cells and verified it by flow cytometry after 4 days. Four days after transduction, puromycin was added to the culture medium to select cells having integrated a copy of pTrip PGK-Pac/CMV-Gfp vector.

After 10 days of selection, we counted a higher number of clones for WT and D167H vectors than that for NILV ($P < 0.001$; **Figure 4a**,**b**). In these conditions, we also observed that the D167H mutant produced about 1.5 times more colonies than a vector with WT IN ($P < 0.001$; **Figure 4a**). The D64V mutant was the least integrative giving about 700 times less colonies than WT and about 5 times less colonies than N and LQ mutants ($P < 0.001$). Mutant Q168A had an intermediary phenotype and was about 15 times less integrative than WT and 10 to 50 times more integrative than mutants LQ and N or D64V ($P < 0.001$).

Thus, of all the vectors compared here, the one containing the D167H mutant IN is the most integrating, while the one containing the D64V mutation is the one allowing the least integration events.

### IN D64V + D167H allows more integration events than IN D64V

To further study if D167H mutation acts on catalytic activity of IN or through any other mechanism, we associated it with the D64V substitution that abolishes IN activity. We thus introduced the two mutations in the IN sequence of the plasmid p8.91 (p8.91-D64V+Q167H). We transduced HEK-293T cells with 50 vRNAc of particles either carrying the single D64V or the double D64V+Q167H mutant IN. Both vectors displayed equivalent transduction efficiency (**Figure 5a**) and a similar nonintegrating phenotype with loss of GFP expression through cell division (**Figure 5b**). Surprisingly, when we measured integration, by selecting puromycin-resistant clones starting at day 4 after transduction, we observed that the double mutant D64V+Q167H led to a significant increase of clones (1.5 times more) as compared with the mutant D64V (**Figure 5c**). This suggests that the D167H mutation introduces a change in a NILV vector toward a form more prone to integration.

### Kinetics of integration of vectors carrying the D167H mutation

To check whether the D167H mutant changes the kinetic of integration of HIV vectors, we transduced HEK-293T cells with WT, D167H, D64V, or D64V+D167H vectors and correlated efficiency of transduction (GFP expression) to integration (puromycin resistant clones) with time (0, 6, 12, 24, 48, and 72 hours). When comparing integration of D64V and D64V+D167H in HEK-293T cells, we observed that the more we progressed in time posttransduction, the higher the difference between the two vectors. At multiplicity of infection

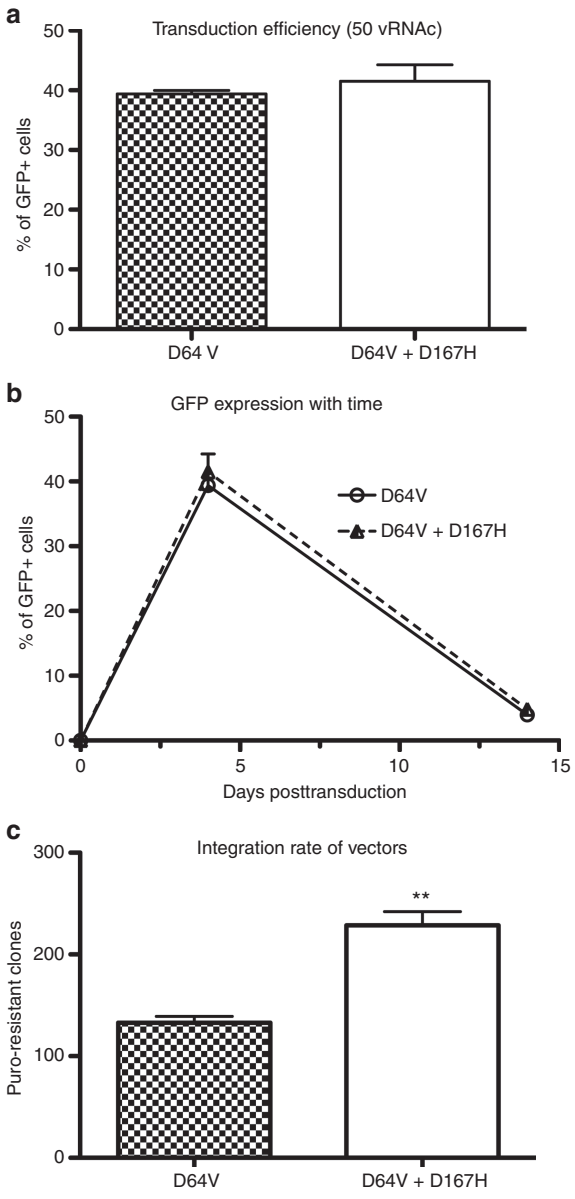


**Figure 5 Comparison of efficiency of transduction and integration, of D64V and D64V + D167H mutant vectors pTrip PGK-Puro/CMV-Gfp (50 vRNAc/cell).** (**a**) Transduction efficiency of 293T cells is equivalent with both vectors. (**b**) Both vectors display a nonintegrating phenotype as GFP expression decreases with time in dividing 293T cells. (**c**) Cells transduced with vectors carrying the IN D167H + D64V display significantly more integrating events as compared with cells transduced with D64V vectors. Experiments were done in triplicate and with two different stocks of vectors for each mutant. Nontransduced cells all died in puromycin medium. Error bars: mean + SEM. Statistics: *t*-test, ns = $P > 0.05$; $^{*}P < 0.005$; $^{**}P < 0.01$; $^{***}P < 0.001$.

(MOI) 50 or 100 vRNAc, integration of D64V vectors mainly occurred between 0 and 48 hours after transduction then seemingly reaching a plateau. Integration of D64V+D167H, instead, kept increasing linearly until the last time point analyzed at 72 hours after transduction (**Figure 6a**,**b**). The kinetics of integration of WT and D167H vectors were also different. The former tended to integrate between 0 and 48

