Molecular Therapy
Nucleic Acids
Original Article

Engineering Next-Generation BET-Independent MLV Vectors for Safer Gene Therapy

Sara El Ashkar,[1,5] Dominique Van Looveren,[2,5] Franziska Schenk,[3] Lenard S. Vranckx,[1] Jonas Demeulemeester,[1,6] Jan De Rijck,[1] Zeger Debyser,[1] Ute Modlich,[3] and Rik Gijsbers[2,4]

[1]Laboratory for Molecular Virology and Drug Discovery, Department of Pharmaceutical and Pharmacological Sciences, 3000 Leuven, KU Leuven, Belgium; [2]Laboratory for Viral Vector Technology and Gene Therapy, Department of Pharmaceutical and Pharmacological Sciences, KU Leuven, 3000 Leuven, Belgium; [3]RG Gene Modification in Stem Cells, LOEWE Center for Cell and Gene Therapy Frankfurt, Paul-Ehrlich-Institute, 63225 Langen, Germany; [4]Leuven Viral Vector Core, KU Leuven, 3000 Leuven, Belgium

Retroviral vectors have shown their curative potential in clinical trials correcting monogenetic disorders. However, therapeutic benefits were compromised due to vector-induced dysregulation of cellular genes and leukemia development in a subset of patients. Bromodomain and extraterminal domain (BET) proteins act as cellular cofactors that tether the murine leukemia virus (MLV) pre-integration complex to host chromatin via interaction with the MLV integrase (IN) and thereby define the typical gammaretroviral integration distribution. We engineered next-generation BET-independent (Bin) MLV vectors to retarget their integration to regions where they are less likely to dysregulate nearby genes. We mutated MLV IN to uncouple BET protein interaction and fused it with chromatin-binding peptides. The addition of the CBX1 chromodomain to MLV IN_{W390A} efficiently targeted integration away from gene regulatory elements. The retargeted vector produced at high titers and efficiently transduced $CD34^+$ hematopoietic stem cells, while fewer colonies were detected in a serial colony-forming assay, a surrogate test for genotoxicity. Our findings underscore the potential of the engineered vectors to reduce the risk of insertional mutagenesis without compromising transduction efficiency. Ultimately, combined with other safety features in vector design, next-generation BinMLV vectors can improve the safety of gammaretroviral vectors for gene therapy.

INTRODUCTION

Stable integration of retroviral vectors encompassing a therapeutic transgene enables gene correction of severe blood and immune disorders. Over the past 25 years, murine leukemia virus (MLV)-based vectors have shown therapeutic benefit in gene therapy studies for primary immunodeficiencies (PIDs), such as X-linked severe combined immunodeficiency (SCID-X1), adenosine deaminase deficiency-severe combined immunodeficiency (ADA-SCID), and Wiskott-Aldrich syndrome (WAS).[1–4] MLV-based vectors were successfully used in the first clinical trials for ADA-SCID.[5–7] This led to the recent European approval of a retrovirus-based gene therapy product (Strimvelis; GSK GlaxoSmithKline Pharmaceuticals) to treat patients that lack a suitable human leukocyte antigen (HLA)-matched related stem cell donor.[8,9] However, in clinical trials for other PIDs, several patients developed leukemia or myelodysplasia, raising concerns about the safety of gene therapy.[10–14] These side effects have been directly attributed to the integration pattern and vector design. Insertional mutagenesis occurred as a consequence of vector integration preference in proximity of proto-oncogenes and activation by strong viral promoter and enhancer elements in the long terminal repeat (LTR) of retroviral vectors. To prevent insertional mutagenesis, self-inactivating (SIN) vectors with deleted enhancer sequences were designed. The lack of promoter/enhancer activity is compensated by weak heterologous promoters to drive transgene expression, such as the elongation factor 1 α short (EFS) and phosphoglycerate kinase (PGK) promoters.[15,16] Additionally, introduction of genetic insulator sequences has improved the safety of viral vectors.[17–22] The efficacy of these modified vectors was confirmed in pre-clinical studies and they are now in phase I/II clinical trials for several PIDs.[19,23–27]

A complementary approach to improve the safety of gene therapy is to alter the integration pattern, directing integration away from potentially unsafe regions. Gammaretroviral integration is not random, but rather is dictated by host cellular cofactors, such as the bromodomain and extraterminal domain (BET)-containing family of proteins (BRD2, BRD3, and BRD4) that serve as anchors on the host chromatin.[28,29] A motif in the unstructured C-terminal tail of MLV integrase (IN) interacts with the extraterminal (ET) domain of BRDs, where the latter tethers the retroviral pre-integration complex (PIC) to chromatin regions enriched in BET proteins and thereby defines the integration profile.[28–30] Deletion of the C-terminal domain (Δ23 amino acids [aa], $IN_{1–380}$) or a single substitution (IN_{W390A}) uncouples the BET interaction, resulting in

Received 18 November 2016; accepted 5 April 2017;
http://dx.doi.org/10.1016/j.omtn.2017.04.002.

[5]These authors contributed equally to this work.

[6]Present address: The Francis Crick Institute, NW1 1AT London, UK

Correspondence: Rik Gijsbers, Laboratory for Viral Vector Technology and Gene Therapy, Department of Pharmaceutical and Pharmacological Sciences, KU Leuven, Kapucijnenvoer 33, VCTB +5, 3000 Leuven, Belgium.
E-mail: rik.gijsbers@kuleuven.be

BET-independent (Bin) MLV vectors that transduce target cells at wild-type (WT) efficiency but with diminished integration in the vicinity of retroviral integration markers.[31]

Here, we developed next-generation BinMLV vectors with a potentially safer integration profile and lower propensity to activate nearby genes in an effort to alleviate the risk of insertional mutagenesis by interfering with the chromatin-tethering process. We linked chromatin binding peptide sequences to the C-terminal end of BinMLV IN and demonstrated that fusion of these peptides to BinMLV IN generates vectors that produce at high titers and transduce cells at wild-type efficiency. The addition of the chromodomain of CBX1 to MLV_{IN_W390A} efficiently retargeted integration away from gene regulatory elements. More importantly, the retargeted vector transduced human $CD34^+$ hematopoietic stem cells (HSCs) at wild-type efficiency, while genotoxicity assays revealed reduced transformation potential.

RESULTS

Efficient Transduction and Integration of Next-Generation BinMLV Vectors

To direct BinMLV integration away from potentially unsafe chromosomal regions, we tailored the chromatin-tethering process by fusing tethering peptides (between 16 and 61 aa long) to the C-terminal end of IN_{W390A} in the MLV packaging plasmid. We opted for peptides that bind histone markers that are widely spread across the chromatin (Figure 1A; Table 1). On one hand, we used peptides derived from cellular proteins that bind specific epigenetic histone modifications, such as the chromodomain of heterochromatin-binding protein 1β (CBX1, aa 20–73) and the chromodomain of Y-like protein (CDYL; aa 1–60),[32,33] giving rise to $IN_{W390A\text{-}CBX}$ and $IN_{W390A\text{-}CDYL}$, respectively. Alternatively, we fused virus-derived peptides, such as the tethering domain of the human papilloma virus (HPV8) E2 protein (aa 240–255)[34] and the N-terminal end of Kaposi sarcoma's latency associated nuclear antigen (LANA; aa 1–31), which bind to core histone 2A and 2B,[35] resulting in $IN_{W390A\text{-}E2}$ and $IN_{W390A\text{-}LANA}$, respectively.

The respective packaging plasmids were subsequently used to produce vesicular stomatitis virus glycoprotein G (VSV-G) pseudotyped MLV-based vectors encoding an LTR-driven EGFP reporter ($MLV_{IN\_W390A\text{-}CBX}$, $MLV_{IN\_W390A\text{-}CDYL}$, $MLV_{IN\_W390A\text{-}E2}$, and $MLV_{IN\_W390A\text{-}LANA}$; Figure 1A). In line with previous results,[31] transduction efficiencies in SupT1 cells were at similar levels for MLV_{IN_W390A} and MLV_{IN_WT} (Figure 1B). The addition of peptide sequences to the C-terminal end of MLV_{IN_W390A} resulted in BinMLV vectors that transduced as efficiently as MLV_{IN_WT} at different MOIs (Figure 1B) and resulted in comparable expression levels (measured as mean fluorescence intensities [MFIs] at day 3; Figure 1C). Transduction efficiencies and MFIs were corroborated at 10 days post-transduction, underscoring stable expression for the respective integrated vectors (Figures S1A and S1B). Similar data were obtained following transduction of HeLa cells (data not shown). Collectively, these results indicate that the addition of peptide sequences at the C-terminal end of MLV_{IN_W390A} does not impair vector integrity nor transduction efficiency compared to MLV_{IN_WT} and MLV_{IN_W390A}.