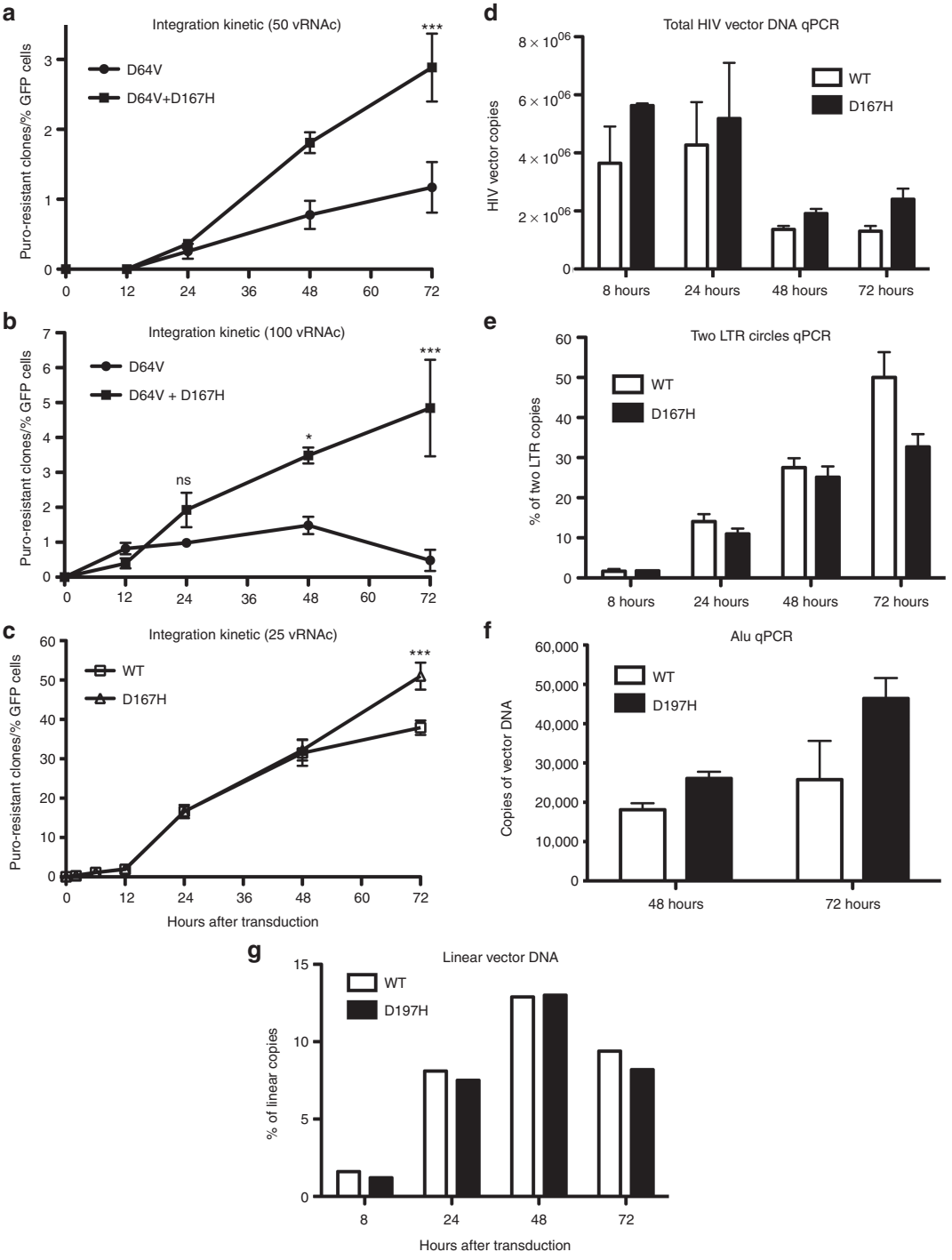


**Figure 6 Kinetics of integration assessed through puromycin resistance and real-time qPCR measurement of vectors pTrip PGK-Puro/CMV-Gfp carrying a D167H IN. Puromycin was added 12, 24, 48, or 72 hours after transduction, and integration was rated as the number of puromycin-resistant colonies normalized with respective percentage of GFP expression. (a,b) 293T cells transduced with vectors carrying a D64V or D64V + D167H IN.** (**a**) Integration rate at each time point after cells transduction with 50 vRNAc or (**b**) 100 vRNAc. (**c**) 293T cells transduced with vectors carrying a WT or a D167H IN. Graph shows vector integration rate at each time point after cell transduction with 25 vRNAc. Nontransduced cells all died in puromycin medium. (**d**–**g**) Real-time qPCR measurement of total vector DNA, two LTR circles, integrated vector DNA, and linear vector DNA at different times after transduction. (**d**) Total vector DNA between 8 and 72 hours after transduction to assess reverse transcription. (**e**) Percentage of two LTR circles at different times after transduction as an inverse correlate of integration. (**f**) Integrated vector DNA normalized with total vector DNA. Note that WT vector does not further integrate between 48 and 72 hours while the amount of integrated vector D167H keeps increasing between 48 and 72 hours. (**g**) Percentage of linear copies of vector DNA. For kinetics of puromycin resistance, experiments were done in triplicate with two different stocks for each vector. Real-time PCR experiments were done in duplicate. Error bars: mean + SEM. Statistics: two-way ANOVA and Bonferroni posttest, ns = $P > 0.05$; *$P < 0.005$; **$P < 0.01$; ***$P < 0.001$.

hours after transduction, as reported before,[11,27] while the mutant D167H kept integrating between 48 and 72 hours after transduction (**Figure 6c**).

This trend was further confirmed with real-time Alu-LTR PCR, which showed that integrated copies of WT vectors remained equivalent at 48 and 72 hours while they kept increasing for vector D167H (**Figure 6f**). At 72 hours after transduction, integrated copies of D167H vectors were about 1.7 times higher than that of WT vectors.

We also measured the proportion of linear vector genomes, which are the substrate for integration. We observed that these forms equally decreased between 48 and 72 hours after transduction for both D167H and WT vectors (**Figure 6g**). This indicates that a significant proportion of D167H linear vectors still integrated at later time points, while during this same period, WT linear vectors were lost. In addition, we also measured the percentage of two LTR circles that are usually inversely associated with integration.[11,27] Interestingly, these forms of HIV vector genomes were decreased for D167H vectors as compared with WT at 72 hours after transduction (**Figure 6e**). These results indicate that IN D167H substitution lengthens the kinetics of integration of HIV vectors.

### LV with D167H substitution improves CD34+ cell transduction

Lentiviral vectors have been used successfully to genetically modify hematopoietic progenitors and treat congenital immunodeficiency.[28] We asked whether HIV vectors bearing a D167H IN could improve the transduction efficiency of human hematopoietic CD34+ cells compared with WT vector. Primary human cord blood CD34+ cells were incubated with equal vRNAc of D167H and WT vectors. After 4 days, cells were analyzed by flow cytometry to assess the percentage of transduction. Particles carrying the D167H IN allowed about two times higher transduction of CD34+ cells compared with WT particles (**Figure 7**). This was true for both low and high vector doses. Altogether, these results indicate that the effect of the D167H substitution in improving transduction efficiency is not restricted to HEK-293T cells.

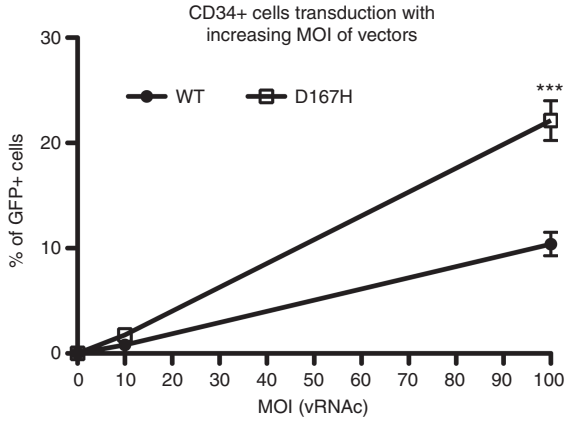


**Figure 7** Human hematopoietic CD34+ cells were transduced with 10 or 100 viral RNA copies of WT or D167H vectors. Transduction efficiency was assessed by using flow cytometry to determine the percentage of GFP expressing cells 4 days after transduction. Experiments were done twice on independent CD34+ cord blood samples. Error bars: mean + SEM. Statistics: *t*-test; ns = $P > 0.05$; *$P < 0.005$; **$P < 0.01$; ***$P < 0.001$.

### Integration profile of LV with different IN mutations

To further compare the properties of mutants and WT vectors, we assessed their respective pattern of integration in HEK-293T cells with nonrestrictive LAM-PCR.[29] Cells were transduced with increasing input of vectors (100, 300, 900, and 2,700 vRNAc/cell), and integration events were isolated with puromycin selection. Resistant clones were counted (**Supplementary Table S1**), pooled, and further cultured before genomic DNA purification. To analyze the pattern of integration of WT and D167H vectors, we used genomic DNA of colonies obtained with the lower vector input (100 vRNAc/cell). For NILV, leading to fewer integration events, we pooled the DNA of all colonies obtained with the different doses of vectors (**Figure 8** and **Supplementary Figure S3**). Nonrestrictive LAM-PCR products were then sequenced using Illumina MiSeq technology, and flanking genomic sequences were characterized by bioinformatical data mining.[30] As expected, the number of IS obtained from samples transduced with NILVs was lower than from samples transduced with an integrating vector (**Supplementary Table S1**). Vector carrying IN D167H, instead, gave ~1.4 times more IS than WT vector. The lower frequency of IS in cells transduced with NILVs was most prominent for D64V, N, and LQ in comparison to data obtained with Q168A vector, which is in line with the higher number of clones obtained after transducing HEK-293T cells with this vector (**Figure 4** and **Supplementary Figure S3**). Previous analysis of integration patterns of lentiviral vectors showed that genomes integrated by a D64V IN suffer about 10 times more deletions within LTRs than that of vector genomes integrated by a WT IN[30,31]. Here, we also observed that IN mutations impairing integration produced an increased number of LTR deletions (Q168A, 1.8%; LQ, 16.8%; D64V, 21.6%; N, 44.7%), while D167H and WT IN displayed the same low frequency of deletion (0.3%; **Supplementary Table S2**). Differences in LTR deletion frequency observed with the different vectors suggest that integration proceeds through mechanisms with variable involvement of IN or IN-interacting factors.

We then characterized the localization of IS of the different vectors within chromosomes. Previous studies showed that HIV and derived vectors integrate preferentially within

**Table 2** Performance rating between WT and D167H vectors IN

| | TE at day 4 with 10 vRNAc | IR in 30% of transduced cells | MFI at day 7 with 100 vRNAc | ICN measured with Alu-PCR | # of IS with 100 vRNAc | # of IS in CIS of higher order | Mean IS # in top 10 CIS |
|---|---|---|---|---|---|---|---|
| WT | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| D167H | 1,42 | 1,44 | 1,42 | 1,7 | 1,38 | 1,8 | 1,6 |

The numbers in the row "D167H" reflect the proportion of increase in transduction efficiency (TE), integration rate (IR), vector copy number (VCN), the amount of transgene expression as assessed through the measurement of mean fluorescence intensity (MFI), the number of integration sites, or the number of integration within common integration sites (CIS), in HEK-293T cells.

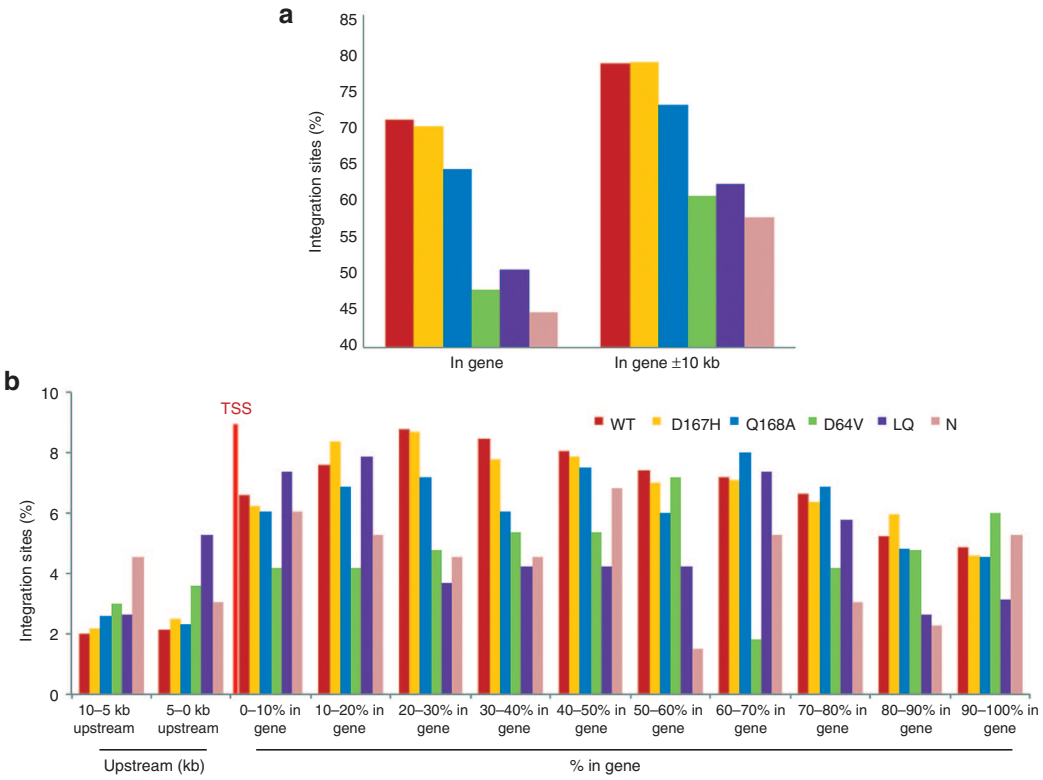


**Figure 8** Integration profile of integration-competent lentiviral vector (bars 1 and 2) and NILV (bars 3, 4, 5, and 6) within or in the vicinity of coding regions of the genome. (**a**) Integration sites (IS) are displayed with respect to frequency of IS located within RefSeq genes or 10 kb surrounding and (**b**) distribution within genes and upstream of genes.