MLV_{IN_W390A}-Peptide Fusions Redistribute Bin Vector Integration

Next, we set out to investigate whether the respective IN chimeras redistribute MLV vector integration. Integration sites were amplified and sequenced, yielding a total of 43,676 unique sites and their computationally generated matched random control (MRC) sites. Initially, we evaluated integration frequencies relative to transcription start sites (TSSs), CpG islands (typically enriched in/near housekeeping gene promoters), and DNase hypersensitive sites (DHSs), both surrogate markers for open areas of active chromatin. In accordance with our previous work,[31] MLV_{IN_W390A} integration near TSSs and CpG islands decreased ~2-fold (Figure 2A; Table S1A). Whereas integration frequencies for $MLV_{IN\_W390A\text{-}CDYL}$ and $MLV_{IN\_W390A\text{-}E2}$ did not differ from MLV_{IN_W390A} near TSSs or CpG islands, the integration frequencies for both $MLV_{IN\_W390A\text{-}CBX}$ and $MLV_{IN\_W390A\text{-}LANA}$ were ~4-fold and ~2-fold lower when compared to MLV_{IN_WT} and MLV_{IN_W390A}, respectively (Figure 2A; Table S1A; $p < 0.001$, compared to MLV_{IN_W390A}). The detargeting effect of $MLV_{IN\_W390A\text{-}CBX}$ and $MLV_{IN\_W390A\text{-}LANA}$ was also evident near DHSs. 33% of MLV_{IN_W390A} integrations occurred in a 2-kb window around DHSs (~13% less than MLV_{IN_WT}), whereas only 25.66% and 27.84% of $MLV_{IN\_W390A\text{-}CBX}$ and $MLV_{IN\_W390A\text{-}LANA}$ integrations occurred in this window (Figure 2A; Table S1A). For comparison, integration datasets of prototype foamy viral vectors (FVs[36]) and HIV-derived lentiviral vectors (LVs[37]) were juxtaposed (Figure 2A). FVs are known to have a lower tendency to integrate near promoter regions compared to MLV vectors.[36] MLV_{IN_W390A} integration near TSSs was comparable to that of FVs (10.17% and 10.3%, respectively), whereas $MLV_{IN\_W390A\text{-}CBX}$ and $MLV_{IN\_W390A\text{-}LANA}$ integration occurred ~2-fold less near TSSs (Figure 2A; Table S1A) yet more frequently than LVs.[37] Similar results were obtained for larger window sizes (4-kb window, data not shown). Together, the data confirm that fusion of the CBX1 chromodomain and LANA peptide to MLV IN shifts vector integration away from the traditional markers associated with MLV integration.

In a next step, we analyzed integration preferences relative to a wider set of genomic features to evaluate overall vector integration (represented by heatmaps, Figure 2B). The tile color depicts the correlation for an integration dataset with the respective genomic feature (left) relative to the MRC, as quantified by the area under the receiver operating characteristic (ROC) curve. Asterisks indicate statistical significance of the integration site distributions of the respective vectors relative to that of MLV_{IN_W390A}. $MLV_{IN\_W390A\text{-}CDYL}$ and $MLV_{IN\_W390A\text{-}E2}$ integration profiles were similar to that of MLV_{IN_W390A} (Figure 2B), whereas $MLV_{IN\_W390A\text{-}CBX}$ and $MLV_{IN\_W390A\text{-}LANA}$ showed significant differences. The addition of CBX1 chromodomain to MLV $_{IN\_W390A}$ ($MLV_{IN\_W390A\text{-}CBX}$) significantly shifted integration for most of the genomic features ($p < 0.001$,

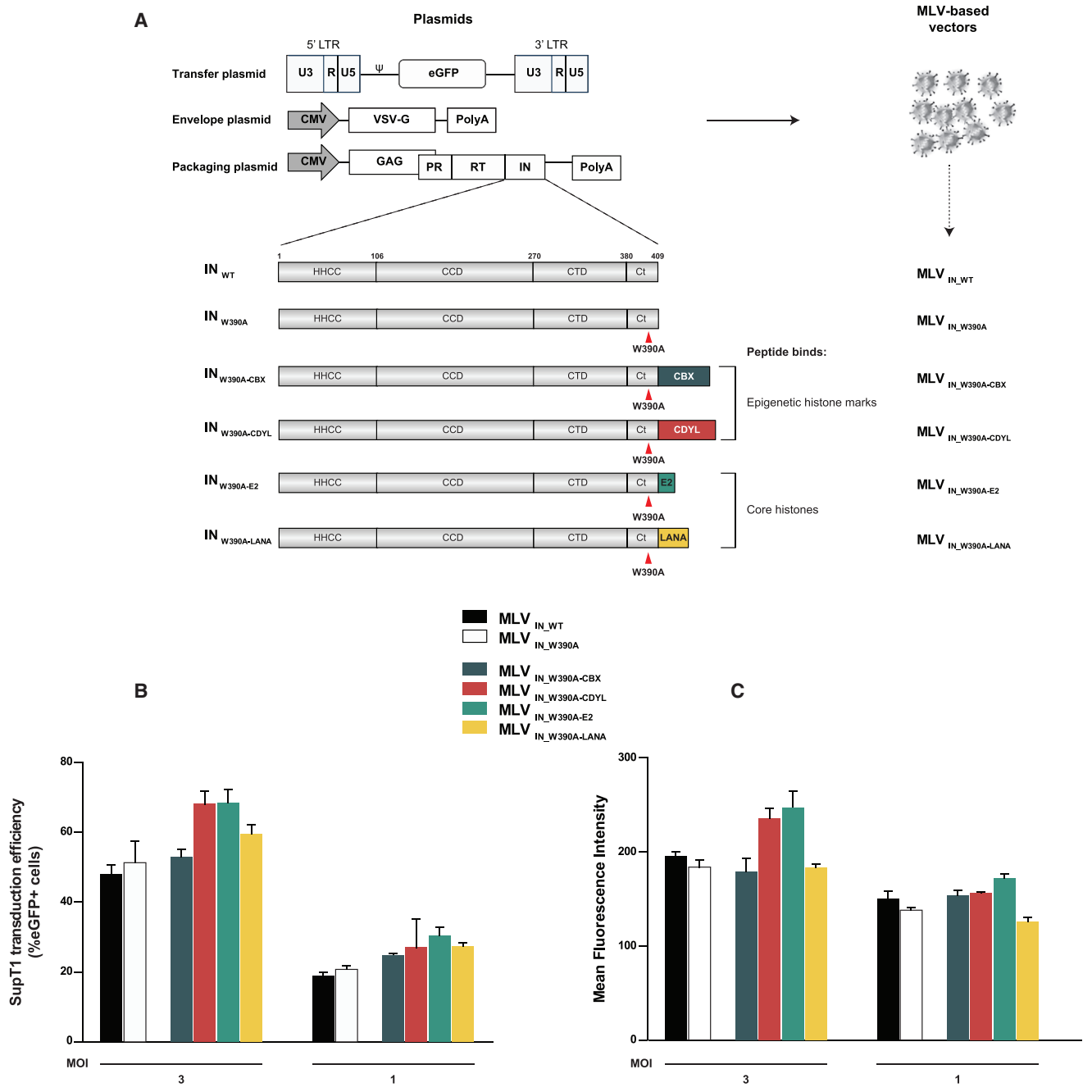


Figure 1. Transduction Efficiencies of Next-Generation BinMLV Vectors

(A) Schematic representation of MLV-based vector production. Structure of IN_{WT}, IN_{W390A}, and the peptides fused to IN_{W390A} are highlighted. The N-terminal HHCC zinc binding domain, the catalytic core domain (CCD), and the C-terminal domain (CTD) are indicated. Red arrowheads indicate the position of the W390A point mutation. The size of the fused peptides is proportionally represented. (B) FACS analysis of SupT1 cells transduced with equal RTUs of the indicated vectors at different MOIs. Three days post-transduction, the percentage of EGFP-positive cells was determined. Average values and standard deviations of triplicate measurements are shown. Data represent measurements from a representative experiment. (C) Mean fluorescence intensity of SupT1 cells transduced with the indicated next-generation BinMLV vectors 3 days post-transduction. Average values and standard deviations of triplicate measurements are shown. Data represent measurements from a representative experiment. ψ, packaging signal; CMV, cytomegalovirus promoter; GAG, group-specific antigen; IN, integrase; LTR, long terminal repeat; PolyA, polyadenylation signal; Pol, polymerase; PR, protease; RT, reverse transcriptase; VSV-G, vesicular stomatitis virus glycoprotein G.

Table 1. Overview of the Tethering Domains Fused to BinMLV Vectors

Protein	Accession Number	Selected Peptide	Binding Site	Sequence
CBX1	P83916	CD_1 CBX	H3K9me2	EYVVEKVLDRRVVKGKVEYLLKWKGFSDEDNT WEPEENLDCPDLIAEFLQSQKT
			H3K9me3	
CDYL	Q9Y232	CD_1 CDYL	H3K9me2	LMTFQASHRSAWGKSRKKNWQYEGPTQKLFL KRNNVSAPDGPSDPSISVSSEQSGAQQPPA
			H3K27me2	
			H3K27me3	
HPV8 E2	P06422	E2	H2A/H2B	QTETKGRRYGRRPSSR
KSHV LANA	E5LC01	LANA	H2A/H2B	MAPPGMRLRSGRSTGAPLTRGSCRKRNRSPE

Overview of the aa sequences and binding sites of CBX1 and CDYL chromodomains, HPV8 E2, and KSHV LANA used in this study. The aa position of each peptide in the respective protein is indicated in superscript. Protein sequences were downloaded from the UniProt database.CBX1, chromobox homolog 1; CD, chromodomain; CDYL; chromodomain protein Y-like; HPV, human papilloma virus; KSHV, Kaposi sarcoma human virus; LANA, latency-associated nuclear antigen.

compared to MLV_{IN_W390A}), resulting in redistributed integration with a more random pattern (Figure 2B, compare tile color of MLV_{IN_W390A} and $MLV_{IN\_W390A\_CBX}$; tile colors shift toward gray). On the other hand, $MLV_{IN\_W390A\text{-}LANA}$ only showed significant effects in smaller window sizes for the typical determinants of MLV vector integration, such as TSSs, CpG islands, and DHSs.