transcribed genes (70%), while this trend is reduced in cells depleted in LEDGF or TNPO3 (refs. 7,32,33) and for vectors carrying a mutant D64V IN[29–31]. Here, we observed that the reduced frequency of integration within gene-dense regions was variable among NILVs as compared to WT and D167H (70% of integration in gene coding regions); it was moderate for Q168A (~65%), but more prominent for the N mutant (~45%), as compared with D64V and LQ (~50%). This indicates that our method for selecting integration events not only allows to isolate a wide range of integration patterns within cells but also points at different alterations of the mechanism of chromatin insertion with these different IN mutations.

### Common integration sites

To uncover integration hotspots, we determined the occurrence and frequency of common integration sites (CIS) in the genome where multiple vector integrations occurred. Several orders of CIS were defined (**Supplementary Table S3)**. The frequency of CIS was in line with the number of IS that could be retrieved for the individual data set. The highest CIS order observed is CIS of the 27th order in the D167H data set to occur on chromosome 8. Also, in the WT data set, the highest CIS of 15th order was identified in the same genomic region of chromosome 8. **Supplementary Table S4**, summarizing the 10 highest CIS orders for individual vector types, shows that the mean of the highest order of CIS is 11.4 for vectors with WT IN but is 1.6 times higher for vectors D167H (18.6) and 2 to 6 times lower for vectors Q168A and other NILVs (**Supplementary Table S4**). This indicates that the overall integration rate of the vector D167H is higher than that of WT, as at equivalent MOI of either vectors (100 vRNAc), D167H leads to an increased copy number within single cell CIS. These data also reveal that in HEK-293T cells, as in human CD34+ cells, typical lentiviral integration hotspot regions like PACS1 are retrieved as CIS of high order,[28,34] but also that particular gene regions, which have not been previously described as lentiviral hotspot, are preferentially targeted, like CIS on chromosome 8 or 16.

## Discussion

In this study, we compared gene transfer properties of HIV-1–derived vectors carrying viral IN with different amino acid substitutions. WT and D64V enzymes represented archetypical IN behaviors, with either optimal or abolished enzyme activity,[10,22,35] in between which a continuum of transduction and integration efficiencies was observed using mutants Q168A, N, or LQ. Interestingly though, IN mutant D167H outperformed WT vector for transduction efficiency and integration.

Viruses with IN D64V, Q168A, N, or LQ mutations are all impaired for replication due to a main defect in integration.[13,18,35–38] The phenotype of HIV-1–derived vectors carrying D64V, Q168A, or N substitutions have been individually described[9,10,22,23] but never compared under similar experimental conditions. We found that vector with D64V substitution was two to six times more efficient for cell transduction than vectors Q168A, N, or LQ. This is explained by the fact that the three mutations also affect reverse transcription (**Table 1**).

We then observed that these IN mutations differently affected HIV vector integration. For instance, D64V IN, a pure catalytically silent enzyme, induced the least integrative phenotype of all NILV, suggesting that vectors with class II mutations, Q168A, N, or LQ, retain some IN activity, which is concordant with previous *in vitro* studies that analyzed some of the mutations used here (**Table 1**). *In vitro*, catalytic activity of a IN with nonconservative substitutions RRK 262–264 DVE was abolished with substitution K264E having the deeper impact on enzyme activity.[39] However, in our study, IN K264H conservative replacement possibly retains some catalytic activity. This is supported by previous experiments showing abolished or conserved replication of HIV-1 carrying either K264E or K264A IN[37]. Most interestingly, mutations LQ and N possibly also affect integration by impairment of IN interaction with key cellular factors implicated in PIC tethering toward chromatin hotspots of integration. Indeed, residues K186, Q214, and Q216 mutated in LQ conform a nuclear localization signal of IN,[20,38] and their mutation impairs IN interaction with importins,[20,40] affecting integration rather than nuclear translocation.[13] Similarly, IN residue K264 and residues RRK262-4 were found necessary for binding to an unidentified factor needed for integration[35] and to TNPO3,[5,6] a nuclear receptor important in HIV integration within active transcription units.[7] Interestingly, our data indicate that a vector with an N mutant IN has a pattern of integration different from that of the other NILVs.

Regarding chromosomal distribution across the genome, WT, D167H, and Q168A vectors were quite indistinguishable, while vectors D64V, LQ, and N displayed a wide degree of variation. When comparing the divergence of integration rate per chromosome, between WT and NILVs, it appeared that LQ, D64V, and N cumulated 21, 27, and 44% variation, respectively (**Supplementary Figure S4**). This rate varied for ≥2% in 9 chromosomes for D64V and LQ vectors, but in 17 chromosomes concerning vector N. Vector N also displayed the more random pattern of integration (*i.e.,* the closest to 30% of integration within genes) of all NILV and integrated consistently more upstream of genes than the other five vectors. Finally, analysis of LTR borders revealed twice more deletions for the N vector than for LQ and D64V vectors. All together, these data indicate that N mutant display a particular phenotype that could be explained by a lack of interaction with a nuclear factor inducing particular consequences on HIV-1 DNA mechanism and/or place of integration. This factor may likely be TNPO3[5–7] or another IN interacting factor as an importin, KAP1, or EED as suggested by Li *et al.*[35]

Vector with mutation Q168A, that reduces *in vitro* IN affinity for LEDGF[18,19,35,36] and TNPO3,[19] was about 15 times less integrative than WT. Its integration within transcription units was however only slightly reduced (65%) as compared with vectors with WT IN (70%). This contrasts with observations in cells depleted for LEDGF or TNPO3 where HIV integration is much more random.[7,33] It thus appears that the pattern of integration of Q168A vector is largely mediated by a remaining ability to interact with these cellular factors *in vivo*. Its nonintegrating phenotype is then possibly due to an unstable IN oligomer mainly affecting catalysis of integration.[12,18]

Of all vectors analyzed here, the one carrying the D167H IN displayed the most intriguing phenotype with an increased integration rate of about 40% as compared with WT vectors (**Table 2**). Residue D167 of IN was mutated in this study because it is well conserved in HIV[41] and it is embedded within a short stretch of amino acids (161–173) with nuclear localization signal properties[17] and important for IN–p75/LEDGF interaction.[12,18,24] We initially chose histidine to replace a polar acid residue by a polar basic/neutral one with the purpose to induce a mild phenotype with an integration rate between WT and Q168A vectors.

Integrase mutations D167A and D167K induced opposite effect on single round infectivity of HIV with D167A increasing it to 140% while D167K reduced it to 5% of WT[18] indicating a critical role played by this residue in the processing of the PIC. The fact that D167A or K substitutions do not reduce IN enzymatic activity,[18] but slightly affect RT[35] and interaction with p75/LEDGF,[18] points to a change in PIC processing due to a modified trend of interactions between IN monomers, between IN and RT, or between IN and a cellular factor such as p75/LEDGF. Histidine is a versatile amino acid that can assume positive charge, but remains usually neutral at physiological pH. Indeed, in our experiments, vector D167H seem to function as D167A mutant[18] with no impact on early events of transduction as RT (**Figure 6d**). In fact, this mutation affects later time points of our analysis with vector D167H still integrating between 48 and 72 hours following transduction. D167H mutation most likely keeps IN catalytic activity unchanged, but rather affects PIC structural features modifying its interaction with a nuclear protein, increasing its stability, its anchoring to chromatin, and consequently integration of the linear vector genome at later time points. This notion is supported by the fact that a NILV D64V+D167H also demonstrated an increased IN-independent integration as compared to a D64V vector, showing that the D167H substitution provides a gain of function independent of catalytic activity of integrase. Moreover, as an equivalent decrease in the linear forms of vector D167H and WT between 48 and 72 hours after transduction (**Figure 6g**) led on one hand to a continuing integration of the former but not of the latter also supports the view of an increased stability of a PIC carrying a D167H IN.

Although D167H substitution of IN represents a net gain of function for HIV vector integration, it is very unusual in HIV, because it probably affects particles release in a complete viral cycle as showed in viruses carrying IN D167A or K substitutions.[18]

The vectors described here illustrate the central role played by IN in determining the efficiency and place of HIV integration. Most notably, vectors D64V, N, and D167H are of greater interest in gene transfer applications. HIV vectors with IN D167H substitution will be preferred in strategies requiring transgene integration in fast dividing cells, as well as to transduce primary cells with low permeability to LV such as CD34+ hematopoietic stem cells. Alternatively, HIV vectors with mutation D64V or N will be chosen for the situation where vectors with a nonintegrating phenotype are required.

## Materials and methods

*Plasmid constructs.* Five different integrase sequences were introduced into the p8.91 plasmid encoding the gag-pol genes of HIV-1 and initially described in ref. 42. Plasmid p8.91_D64V was a kind gift from DB. Kohn.[22] Plasmid p8.91_N

was described earlier,[9] plasmids p_8.91_LQ, p8.91_Q168A, p8.91_D167H, and p8.91_D64V_D167H were constructed by inserting a SwaI-AflII synthesized DNA fragments (Genescript, Piscataway, NJ) from p8.91 containing a large part of the integrase and including LQ, Q168A, D167H, or D64V and D167H substitutions. These fragments cloned in pUC57 were digested with SwaI-AflII restriction enzymes and cloned in p8.91 linearized with SwaI-AflII enzymes. The pTrip PGK puro polyA/CMV eGFP Delta U3 was generated using pTrip CMV eGFP Delta U3 as backbone[43] in which a PGK_puromycin_polyA cassette was inserted in the reverse orientation on the 5′ side of the cppt sequence. Briefly, a linker containing *Eco*RI-compatible *Sac*II, *Sal*I *Bam*HI, *Mlu*I, and *Eco*RI restriction sites was first inserted in the *Eco*RI site of pTrip CMV eGFP Delta U3. Next, the polyA signal from human β-globin was amplified by PCR from pCDNA3.0 vector (Life Technologies, Carlsbad, CA) with primers containing *Sac*II and *Sal*I overhanging ends and cloned in the inserted linker. The PGK promoter sequence was digested from pTrip PGK eGFP Delta U3 vector,[44] using *Mlu*I and *Bam*HI and cloned in the polyA-containing vector. Finally, the puromycin resistance gene (pac) was amplified by PCR from pLKO puro vector (Addgene, Cambridge, MA) with primers containing *Bam*HI and *Sal*I ends and cloned between the PGK promoter and the polyA sequences.

*Vectors production.* For vector production, we used HEK-293T cells (human embryonic kidney cell line, ATTC–CRL-11268) grown in DMEM (Life Technologies) supplemented with 10% SVF and 1% penicillin and streptomycin. Trancomplementation plasmids expressing structural proteins and enzymes of HIV-1 but deleted of accessory genes (p8.91) and a plasmid expressing the vesicular stomatitis virus envelop glycoprotein (pVSV) were used to produce VSV-G pseudotyped pTrip PGK-Pac/CMV-Gfp self-inactivating vector particles as described in ref. 42. Briefly, batches of HIV-1–derived vectors were produced by transient cotransfection of the plasmids mentioned above in HEK-293T cells cultured in 10 cm Petri dishes, with the $Ca_3(PO_4)_2$ precipitate procedure. Supernatant was harvested after 48 hrs, filtered, and ultracentrifuged at 19,000 rpm for 1.5 hours at 4 °C. The supernatant was discarded, and the pellet was concentrated a 1,000 times in PBS and 1% BSA and then frozen at –80 °C until use.