Since some of the peptides recognize specific chromatin marks, we also analyzed integration preferences near a collection of epigenetic features (Figure 2C). In line with previous data, uncoupling of BET interaction (MLV_{IN_W390A}) yields a more random integration pattern (compare colors between MLV_{IN_WT} and MLV_{IN_W390A}; tile colors shift toward black).[31] Nonetheless, MLV_{IN_W390A} integration still correlates with histone marks associated with open and transcriptionally active chromatin, such as H3 acetylation, H3K4 mono-, di-, and trimethylation, and H3K36 mono- and trimethylation albeit to a lesser extent (tile colors shift to darker blue; integration is enriched compared to MRC), while disfavoring transcriptionally silent regions or heterochromatin, such as di- and trimethylation of H3K9, H3K27, and H3K79[38] (tile colors shift to yellow; integration is depleted compared to MRC). Fusion of the CBX1 chromodomain and LANA peptide to IN_{W390A} shifted integration away from markers correlating with transcriptionally active open chromatin (blue tiles overall shift toward darker blue; $p < 0.001$, compared to MLV_{IN_W390A}), whereas the addition of the other peptides had no effect ($MLV_{IN\_W390A\text{-}CDYL}$ and $MLV_{IN\_W390A\text{-}E2}$, Figure 2C). CBX1 is known to bind H3K9me2/3 epigenetic marks via its chromodomain.[32] Interestingly, $MLV_{IN\_W390A\text{-}CBX}$ shifts integration more into transcriptionally silent heterochromatin regions, which is generally disfavored for integration, marked by di- and/or tri-methylation of H3K9 and H3K27 ($p < 0.001$, compared to MLV_{IN_W390A}; yellow tiles are darker). Together, these data show that the fusion of peptide tethers to the C-terminal end of MLV IN_{W390A} effectively retargets integration. As integration is shifted away from traditional MLV integration markers, known to associate with insertional mutagenesis in gene therapeutic trials,[39] a potentially safer integration site profile might be obtained.

Addition of Small Peptides to $MLV_{IN\_1\text{–}380}$ Rescues Its Transduction Defect

Apart from W390A substitution, deletion of the C-terminal tail of MLV IN (Δ23 aa, $IN_{1\text{–}380}$) similarly detargeted MLV integration, but with a transduction efficiency for $MLV_{IN\_1\text{–}380}$ that was ~3-fold lower than for MLV_{IN_WT}.[31] Therefore, we fused the respective peptides as an alternative C-terminal tail to $IN_{1\text{–}380}$ to assess whether transduction efficiency could be improved ($MLV_{IN\_1\text{–}380\text{-}CBX}$, $MLV_{IN\_1\text{–}380\text{-}CDYL}$, $MLV_{IN\_1\text{–}380\text{-}E2}$, and $MLV_{IN\_1\text{–}380\text{-}LANA}$; Figure 3A). Indeed, the addition of all peptides improved transduction efficiency to near wild-type levels at different MOIs (Figures 3B and S1C) and resulted in similar expression levels (Figures 3C and S1D), underscoring the importance of the C-terminal tail.

Next, we amplified 28,607 unique integration sites (Table S1B) to evaluate whether these IN chimeras also redistributed vector integration. In line with earlier data, $MLV_{IN\_1\text{–}380}$ integration was decreased near TSSs, CpG islands, and DHSs compared to MLV_{IN_WT}[31] (Figure S2A; Table S1B). Fusion of CDYL or E2 peptides to $MLV_{IN\_1\text{–}380}$ did not alter integration site distribution, whereas fusion of the CBX1 chromodomain and LANA peptide redistributed integration sites in a similar fashion as for the MLV_{IN_W390A} fusions (compare Figures 2A and S2A; compare Tables S1A and S1B). Additionally, $MLV_{IN\_1\text{–}380\text{-}CBX}$ integration redistributed similar to $MLV_{IN\_W390A\text{-}CBX}$ for a wide range of genomic features (Figure S2B) and histone modifications (Figure S2C), highlighting the specificity of the detargeting effects achieved by fusion of the CBX1 chromodomain or the LANA peptide to MLV_{IN_W390A} and $MLV_{IN\_1\text{–}380}$.

Addition of Peptide Tethers Does Not Alter the Local MLV Integration Site Sequence

Retroviral INs show weak but discernable target sequence preferences surrounding the site of integration. This local integration site sequence is mainly determined by IN contacts with the (nucleosomal) DNA template.[40,41] To assess whether the addition of alternative

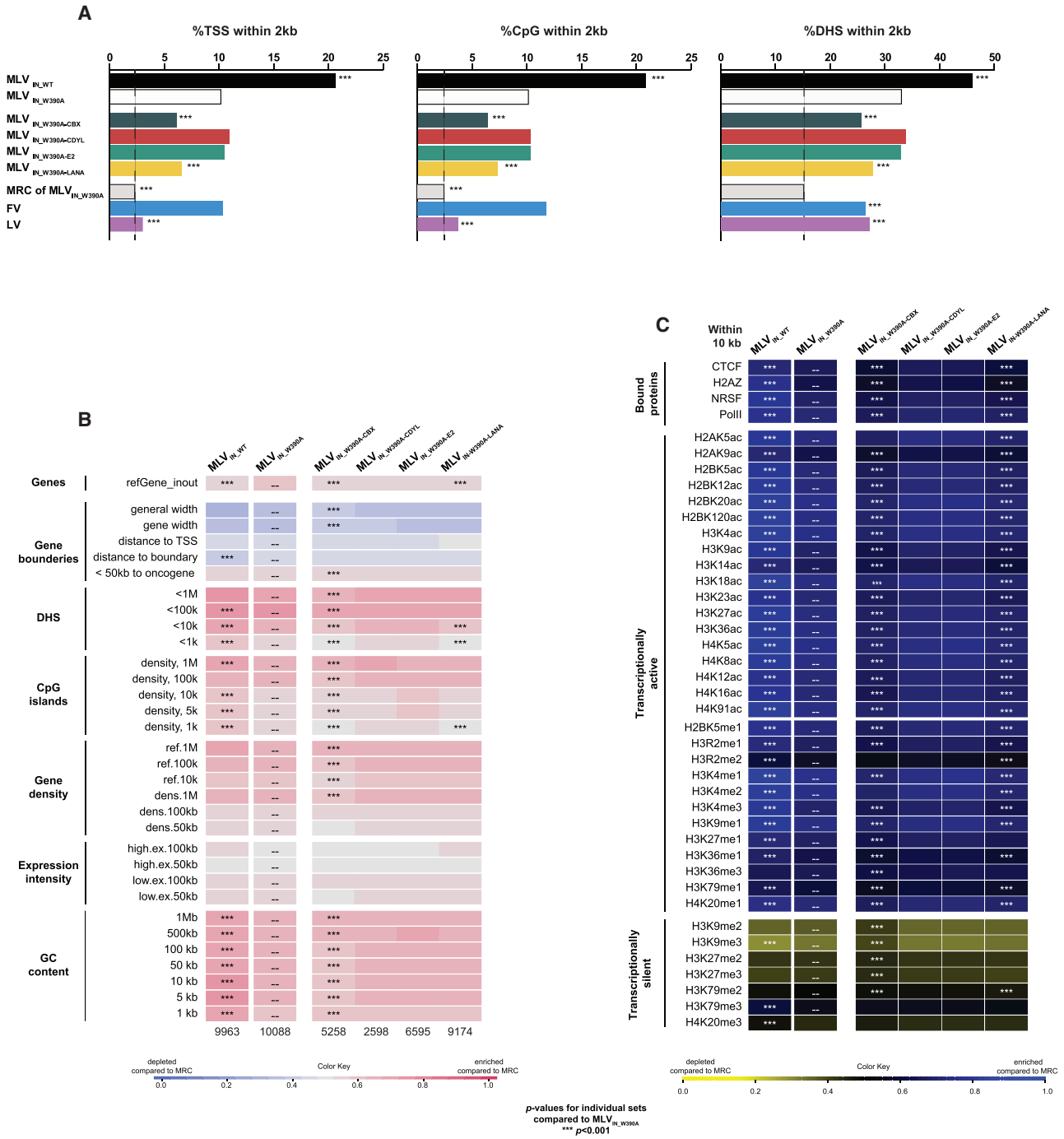


Figure 2. Integration Site Distribution of Next-Generation BinMLV Vectors

(A) Murine leukemia virus (MLV)-based vector integration sites obtained from SupT1 cells and their genomic distribution. Integration percentages in 2-kb windows around TSSs, CpG island midpoints, and DHSs are listed. For comparison, integration of computer-generated match random controls (MRCs) for MLV_{IN_W390A}, prototype foamy viral vectors (FVs), and HIV-derived lentiviral vectors (LVs) were included. *P* values show significant departures (***p < 0.001, pairwise Fisher test) from MLV_{IN_W390A}. (B and C) Heatmaps summarizing the relation between vector integration site frequency and different genomic (B) or epigenetic features within a10-kb interval (C) in SupT1 cells. Evaluated vectors are indicated above the columns. Features analyzed are shown to the left of the corresponding row of the heatmap. Tile colors indicate whether a particular feature is favored or disfavored for integration of the respective datasets relative to their MRCs, as detailed in the colored ROC area scale at the bottom of the panel. Number of integration sites is indicated below each column for the respective vector. *P* values show the significance of departures from MLV_{IN_W390A} integration sites in SupT1 cells (***p < 0.001, Wald statistics referred to χ^2 distribution). CpG, CpG-rich island; DHS, DNase I-hypersensitive site; TSS, transcription start site.