*Lentiviral vector titration.* To determine recombinant particle content in concentrated supernatants, we used an ELISA assay of p24 antigen (Gentaur, France) according to the manufacturer's instructions. Furthermore, the number of LV RNA copies was tittered by real-time quantitative PCR. Concentrated viral suspension (2 µl) was suspended in 353 µl DNase- and RNase-free water, 5 µl of DNAseI and 50 pg of RNaseA, and incubated for 10 minutes at room temperature before adding 20 µl of RNasin (Promega, Madison, WI). The mixture was incubated for 10 minutes at 37 °C, and DNase was inactivated by addition of 10 µl of 25 mmol/l EDTA and heating at 70 °C for 10 minutes. RNA was extracted with a Pure link RNA mini Kit (Ambion, Carlsbad, CA) according to guidelines. At the end of the procedure, the purified RNA was eluted in 30 µl of the kit solution. Purified RNA (5 µl) was added to 96-well plate for reverse transcription and quantitative PCR with the Express Superscript mix and Express SYBR green ER supermix (Life technologies). As a negative control, the RNA was added to the well without reverse transcriptase. PCR was carried out as follows: after reverse transcription for 5 minutes at 50 °C, 2 minutes at 95 °C, and 40 cycles of 15 seconds at 95 °C, and 1 minute at 60 °C (amplification) with the sense primer 5′-TGTGTGCCCGTCTGTTGTGT-3′, and the antisense primer 5′-GAGTCCTGCGTCGAGAGAGC-3′. The number of RNA copies was deduced from a standard curve for known numbers of copies of DNA plasmid treated in the same way as the samples. The relative titer/ml was calculated as the number of RNA copies/volume in ml × dilution of vector preparation. To determine transducing units/ml of vectors stock, serial dilutions from $10^{-3}$ to $10^{-9}$ of each stock were used to transduced HEK-293T cells ($10^5$ cells/well in a 24 multiwell plate) in triplicate. After 72 hours, cells were analyzed with flow cytometry to determine the percentage of GFP-positive cells, using the Cyflow Space device (Partec, Germany), and data were analyzed with FlowJo software (FlowJO, FlowJO LLC, Ashland Oregon).

*Determination of residual integration.* HEK-293T cells were plated ($10^5$ cells/well in a 24 multiwell plate), and the next day, they were transduced with a MOI of vectors (RNA copy/cell) allowing to transduce 30% of cells. After 4 days, cells were trypsinyzed and resuspended in 1 ml of medium and split into two Petri dishes (10 cm) with 100 and 900 µl, respectively. As integration was assessed through measurement of puromycin resistance and to minimize selecting only cells expressing high levels of PAC, we determined the lowest dose of drug necessary to kill all 293T cells, which was 2 µg/ml. Puromycin (Sigma, St Louis, MO) selection agent was added in the afternoon at a final concentration of 2 µg/ml, and resistant clones were counted after 10 days. The number of clones was normalized with the percentage of GFP expressing cells at 4 days after transduction. In all cases, the percentage of GFP cells was between 25 and 30%. For kinetic of integration, transduction was performed in HEK-293T cells in suspension. For 1 h, $2.5 \times 10^5$ cells/tube were incubated with the vector and then plated in a 10 cm Petri dish. Puromycin was added to the cell media at 2, 6, 12, 24, 48, or 72 hours after transduction. Puromycin-resistant clones were counted after 10 days. The number of clones was normalized with that of percentage of GFP expression at each time point. Nontransduced cells were used as control.

*Transduction efficiency and flow cytometry.* HEK-293T cells were plated ($10^5$ cells/well in a 24 multiwell plate). The next day, cells were transduced with a number of RNA copy/cell allowing to transduce 30% of the cells (residual integration) or particular MOIs of vectors (transduction efficiency). Transduced cells were harvest at different times posttransduction (as stated above), trypsinyzed, and fixed with formaldehyde 1% final. Then, cells were analyzed with flow cytometry to determine the percentage of cells positive for GFP.

*Real-time PCR to quantify kinetics of total and integrated vector DNA.* Real-time PCR was used to quantify different forms of vector DNA in HEK-293T cells at different times after transduction. Cells were transduced with 50 vRNAc of vector

particles carrying a WT or a D167H IN. Transduced cells were harvested at 20 minutes and 8, 24, 48, and 72 hours after transduction, and DNA was purified using a Dneasy Blood and Tissue Kit (Qiagen, Valencia, CA) as recommended by the manufacturer. All samples were done in duplicate. Fifty micrograms of DNA of each samples was used to assess total vector DNA, integrated vector DNA, two LTR circular vector DNA, and linear DNA using quantitative real-time PCR as described before.[45,11] Detailed procedures and materials can be found in **Supplementary Materials and Methods**.

*LAM-PCR and IS analysis.* Ten thousand HEK-293T cells were transduced with vector pTrip PGK-Pac/CMV-Gfp carrying WT or mutants IN (D167H, Q168A, N, or LQ) at MOI of 100, 300, 900, and 2,700 viral RNA copies/cell (vRNAc/cell). Cells were plated at 1/10 or 9/10 in 10 cm plates, and puromycin was added 4 days after transduction (2 µg/ml), and resistant clones were counted after 10 days in selection. This experiment was done in duplicate. After counting, clones of each dish were trypsynized, pooled, and grown as populations until reaching confluence. Genomic DNA of each population was purified using cell lysis with proteinase K (Sigma) and phenol–chlorophorm extraction and ethanol precipitation.

Analysis of integration profile of HIV vectors, carrying either WT or mutant IN in HEK-293T cells, was carried out at the National Center for Tumor Diseases in Heidelberg Germany using nonrestrictive (nr) LAM-PCR[29] coupled to Illumina MiSeq pyrosequencing. Data were analyzed using subsequent semi-automated bioinformatical data mining, including trimming of sequences, alignment using UCSC BLAT analysis tools, and identification of nearby genes and other genomic features, allowing high-throughput analysis of viral IS.[30] Detailed procedures and materials can be found in **Supplementary Materials and Methods**.

*Statistics.* To evaluate statistically significant differences between vector titers, transduction efficiencies, integration rates, or integration kinetics, data were analyzed using standard statistical tests calculated using GraphPad Prism 5 software (GraphPad Software, Inc, La Jolla, CA). Employed tests and statistical significance are detailed in figure legends.

### Supplementary material

**Figure S1.** Titers of pTrip PGK-Pac/CMV-Gfp vector stocks with different IN, assessed as measurements of p24/ul, vRNAc and TU/ul. Thirty stocks of vectors were analyzed to determine their content in viral capsid protein p24, in viral RNA copies (vRNAc) and transducing units (TU) per volume of supernatant.
**Figure S2.** Analysis of pTrip PGK-Pac/CMV-Gfp transduction efficiencies in HEK-293T cells with vectors bearing different integrases with respect to time and vector input.
**Figure S3.** Comparison of transduction efficiency and integration rate of different IDLV in HEK-293T cells after transduction with increasing input of vector pTrip PGK-Pac/CMV-Gfp.
**Figure S4.** Chromosomal distribution of integration sites.
**Table S1.** Overview of sequencing results.
**Table S2.** Nucleotides deletion in LTR.
**Table S3.** Common integration site (CIS) analysis.
**Table S4.** Top 10 CIS of individual vector data set.
**Material and Methods**

**Acknowledgments.** This work was founded by l'Institut National de la Santé et de la Recherche Médicale (INSERM) and the Higher Education Commission of Pakistan. We thank Nathalie Cartier and Yves Bigot for their advice in the course of this work and their critical comments on the manuscript. Authors declare no conflict of interest.

1. Naldini, L, Blömer, U, Gallay, P, Ory, D, Mulligan, R, Gage, FH *et al.* (1996). *In vivo* gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* **272**: 263–267.
2. Arhel, N (2010). Revisiting HIV-1 uncoating. *Retrovirology* **7**: 96.
3. Hare, S and Cherepanov, P (2009). The interaction between lentiviral integrase and LEDGF: structural and functional insights. *Viruses* **1**: 780–801.
4. Fields, BN, Knipe, DM, and Howley, PM (2007). *Fields Virology*. Wolters Kluwer Health/Lippincott Williams & Wilkins: Philadelphia, PA.
5. De Houwer, S, Demeulemeester, J, Thys, W, Taltynov, O, Zmajkovicova, K, Christ, F *et al.* (2012). Identification of residues in the C-terminal domain of HIV-1 integrase that mediate binding to the transportin-SR2 protein. *J Biol Chem* **287**: 34059–34068.
6. Larue, R, Gupta, K, Wuensch, C, Shkriabai, N, Kessl, JJ, Danhart, E *et al.* (2012). Interaction of the HIV-1 intasome with transportin 3 protein (TNPO3 or TRN-SR2). *J Biol Chem* **287**: 34044–34058.
7. Ocwieja, KE, Brady, TL, Ronen, K, Huegel, A, Roth, SL, Schaller, T *et al.* (2011). HIV integration targeting: a pathway involving Transportin-3 and the nuclear pore protein RanBP2. *PLoS Pathog* **7**: e1001313.
8. Meehan, AM, Saenz, DT, Morrison, JH, Garcia-Rivera, JA, Peretz, M, Llano, M *et al.* (2009). LEDGF/p75 proteins with alternative chromatin tethers are functional HIV-1 cofactors. *PLoS Pathog* **5**: e1000522.
9. Philippe, S, Sarkis, C, Barkats, M, Mammeri, H, Ladroue, C, Petit, C *et al.* (2006). Lentiviral vectors with a defective integrase allow efficient and sustained transgene expression *in vitro* and *in vivo*. *Proc Natl Acad Sci USA* **103**: 17684–17689.
10. Yáñez-Muñoz, RJ, Balaggan, KS, MacNeil, A, Howe, SJ, Schmidt, M, Smith, AJ *et al.* (2006). Effective gene therapy with nonintegrating lentiviral vectors. *Nat Med* **12**: 348–353.
11. Munir, S, Thierry, S, Subra, F, Deprez, E and Delelis, O (2013). Quantitative analysis of the time-course of viral DNA forms during the HIV-1 life cycle. *Retrovirology* **10**: 87.
12. Busschots, K, Voet, A, De Maeyer, M, Rain, JC, Emiliani, S, Benarous, R *et al.* (2007). Identification of the LEDGF/p75 binding site in HIV-1 integrase. *J Mol Biol* **365**: 1480–1492.
13. Petit, C, Schwartz, O and Mammano, F (2000). The karyophilic properties of human immunodeficiency virus type 1 integrase are not required for nuclear import of proviral DNA. *J Virol* **74**: 7119–7126.
14. Sarkis, C, Philippe, S, Mallet, J and Serguera, C (2008). Non-integrating lentiviral vectors. *Curr Gene Ther* **8**: 430–437.
15. Wanisch, K and Yáñez-Muñoz, RJ (2009). Integration-deficient lentiviral vectors: a slow coming of age. *Mol Ther* **17**: 1316–1332.
16. Engelman, A (1999). *In vivo* analysis of retroviral integrase structure and function. *Adv Virus Res* **52**: 411–426.
17. Bouyac-Bertoia, M, Dvorin, JD, Fouchier, RA, Jenkins, Y, Meyer, BE, Wu, LI *et al.* (2001). HIV-1 infection requires a functional integrase NLS. *Mol Cell* **7**: 1025–1035.
18. Rahman, S, Lu, R, Vandegraaff, N, Cherepanov, P and Engelman, A (2007). Structure-based mutagenesis of the integrase-LEDGF/p75 interface uncouples a strict correlation between *in vitro* protein binding and HIV-1 fitness. *Virology* **357**: 79–90.
19. Cribier, A, Ségéral, E, Delelis, O, Parissi, V, Simon, A, Ruff, M *et al.* (2011). Mutations affecting interaction of integrase with TNPO3 do not prevent HIV-1 cDNA nuclear import. *Retrovirology* **8**: 104.
20. Gallay, P, Hope, T, Chin, D and Trono, D (1997). HIV-1 infection of nondividing cells through the recognition of integrase by the importin/karyopherin pathway. *Proc Natl Acad Sci USA* **94**: 9825–9830.
21. Geraerts, M, Willems, S, Baekelandt, V, Debyser, Z and Gijsbers, R (2006). Comparison of lentiviral vector titration methods. *BMC Biotechnol* **6**: 34.
22. Nightingale, SJ, Hollis, RP, Pepper, KA, Petersen, D, Yu, XJ, Yang, C *et al.* (2006). Transient gene expression by nonintegrating lentiviral vectors. *Mol Ther* **13**: 1121–1132.
23. Vandekerckhove, L, Christ, F, Van Maele, B, De Rijck, J, Gijsbers, R, Van den Haute, C *et al.* (2006). Transient and stable knockdown of the integrase cofactor LEDGF/p75 reveals its role in the replication cycle of human immunodeficiency virus. *J Virol* **80**: 1886–1896.
24. Hare, S, Shun, MC, Gupta, SS, Valkov, E, Engelman, A and Cherepanov, P (2009). A novel co-crystal structure affords the design of gain-of-function lentiviral integrase mutants in the presence of modified PSIP1/LEDGF/p75. *PLoS Pathog* **5**: e1000259.