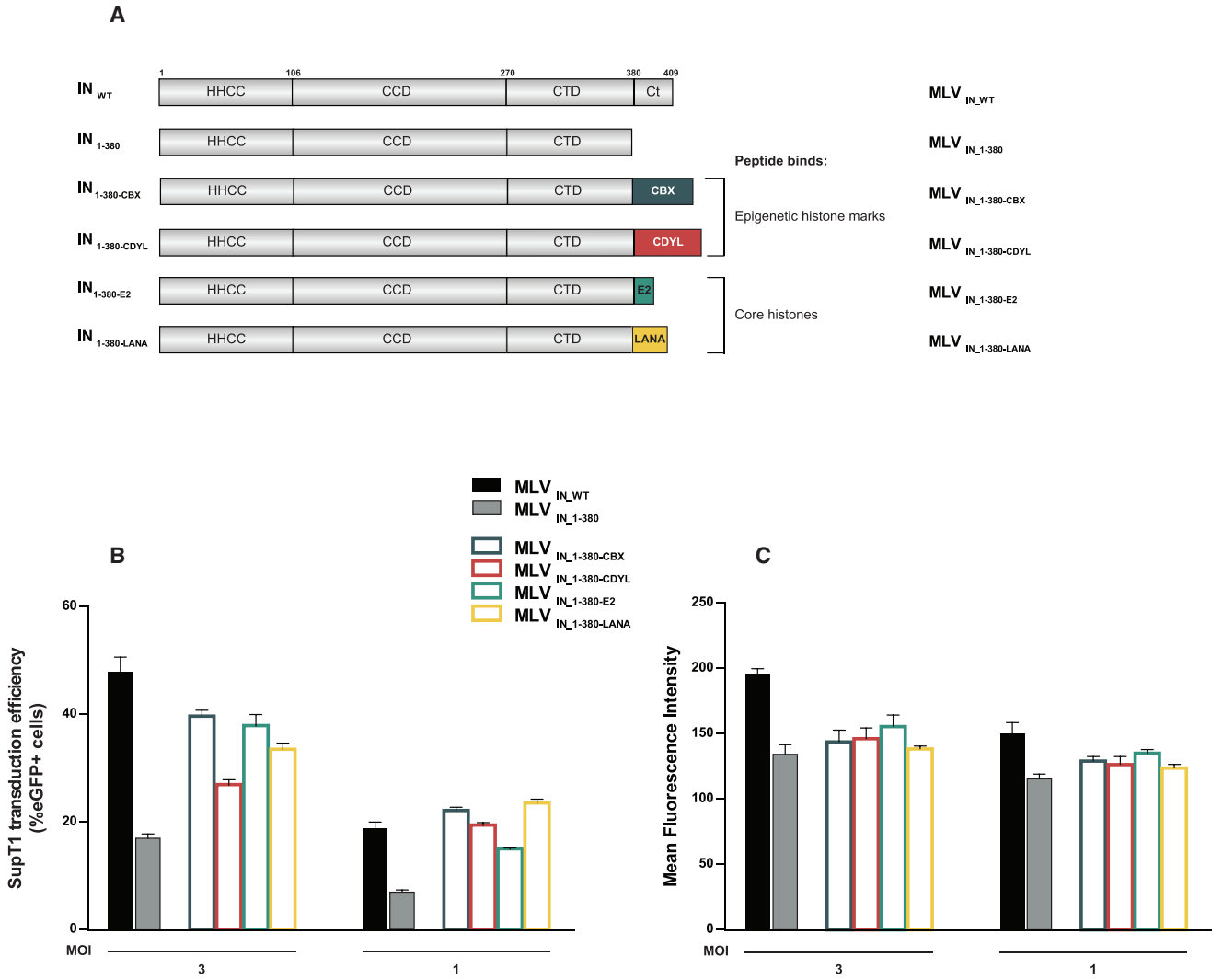


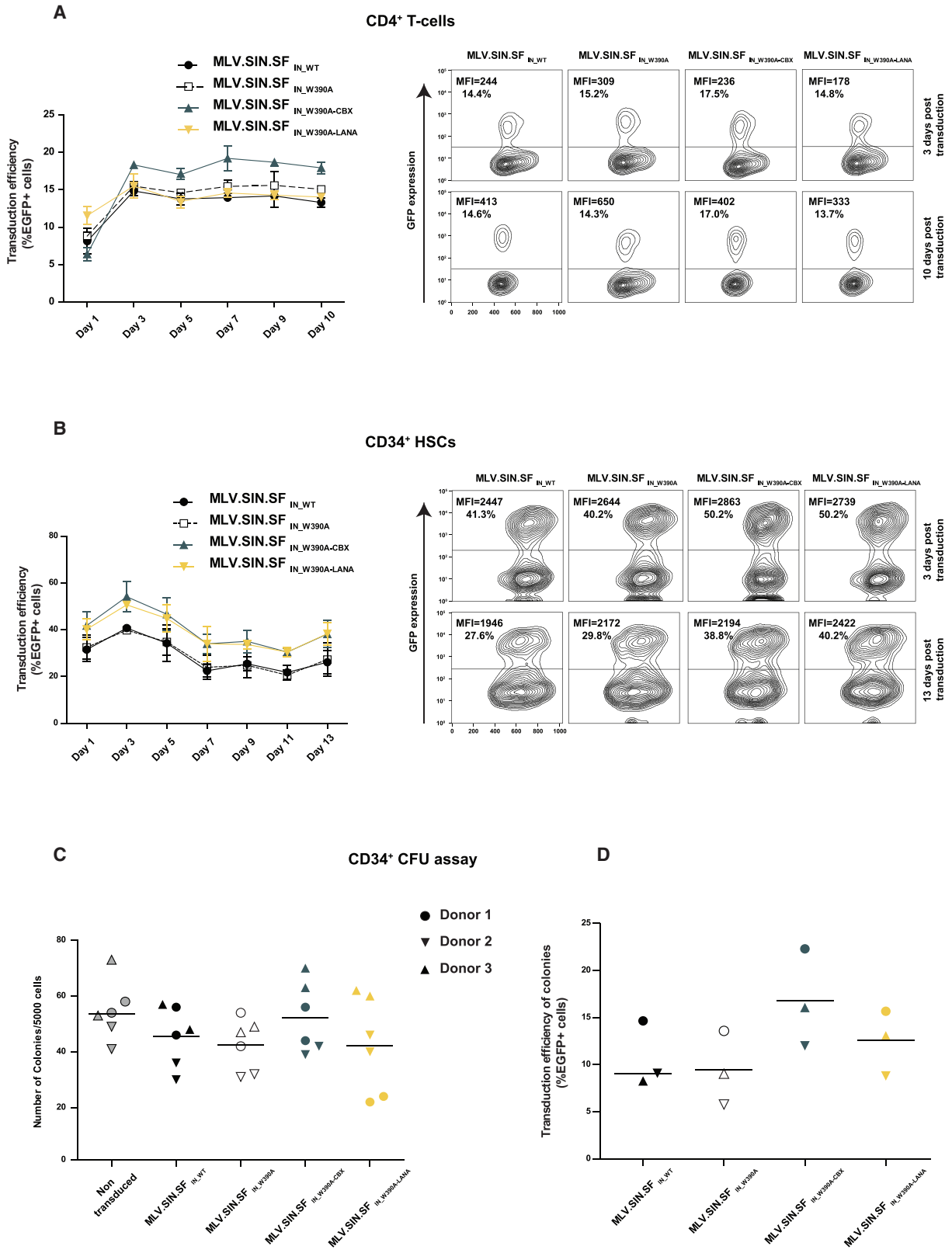
Figure 3. Effects of Fusing the Chromatin Binding Peptides in $MLV_{IN\_1–380}$ Backbone

(A) Schematic representation of IN_{WT}, $IN_{1–380}$, and peptides fused to $IN_{1–380}$. The N-terminal HHCC zinc binding domain, the catalytic core domain (CCD), and the C-terminal domain (CTD) are indicated. The size of the fused peptides is proportionally represented. (B and C) TE (B) and mean fluorescence intensity (C) in SupT1 cells 3 days post-transduction at different MOIs. Average values and standard deviations of triplicate measurements are shown. Data represent measurements from a representative experiment.

peptide tethers to MLV_{IN_W390A} influenced the local integration site sequence, we constructed sequence logos (Figure S3). Results indicate that the local integration site sequence preferences remained unaffected.[42] Similar results were obtained for $MLV_{IN\_1–380}$ peptide fusions (data not shown).

Evidence for Safer Integration of Next-Generation $BinMLV_{IN\_W390A-CBX}$ Vector

Comparative integrome analysis identified $MLV_{IN\_W390A-CBX}$ and $MLV_{IN\_W390A-LANA}$ as next-generation BinMLV vectors with substantial detargeting from traditional MLV markers without compromising transduction efficiency. Initially, we sought to assess whether the altered integration profiles of next-generation BinMLV vectors are potentially safer. We determined integration frequencies near a set of previously defined genomic regions[43,44] that should be avoided to prevent cellular transformation, such as regions proximal to TSSs (<50 kb TSS), cancer-related oncogenes (<300 kb AllOnco), and microRNA (miRNA) coding regions (<300 kb miRNA), or within transcription units (TUs) and ultraconserved regions (UCRs). For each integrome dataset, the percentage of integrations near these features was determined (Table S2). When assessing the individual safe harbor criteria, $MLV_{IN\_W390A-CBX}$ integrated less frequently near each of these features when compared to MLV_{IN_WT}, MLV_{IN_W390A}, and LV integration. Taken together, this computational analysis suggests an overall safer integration profile that is less likely to disturb nearby genes.



(legend on next page)

Therapeutic Potential of Next-Generation BinMLV Vectors

An essential factor in engineering new viral vectors includes evaluating their therapeutic potential to transduce clinically relevant cells. Therefore, we challenged primary human $CD4^+$ T cells and $CD34^+$ HSCs with MLV_{IN_WT}, MLV_{IN_W390A}, $MLV_{IN\_W390A\text{-}CBX}$, and $MLV_{IN\_W390A\text{-}LANA}$ vectors carrying a more relevant SIN gammaretroviral vector genome (MLV.SIN) driving reporter gene expression from a spleen focus-forming virus (SF) enhancer/promoter,[21,45] referred to as MLV.SIN.SF_{IN_WT}, MLV.SIN.SF_{IN_W390A}, MLV.SIN.$SF_{IN\_W390A\text{-}CBX}$, and MLV.SIN.$SF_{IN\_W390A\text{-}LANA}$, respectively (Figure S4A) and monitored transduction efficiency (percentage of $eGFP^+$ cells) over time (MOI of 4.5, Figures 4A and 4B, respectively). MLV.SIN.$SF_{IN\_W390A\text{-}CBX}$ and MLV.SIN.$SF_{IN\_W390A\text{-}LANA}$ transduced $CD4^+$ T cells and $CD34^+$ HSCs to the same extent as MLV.SIN.SF_{IN_WT} and MLV.SIN.SF_{IN_W390A} and this was sustained over time (fluorescence-activated cell sorting [FACS] plots, Figures 4A and 4B, respectively; compare transduction efficiency (TE) and MFI at day 3 and day 10), indicating the absence of silencing effects due to the retargeted integration preference of next-generation BinMLV vectors. Additionally, we performed colony-forming unit (CFU) assays for $CD34^+$ HSCs harvested from three different donors to determine whether the retargeted next-generation BinMLV vectors efficiently transduced hematopoietic progenitor cells (Figure 4C). After 14 days in culture, the number of colonies derived from $CD34^+$ HSCs transduced with MLV.SIN.SF_{IN_W390A}, MLV.SIN.$SF_{IN\_W390A\text{-}CBX}$, and MLV.SIN.$SF_{IN\_W390A\text{-}LANA}$ was in line with MLV.SIN.SF_{IN_WT}, indicating that BinMLV vectors do not affect the functionality of $CD34^+$ HSCs (Figure 4C). FACS analysis of the respective pooled transduced CFUs revealed comparable percentages of EGFP-positive cells (Figure 4D) and no difference in MFIs could be observed (Table S3) for the different BinMLV vectors, confirming the lack of increased transgene silencing.

Reduced Transformation Potential of Next-Generation $BinMLV_{IN\_W390A\text{-}CBX}$ Vector

The intrinsic integration preference of MLV-based vectors has been shown previously to be one of the driving factors of vector-mediated genotoxicity that occurred when the integrated vector dysregulated host genes, leading to oncogenic transformation. Therefore, we evaluated the genotoxic potential of next-generation BinMLV vectors carrying a MLV.SIN-vector genome with an internal SF enhancer/promoter, which is known to trigger insertional transformation events[21,45] (Figure S4A) in the in vitro immortalization assay (IVIM)[46] and the transformational incidence in a murine CFU assay (Figure S4B).[47,48] To ensure the chance of immortalization, at least 55% of cells were transduced, corresponding to a mean vector copy number (VCN) of > 2.[49] Murine hematopoietic lineage-depleted (lin^-) bone marrow cells were transduced with the indicated vectors at different MOIs in six independent transductions in three IVIM assays (Figure S4C). Transduction efficiencies and integrated VCNs were comparable for the different vectors (Figures 5A and 5B). The IVIM assay revealed immortalization for all MLV vectors carrying the MLV.SIN.SF vector architecture (Figures 5C and S4C). The replating frequency is a measure for the fitness of clones, while the number of positive assays reflects the incidence of immortalization events. MLV.SIN.SF_{IN_W390A} and MLV.SIN.$SF_{IN\_W390A\text{-}LANA}$ resulted in replating clone numbers in line with MLV.SIN.SF_{IN_WT}, whereas a reduction in clone numbers was observed for MLV.SIN.$SF_{IN\_W390A\text{-}CBX}$ (mean number of positive wells; mean replating frequency of 2.7×10^{-3}, Figure S4C). MLV.SIN.$SF_{IN\_W390A\text{-}CBX}$ also displayed a reduced replating frequency/copy number (1.4-fold or ~30% lower) compared to MLV.SIN.SF_{IN_WT}, although it was not statistically significant ($p > 0.05$, compared to MLV_{IN_WT}, Mann-Whitney U test, Figures 5C and S4C).

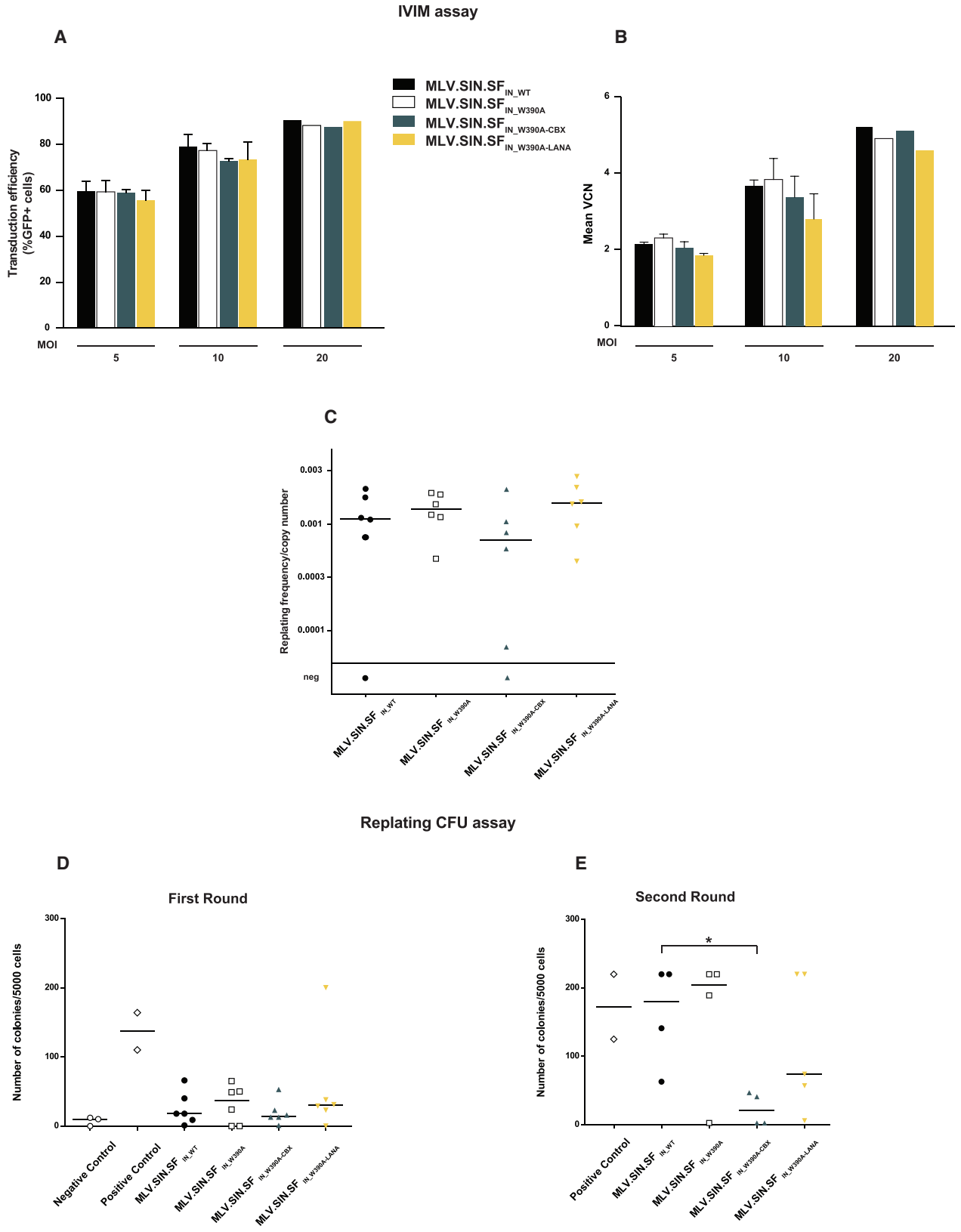
The serial replating CFU assay is a method used to confirm cellular anchorage-independent growth in vitro. The assay provides a stringent method for the detection of the tumorigenic potential of transformed murine HSCs in semi-solid medium.[47,48] In a parallel approach, we sought to assess the serial colony-forming capacity of the transduced lineage marker-negative cells 2 weeks after expansion in the IVIM assay. We included cells transduced with an LTR-driven MLV-based vector as an additional positive control.[46] Cells were plated at a density of 5,000 cells/well in semi-solid medium (Figure S4B). After 10 days in culture (first round, Figures 5D and S4B), the number of colonies obtained for all MLV.SIN.SF vectors was not significantly different from that of non-transduced cells ($p > 0.05$, compared to negative control, Mann-Whitney U test, Figure 5D), whereas the number of colonies for the positive control was significantly higher ($p < 0.05$, compared to negative control, Mann-Whitney U test, Figure 5D). After isolating the cells from the colony assays, 5,000 cells were replated (second round, Figures 5E and S4B). MLV.SIN.SF_{IN_WT}- and MLV.SIN.SF_{IN_W390A}-transduced cells formed numbers of colonies comparable to the positive control ($p > 0.05$, compared to positive control, Mann-Whitney U test, Figure 5E), whereas the number of colonies from MLV.SIN.$SF_{IN\_W390A\text{-}CBX}$ transduced cells was significantly reduced ($p < 0.05$ compared to MLV.SIN.SF_{IN_WT}, Mann-Whitney U test, Figure 5E). Taken together, these data indicate that engineering the MLV vector configuration by modifying the IN protein ($IN_{W390A\text{-}CBX}$) reduced the outgrowth of replating clones, indicating a potentially safer profile.