25. Kantor, B, Ma, H, Webster-Cyriaque, J, Monahan, PE and Kafri, T (2009). Epigenetic activation of unintegrated HIV-1 genomes by gut-associated short chain fatty acids and its implications for HIV infection. *Proc Natl Acad Sci USA* **106**: 18786–18791.
26. Charrier, S, Ferrand, M, Zerbato, M, Précigout, G, Viornery, A, Bucher-Laurent, S *et al.* (2011). Quantification of lentiviral vector copy numbers in individual hematopoietic colony-forming cells shows vector dose-dependent effects on the frequency and level of transduction. *Gene Ther* **18**: 479–487.
27. Butler, SL, Hansen, MS and Bushman, FD (2001). A quantitative assay for HIV DNA integration in vivo. *Nat Med* **7**: 631–634.
28. Cartier, N, Hacein-Bey-Abina, S, Bartholomae, CC, Veres, G, Schmidt, M, Kutschera, I *et al.* (2009). Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy. *Science* **326**: 818–823.
29. Paruzynski, A, Arens, A, Gabriel, R, Bartholomae, CC, Scholz, S, Wang, W *et al.* (2010). Genome-wide high-throughput integrome analyses by nrLAM-PCR and next-generation sequencing. *Nat Protoc* **5**: 1379–1395.
30. Arens, A, Appelt, JU, Bartholomae, CC, Gabriel, R, Paruzynski, A, Gustafson, D *et al.* (2012). Bioinformatic clonality analysis of next-generation sequencing-derived viral vector integration sites. *Hum Gene Ther Methods* **23**: 111–118.
31. Mátrai, J, Cantore, A, Bartholomae, CC, Annoni, A, Wang, W, Acosta-Sanchez, A *et al.* (2011). Hepatocyte-targeted expression by integrase-defective lentiviral vectors induces antigen-specific tolerance in mice with low genotoxic risk. *Hepatology* **53**: 1696–1707.
32. Schrijvers, R, Vets, S, De Rijck, J, Malani, N, Bushman, FD, Debyser, Z *et al.* (2012). HRP-2 determines HIV-1 integration site selection in LEDGF/p75 depleted cells. *Retrovirology* **9**: 84.
33. Shun, MC, Raghavendra, NK, Vandegraaff, N, Daigle, JE, Hughes, S, Kellam, P *et al.* (2007). LEDGF/p75 functions downstream from preintegration complex formation to effect gene-specific HIV-1 integration. *Genes Dev* **21**: 1767–1778.
34. Aiuti, A, Biasco, L, Scaramuzza, S, Ferrua, F, Cicalese, MP, Baricordi, C *et al.* (2013). Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome. *Science* **341**: 1233151.
35. Li, X, Koh, Y and Engelman, A (2012). Correlation of recombinant integrase activity and functional preintegration complex formation during acute infection by replication-defective integrase mutant human immunodeficiency virus. *J Virol* **86**: 3861–3879.
36. Emiliani, S, Mousnier, A, Busschots, K, Maroun, M, Van Maele, B, Tempé, D *et al.* (2005). Integrase mutants defective for interaction with LEDGF/p75 are impaired in chromosome tethering and HIV-1 replication. *J Biol Chem* **280**: 25517–25523.
37. Lu, R, Ghory, HZ and Engelman, A (2005). Genetic analyses of conserved residues in the carboxyl-terminal domain of human immunodeficiency virus type 1 integrase. *J Virol* **79**: 10356–10368.
38. Lu, R, Limón, A, Devroe, E, Silver, PA, Cherepanov, P and Engelman, A (2004). Class II integrase mutants with changes in putative nuclear localization signals are primarily blocked at a postnuclear entry step of human immunodeficiency virus type 1 replication. *J Virol* **78**: 12735–12746.
39. Lutzke, RA, Vink, C and Plasterk, RH (1994). Characterization of the minimal DNA-binding domain of the HIV integrase protein. *Nucleic Acids Res* **22**: 4125–4131.
40. Hearps, AC and Jans, DA (2006). HIV-1 integrase is capable of targeting DNA to the nucleus via an importin alpha/beta-dependent mechanism. *Biochem J* **398**: 475–484.
41. Brockman, MA, Chopera, DR, Olvera, A, Brumme, CJ, Sela, J, Markle, TJ *et al.* (2012). Uncommon pathways of immune escape attenuate HIV-1 integrase replication capacity. *J Virol* **86**: 6913–6923.
42. Zennou, V, Serguera, C, Sarkis, C, Colin, P, Perret, E, Mallet, J *et al.* (2001). The HIV-1 DNA flap stimulates HIV vector-mediated cell transduction in the brain. *Nat Biotechnol* **19**: 446–450.
43. Castaing, M, Guerci, A, Mallet, J, Czernichow, P, Ravassard, P and Scharfmann, R (2005). Efficient restricted gene expression in beta cells by lentivirus-mediated gene transfer into pancreatic stem/progenitor cells. *Diabetologia* **48**: 709–719.
44. Norrman, K, Fischer, Y, Bonnamy, B, Wolfhagen Sand, F, Ravassard, P and Semb, H (2010). Quantitative comparison of constitutive promoters in human ES cells. *PLoS One* **5**: e12413.
45. Brussel, A, Delelis, O and Sonigo, P (2005). Alu-LTR real-time nested PCR assay for quantifying integrated HIV-1 DNA. *Methods Mol Biol* **304**: 139–154.
46. Masuda, T, Planelles, V, Krogstad, P and Chen, IS (1995). Genetic analysis of human immunodeficiency virus type 1 integrase and the U3 att site: unusual phenotype of mutants in the zinc finger-like domain. *J Virol* **69**: 6687–6696.
47. Ao, Z, Fowke, KR, Cohen, EA and Yao, X (2005). Contribution of the C-terminal tri-lysine regions of human immunodeficiency virus type 1 integrase for efficient reverse transcription and viral DNA nuclear import. *Retrovirology* **2**: 62.
48. Leavitt, AD, Robles, G, Alesandro, N and Varmus, HE (1996). Human immunodeficiency virus type 1 integrase mutants retain *in vitro* integrase activity yet fail to integrate viral DNA efficiently during infection. *J Virol* **70**: 721–728.
49. Wiskerchen, M and Muesing, MA (1995). Human immunodeficiency virus type 1 integrase: effects of mutations on viral ability to integrate, direct viral gene expression from unintegrated viral DNA templates, and sustain viral propagation in primary cells. *J Virol* **69**: 376–386.



Supplementary Information accompanies this paper on the Molecular Therapy–Nucleic Acids website (http://www.nature.com/mtna)