DISCUSSION

Retroviral vector technology offers great potential to treat genetic disorders and is a powerful tool for long-term correction of genetic defects in a variety of severe hematological disorders.[1,50,51] Despite the

Figure 4. Translational Potential of Next-Generation BinMLV Vectors

(A and B) TE of next-generation BinMLV vectors in primary $CD4^+$ T cells (A) and $CD34^+$ HSCs (B) at different time points. Percentage of EGFP-positive cells and mean fluorescence intensities (MFIs) are indicated 3 days and 10 or 13 days post-transduction for $CD4^+$ T cells and $CD34^+$ cells, respectively. Data represent measurements from a representative experiment. (C) Colony-forming unit (CFU) assay of human $CD34^+$ HSCs harvested from three donors and transduced with the indicated vectors. The number of colonies was scored after 14 days. (D) TE (percentage of EGFP-positive cells) of CFU colonies (described in C) at 17 days post-transduction from three donors.



(legend on next page)

initial success, a subset of patients developed serious adverse events, such as leukemia or myelodysplasia, that could be directly related to the design and characteristics of the viral vector used.[10–14] Understanding the mechanisms of retroviral vector genotoxicity is, therefore, essential to engineer improved viral vectors with reduced genotoxic potential.

Vector-mediated genotoxicity is defined by (1) the specific integration profile and (2) the design of the integrating proviral genome. Each retroviral family displays a specific integration profile. MLV integration is significantly enriched near TSSs and active enhancer regions[38] and thus potentially causes insertional mutagenesis. Significant efforts have been made to develop next generations of retroviral vectors with reduced genotoxic potential, such as SIN vectors and weaker internal promoters.[19,21,52–58] Nonetheless, their integration is still targeted to gene regulatory regions,[38,59] where they have the potential to disrupt or dysregulate the transcription of nearby genes by other mechanisms.[60] Thus, the development of viral vectors that integrate away from genes may be safer for clinical applications. One approach is the development of other viral vector platforms with more favorable genomic distributions, like lentiviral, foamy, and alpharetroviral vectors.[25,60–62] Alternatively, retroviral vectors can be re-engineered to obtain a more desirable integration pattern that is detargeted from its traditional chromosomal locations and ultimately only occurs at the preferred sites of the host-cell chromosome.

Retroviral integration site selection is dictated by the interaction between the viral IN as part of the PIC and cellular cofactors. Previously, we re-engineered the MLV- and HIV-cellular tethering cofactors (BET and LEDGF/p75, respectively) and demonstrated efficient redistribution of retroviral integration without compromising transgene expression.[29,63–65] However, this approach requires the introduction (at least transient) of artificial anchors in target cells prior to application of the therapeutic vectors,[63,64] which is not always desirable in a clinical setting. A more straightforward strategy is to directly engineer vector particles to contain proteins with adapted or unique binding domains to direct integration. Other groups attempted to redistribute MLV integration through modifications of the MLV Gag protein by fusion of tethering peptides to rescue a mutated MLV p12 protein.[66,67] However, integration site distribution of the engineered MLV p12 chimeras was not altered,[66,67] suggesting that the primary role of MLV p12 is tethering the virus/vector to the condensed host-cell chromatin rather than targeting the genomic integration (reviewed in Rein[68]). Analysis of MLV IN-BET interaction[29] enabled us to generate Bin MLV vectors.[31] Interaction with BET proteins was uncoupled either by deletion of the C-terminal tail of MLV IN ($MLV_{IN\_1–380}$) or by a single point mutation (MLV_{IN_W390A}).

Here, we successfully modified the MLV integration profile by fusing alternative chromatin binding peptides to BinMLV IN, resulting in a significantly different integration site pattern (Figure 2). The addition of the CBX1 chromodomain or LANA peptide redistributed MLV integration to the same extent relative to the typical MLV markers (TSSs, DHSs, and CpG islands), with a more than 3-fold reduction in integration frequency near TSSs and CpG islands compared to wild-type MLV (Figures 2A and S2A; Table S1). Comparative integrome analysis near a range of genomic and epigenetic features corroborated that IN_{LANA} vectors ($MLV_{IN\_W390A\text{-}LANA}$ and $MLV_{IN\_1–380\text{-}LANA}$) and IN_{CBX} vectors ($MLV_{IN\_W390A\text{-}CBX}$ and $MLV_{IN\_1–380\text{-}CBX}$) redistribute along the chromatin, although the effect was much more pronounced for IN_{CBX} vectors (Figures 2B and 2C and S2B and S2C). A similar strategy was employed by Hocum et al.,[69] where the full CBX1 protein was fused to the IN of foamy virus and additional modifications in the FV Gag protein were required to achieve significant effects on retargeting FV integration. The newly generated foamy retroviral vectors (FV) integrated ~2-fold less frequent near genes and proto-oncogenes.[69] In the case of MLV, fusion of a single CBX1 chromodomain to the C-terminal end of MLV IN_{W390A} was sufficient to shift integration toward epigenetic markers for transcriptionally silent regions (H3K9me2/3 and H3K27me3), known to be bound by CBX1[32] (Figure 2C; Table S1).

Next to the retargeting effect, next-generation BinMLV vectors can be produced at high titers and efficiently transduce clinically relevant cells such as primary $CD4^+$ T cells and $CD34^+$ HSCs without any apparent transgene silencing (Figure 4), highlighting the translational potential of these vectors. Finally, we assessed the newly engineered Bin MLV vectors ($MLV_{IN\_W390A\text{-}CBX}$ and $MLV_{IN\_W390A\text{-}LANA}$) in the IVIM assay as well as in a murine serial replating CFU assay to predict the genotoxic profile (Figure S4B).[46] BinMLV and wild-type MLV-based vectors were produced carrying the same SIN gammaretroviral vector genome (MLV.SIN.SF) with an internal SF enhancer/promoter to drive EGFP expression (Figure S4A), a design known to trigger insertional transformation.[21,45] Whereas $MLV.SIN.SF_{IN\_WT}$, $MLV.SIN.SF_{IN\_W390A}$, and $MLV.SIN.SF_{IN\_W390A\text{-}LANA}$ displayed a comparable frequency of IVIM clones, immortalization events were slightly reduced for $MLV.SIN.SF_{IN\_W390A\text{-}CBX}$ (Figures 5C and S4C) but were not significantly different. The results of the IVIM assay in our experiments confirmed that the strong spleen focus-forming

Figure 5. Assessment of Genotoxicity Profile of Next-Generation BinMLV Vectors

(A and B) TE at day 4 (A) and mean vector copy number at day 5 post-transduction (B) of next-generation BinMLV vectors in murine hematopoietic lineage marker-negative (lin^-) bone marrow cells at increasing MOIs. MOI 20 was performed only once. (C) Replating frequencies corrected for the mean vector copy number, measured in DNA of mass cultures at day 5; horizontal lines indicate the median for a given vector design. (D and E) Serial CFUs assay. The number of colonies per 5,000 cells plated in methylcellulose is shown. Each dot represents the number of colonies formed for an independently transduced culture. A negative control represents non-transduced lin^- cells, while a positive control represents lin^- cells transduced with an LTR-SFFV-driven MLV-based vector. After the first round (D), the number of colonies was scored and cells from positive assays were harvested and re-seeded at 5,000 cells for second round colony formation (E). Colony counts > 200 colonies are shown as 220. The experimental setup is shown in Figure S4B. VCN, vector copy number.

virus (SFFV) enhancer/promoter present in all vector genomes produced replating clones with similar efficiency in all cultures independent of the IN, while the low replating in two assays of $MLV_{IN\_W390A\text{-}CBX}$ may be attributed to the detargeted integration pattern. Hence, at least for SFFV, promoter strength has a greater impact on the generation of insertional mutants than integration site preference in the IVIM assay.

To further address whether detargeted BinMLV vectors may be less genotoxic, we employed serial replating CFU assays of expanded murine HSCs[47,48] to assess vector integration-related genotoxicity. The murine CFU assays demonstrated that MLV.SIN.SF_{IN_WT} and MLV.SIN.SF_{IN_W390A} potently transformed murine HSCs, whereas fusion of the CBX1 chromodomain to MLV IN_{W390A} (MLV.SIN.$SF_{IN\_W390A\text{-}CBX}$) resulted in a significantly reduced transformational incidence ($p < 0.05$, Figure 5E).

In conclusion, we demonstrate the potential to engineer MLV-based vectors that detarget from unsafe regions by fusing peptide fragments to the C-terminal end of MLV IN. The mere addition of the CBX1 chromodomain (and LANA peptide, to a lesser extent) was sufficient to detarget integration preference away from the traditional markers of MLV integration. In order to translate these findings into suitable vectors for the (pre)clinical field, the performance of the new Bin MLV should be evaluated in more relevant pre-clinical safety assays and bone marrow/HSC transplantation assays to validate their improved safety profile. Together, our findings will help achieve better control of MLV-based vector integration preferences. Combining next-generation BinMLV packaging constructs with next-generation SIN gammaretroviral vector architectures that incorporate a weaker physiological promoter less likely to dysregulate nearby genes[19,70] will lead to an additional reduction in genotoxicity.

MATERIALS AND METHODS

Plasmids

BinMLV IN was cloned as previously described.[31] Chromodomians of CBX and CDYL fusions were cloned with gBlocks (IDT) in PacI- and NotI-digested pcDNA3.MLV.gp packaging plasmid, a kind gift from Prof. Axel Schambach.[71] HPV8 E2 and LANA peptide fusions were introduced into the indicated vectors by an oligonucleotide annealing strategy using the same restriction sites. Oligonucleotide sequences are listed in Table S4. All enzymes were purchased from Thermo Fisher Scientific. The integrity of all plasmids was verified by DNA sequencing.

Cell Culture

SupT1 cells were cultured in RPMI-1640 medium (Gibco-BRL/Life Technologies) supplemented with 10% heat-inactivated fetal bovine serum (Gibco-BRL) and gentamicin (50 μg/mL; Gibco-BRL). HeLa cells were cultured in DMEM (Gibco-BRL) supplemented with 8% heat-inactivated fetal bovine serum and gentamicin. All cells were grown in a humidified atmosphere with 5% CO_2 at 37°C.

Primary Cell Purification

Human peripheral blood mononuclear cells (PBMCs) were purified from buffy coats of three different donors, obtained from the Red Cross blood transfusion center, using density-gradient centrifugation (Lymphoprep; Axis-Shield). Primary $CD4^+$ T cells were selectively enriched using bi-specific monoclonal antibody (mAb) CD3.8 (0.5 μg/mL, NIH AIDS Reagents Program; https://www.aidsreagent.org) for 5 days. $CD4^+$ T cells were cultured in RPMI medium supplemented with 15% FBS, gentamycin, interleukin (IL)-2 (100 U/mL; Peprotech), and MEM Non-Essential Amino Acids Solution (MEM NEAA) (50 μg/mL; Gibco-BRL), referred to as T-cell medium (TCM). $CD34^+$ HSCs were positively selected with anti-CD34-conjugated microbeads according to the manufacturer's instructions (MACS; Miltenyi Biotec) and stimulated for 48 hr in StemSpan SFEMII medium containing CC100 Cytokine Cocktail (STEMCELL Technologies).

Retroviral Vector Production and Transduction

Viral vectors were produced as previously described.[72] Briefly, MLV-based vectors were produced by a triple polyethylenimine (PEI)-based or Ca-phosphate transfection of 293T cells with a pVSV-G envelope, pcDNA3.MLV.gp packaging plasmids or their derived fusions (see above), and p450-GFP transfer plasmid (kindly provided by F.D. Bushman) encoding an LTR-driven EGFP reporter. For the IVIM assay, the transfer plasmid pSRS11.SF.GFP.pre (referred to here as MLV.SIN.SF.EGFP.pre) was used, which was kindly provided by Axel Schambach.[71]

Produced vectors were concentrated by tangential flow filtration and normalized based on RT units (RTUs; non-functional titration) by the SYBR Green I product-enhanced reverse transcriptase assay (SG-PERT).[73] Subsequently, functional transducing titers were determined in SupT1 cells reaching titers > 10^7 TU/mL. For transduction of laboratory cell lines, SupT1 cells (12×10^4/well) and HeLa cells (2×10^4/well) were seeded in 96-well plates and transduced with a MOI of 1 and 3 of the respective vectors.

Seventy-two hours post-transduction, 50% of the cells were harvested for FACS analysis, while the remaining 50% were cultured for 10 days post-transduction for a second FACS analysis and to perform integration site sequencing. Prior to primary cell transduction, $CD4^+$ T cells (25.104/well) and $CD34^+$ HSCs (10.104/well) were pre-stimulated for 5 days in TCM and 2 days in StemSpan medium enriched with CC100 Cytokine Cocktail, respectively. An MOI of 4.5 of the different vectors was applied by spinoculation (2 hr, 1,200 g). Cells were analyzed for EGFP expression by flow cytometry at the indicated time points.

Genomic DNA Isolation and qPCR

Genomic DNA (gDNA) isolation and qPCR were performed as previously described.[31] Briefly, 2 million cells were pelleted and genomic DNA was extracted using a Mammalian Genomic DNA Miniprep Kit (Sigma-Aldrich). Samples corresponding to 700 ng genomic DNA were used for analysis. Each reaction contained 12.5 μL iQ Supermix (Bio-Rad), 40 nM forward and reverse EGFP primer, and 40 nM

EGFP probe in a final volume of 25 μL. RNaseP or β-actin was quantified as the endogenous control (TaqMan RNaseP control reagent; Applied Biosystems). Samples were run in triplicate for 3 min at 95°C followed by 50 cycles of 10 s at 95°C and 30 s at 55°C in a LightCycler 480 (Roche Applied Science). Analysis was performed using the LightCycler 480 software supplied by the manufacturer.

Recovery of Integration Sites and Analysis of Integration Site Distributions

Recovery of integration sites was performed as previously described.[29] Briefly, linkers were ligated to restriction enzyme-digested (MseI) genomic DNA isolated from transduced cells and virus-host DNA junctions were amplified by nested PCR. Samples were individually barcoded with the second pair of PCR primers to generate 454 libraries. PCR products were purified and sequenced using 454/Roche pyrosequencing (titanium technology). Reads were quality-filtered by requiring perfect matches to the LTR linker, barcode, and flanking LTR and were subsequently mapped to the human/mouse genome. All sites were required to align to the reference genome within 3 bp of the LTR edge. To control for possible biases in the datasets due to the choice of the MseI restriction endonuclease in cloning integration sites, random control sites were generated computationally and matched to experimental sites with respect to the distance to the nearest MseI cleavage site (MRC).[74,75] To do so, each experimental integration site was paired with three random control sites in the genome with respect to the distance to the nearest MseI cleavage site in the genome. A more detailed explanation can be found in the supplemental guidelines included in Ocwieja et al.[76]

Analyses were carried out as described in Marshall et al.[77] A detailed account of the statistical methods used and the methods for forming and analyzing heatmaps using ROC curves can be found in Brady et al.[78] Consensus sequence analysis at the point of integration was performed using WebLogo3 (http://140.114.98.75/weblogo/). For association with specific genomic features, the distance of each integration site (in kilobases) to the respective genomic feature was calculated (midpoint of the CpG island or DHS, and the X5-end of genes as a measure for the TSS). Integration sites upstream of the genomic feature were given negative kilobase values, while downstream integration sites were calculated as positive. For heatmaps, comparisons were carried out over three different interval sizes surrounding each integration site (5 kb, 10 kb, and 50 kb), since previous studies have shown that the interval sizes chosen for comparison can influence the conclusions. In this study, results were in line for each interval size examined (data not shown). Only the data for the 10-kb interval are shown. In the heatmap, the distribution of experimental MLV sites is normalized to that of the MRC sites, as a control for recovery bias due to cleavage by restriction enzymes (in our case, MseI).[74,75] Results of statistical tests comparing the distributions of integration sites to the reference dataset are summarized as asterisks on each tile of the heatmap. Datasets used in the safe harbor analysis were retrieved from the Ensembl and/or UCSC databases (TxDB knownGenes, miRNA biotype, UCR; hg19) using BioMart.[79] The AllOnco list was used for oncogenes as published in Sadelain Analysis was performed according to the parameters defined in Papapetrou.

$CD34^+$ HSC CFU Assay

For human $CD34^+$ HSC CFU assays, freshly purified cells were seeded at 5,000 cells/well in human methylcellulose medium (methocult H4230; STEMCELL Technologies) supplemented with CC100 Cytokine Cocktail. Cells were plated in 35-mm petri dishes and cultured in a fully humidified atmosphere with 5% CO_2 at 37°C for 14 days. The number of colonies was scored after 14 days. For FACS analysis of the CFUs, colonies were harvested, washed twice with PBS, and evaluated for EGFP expression.

IVIM Assay

The IVIM assay was performed as described earlier.[46] Briefly, murine lineage marker-negative bone marrow cells were isolated from the tibias and femurs of C57BL/6 mice and enriched for stem and progenitor cells (mouse lineage cell depletion kit; Miltenyi Biotec) and frozen in aliquots. After thawing and 48 hr of prestimulation, 1×10^5 lin^- cells were transduced on 2 consecutive days on RetroNectin-coated (Takara; Clontech) wells with a MOI of 5, 10, or 20. TE was analyzed by flow cytometry 4 days thereafter. Cells were expanded for 2 weeks in IMDM, 10% FCS, 1 mM glutamine, 1% (v/v) penicillin/streptomycin, murine stem cell factor (mSCF) (50 ng/mL), human FMS-like tyrosine kinase 3 ligand (hFlt3L) (100 ng/mL), murine IL (mIL)-3 (20 ng/mL), and human IL (hIL)-11 (100 ng/mL; all cytokines purchased from Peprotech) and diluted to a cell density of 500,000 cells/mL approximately twice a week. Cells were the seeded on 96-well suspension plates at a density of 100 cells per well (48 wells seeded from each culture). Replating clones were detected by microscopic scoring. The replating frequency (according to Poisson distribution) was calculated with L-calc (STEMCELL Technologies) and normalized by VCN as determined 5 days post-transduction.

Murine Serial CFU Assay

Murine lineage marker-negative bone marrow (BM) cells were purified and transduced as in the IVIM assay. After 2 weeks of expansion, cells were plated (5,000 cells/well) in methylcellulose (HSC006; R&D Systems) supplemented with 20 ng/mL interleukin-3, and 50 ng/mL murine stem cell factor. The number of colonies was scored after 10 days. For serial replating, colonies were harvested and replated in fresh methylcellulose medium at the same density of 5,000 cells/well for the subsequent round.

SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures and four tables and can be found with this article online at http://dx.doi.org/10.1016/j.omtn.2017.04.002.

AUTHOR CONTRIBUTIONS

S.E.A., D.V.L., and R.G. designed the experiments. S.E.A., D.V.L., and R.G. wrote the manuscript. S.E.A., D.V.L., J.D.R., F.S., U.M., and R.G. performed experiments and analyzed data. J.D. and L.V. performed

bioinformatics for integration site analysis. R.G. and Z.D. supervised the project. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

We thank Paulien Van de Velde and Irina Thiry for excellent technical assistance, as well as Dr. Annelies Michiels for critical reading of the manuscript. Viral vector production was performed at the Leuven Viral Vector Core. This work was supported by grants from the KU Leuven Research Council (OT/13/098-3M130157), the KU Leuven Interdisciplinary Research Programmes (IDO; Interdisciplinaire onderzoeksprogramma's) (IDO/12/008-3E130241), and the Belgian IAP Belvir (P7/45-P-3M120222), IWT (3M140565 SB/151139), and FWO Vlaanderen (3M120412 Asp/12 and G0B3516N). U.M. and F.S. are supported by funding from the LOEWE Centre of Cell and Gene Therapy Frankfurt (HMWK III L 4-518/17.004 [2013]). S.EA is a doctoral fellow supported by Interdisciplinary Research Programmes. D.V.L and L.V. are doctoral fellows of the Research Foundation–Flanders (FWO-SB), and J.D. is a postdoctoral fellow of the Research Foundation–Flanders (FWO).

REFERENCES

1. Cavazzana-Calvo, M., Hacein-Bey, S., de Saint Basile, G., Gross, F., Yvon, E., Nusbaum, P., Selz, F., Hue, C., Certain, S., Casanova, J.L., et al. (2000). Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. Science *288*, 669–672.
2. Gaspar, H.B., Parsley, K.L., Howe, S., King, D., Gilmour, K.C., Sinclair, J., Brouns, G., Schmidt, M., Von Kalle, C., Barington, T., et al. (2004). Gene therapy of X-linked severe combined immunodeficiency by use of a pseudotyped gammaretroviral vector. Lancet *364*, 2181–2187.
3. Hacein-Bey-Abina, S., Le Deist, F., Carlier, F., Bouneaud, C., Hue, C., De Villartay, J.-P., Thrasher, A.J., Wulffraat, N., Sorensen, R., Dupuis-Girod, S., et al. (2002). Sustained correction of X-linked severe combined immunodeficiency by ex vivo gene therapy. N. Engl. J. Med. *346*, 1185–1193.
4. Cicalese, M.P., and Aiuti, A. (2015). Clinical applications of gene therapy for primary immunodeficiencies. Hum. Gene Ther. *26*, 210–219.
5. Aiuti, A., Slavin, S., Aker, M., Ficara, F., Deola, S., Mortellaro, A., Morecki, S., Andolfi, G., Tabucchi, A., Carlucci, F., et al. (2002). Correction of ADA-SCID by stem cell gene therapy combined with nonmyeloablative conditioning. Science *296*, 2410–2413.
6. Aiuti, A., Cassani, B., Andolfi, G., Mirolo, M., Biasco, L., Recchia, A., Urbinati, F., Valacca, C., Scaramuzza, S., Aker, M., et al. (2007). Multilineage hematopoietic reconstitution without clonal selection in ADA-SCID patients treated with stem cell gene therapy. J. Clin. Invest. *117*, 2233–2240.
7. Aiuti, A., Cattaneo, F., Galimberti, S., Benninghoff, U., Cassani, B., Callegaro, L., Scaramuzza, S., Andolfi, G., Mirolo, M., Brigida, I., et al. (2009). Gene therapy for immunodeficiency due to adenosine deaminase deficiency. N. Engl. J. Med. *360*, 447–458.
8. Cicalese, M.P., Ferrua, F., Castagnaro, L., Pajno, R., Barzaghi, F., Giannelli, S., Dionisio, F., Brigida, I., Bonopane, M., Casiraghi, M., et al. (2016). Update on the safety and efficacy of retroviral gene therapy for immunodeficiency due to adenosine deaminase deficiency. Blood *128*, 45–54.
9. Hoggatt, J. (2016). Gene therapy for "bubble boy" disease. Cell *166*, 263.
10. Howe, S.J., Mansour, M.R., Schwarzwaelder, K., Bartholomae, C., Hubank, M., Kempski, H., Brugman, M.H., Pike-Overzet, K., Chatters, S.J., de Ridder, D., et al. (2008). Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. J. Clin. Invest. *118*, 3143–3150.
11. Hacein-Bey-Abina, S., Garrigue, A., Wang, G.P., Soulier, J., Lim, A., Morillon, E., Clappier, E., Caccavelli, L., Delabesse, E., Beldjord, K., et al. (2008). Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. J. Clin. Invest. *118*, 3132–3142.
12. Ott, M.G., Schmidt, M., Schwarzwaelder, K., Stein, S., Siler, U., Koehl, U., Glimm, H., Kühlcke, K., Schilz, A., Kunkel, H., et al. (2006). Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1. Nat. Med. *12*, 401–409.
13. Stein, S., Ott, M.G., Schultze-Strasser, S., Jauch, A., Burwinkel, B., Kinner, A., Schmidt, M., Krämer, A., Schwäble, J., Glimm, H., et al. (2010). Genomic instability and myelodysplasia with monosomy 7 consequent to EVI1 activation after gene therapy for chronic granulomatous disease. Nat. Med. *16*, 198–204.
14. Braun, C.J., Boztug, K., Paruzynski, A., Witzel, M., Schwarzer, A., Rothe, M., Modlich, U., Beier, R., Göhring, G., Steinemann, D., et al. (2014). Gene therapy for Wiskott-Aldrich syndrome–long-term efficacy and genotoxicity. Sci. Transl. Med. *6*, 227ra33.
15. Hannan, G.N., Lehnert, S.A., MacAvoy, E.S., Jennings, P.A., and Molloy, P.L. (1993). An engineered PGK promoter and lac operator-repressor system for the regulation of gene expression in mammalian cells. Gene *130*, 233–239.
16. Montiel-Equihua, C.A., Zhang, L., Knight, S., Saadeh, H., Scholz, S., Carmo, M., Alonso-Ferrero, M.E., Blundell, M.P., Monkeviciute, A., Schulz, R., et al. (2012). The β-globin locus control region in combination with the EF1α short promoter allows enhanced lentiviral vector-mediated erythroid gene expression with conserved multilineage activity. Mol. Ther. *20*, 1400–1409.
17. Gaussin, A., Modlich, U., Bauche, C., Niederländer, N.J., Schambach, A., Duros, C., Artus, A., Baum, C., Cohen-Haguenauer, O., and Mermod, N. (2012). CTF/NF1 transcription factors act as potent genetic insulators for integrating gene transfer vectors. Gene Ther. *19*, 15–24.
18. Zhang, F., Frost, A.R., Blundell, M.P., Bales, O., Antoniou, M.N., and Thrasher, A.J. (2010). A ubiquitous chromatin opening element (UCOE) confers resistance to DNA methylation-mediated silencing of lentiviral vectors. Mol. Ther. *18*, 1640–1649.
19. Hacein-Bey-Abina, S., Pai, S.-Y., Gaspar, H.B., Armant, M., Berry, C.C., Blanche, S., Bleesing, J., Blondeau, J., de Boer, H., Buckland, K.F., et al. (2014). A modified γ-retrovirus vector for X-linked severe combined immunodeficiency. N. Engl. J. Med. *371*, 1407–1417.
20. Thornhill, S.I., Schambach, A., Howe, S.J., Ulaganathan, M., Grassman, E., Williams, D., Schiedlmeier, B., Sebire, N.J., Gaspar, H.B., Kinnon, C., et al. (2008). Self-inactivating gammaretroviral vectors for gene therapy of X-linked severe combined immunodeficiency. Mol. Ther. *16*, 590–598.
21. Zychlinski, D., Schambach, A., Modlich, U., Maetzig, T., Meyer, J., Grassman, E., Mishra, A., and Baum, C. (2008). Physiological promoters reduce the genotoxic risk of integrating gene vectors. Mol. Ther. *16*, 718–725.
22. Browning, D.L., Everson, E.M., Leap, D.J., Hocum, J.D., Wang, H., Stamatoyannopoulos, G., and Trobridge, G.D. (2017). Evidence for the in vivo safety of insulated foamy viral vectors. Gene Ther. *24*, 187–198.
23. Nelson, E.J.R., Tuschong, L.M., Hunter, M.J., Bauer, T.R., Jr., Burkholder, T.H., and Hickstein, D.D. (2010). Lentiviral vectors incorporating a human elongation factor 1alpha promoter for the treatment of canine leukocyte adhesion deficiency. Gene Ther. *17*, 672–677.
24. Carbonaro, D.A., Zhang, L., Jin, X., Montiel-Equihua, C., Geiger, S., Carmo, M., Cooper, A., Fairbanks, L., Kaufman, M.L., Sebire, N.J., et al. (2014). Preclinical demonstration of lentiviral vector-mediated correction of immunological and metabolic abnormalities in models of adenosine deaminase deficiency. Mol. Ther. *22*, 607–622.
25. Aiuti, A., Biasco, L., Scaramuzza, S., Ferrua, F., Cicalese, M.P., Baricordi, C., Dionisio, F., Calabria, A., Giannelli, S., Castiello, M.C., et al. (2013). Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome. Science *341*, 1233151.
26. Hacein-Bey Abina, S., Gaspar, H.B., Blondeau, J., Caccavelli, L., Charrier, S., Buckland, K., Picard, C., Six, E., Himoudi, N., Gilmour, K., et al. (2015). Outcomes following gene therapy in patients with severe Wiskott-Aldrich syndrome. JAMA *313*, 1550–1563.
27. Huston, M.W., van Til, N.P., Visser, T.P., Arshad, S., Brugman, M.H., Cattoglio, C., Nowrouzi, A., Li, Y., Schambach, A., Schmidt, M., et al. (2011). Correction of murine SCID-X1 by lentiviral gene therapy using a codon-optimized IL2RG gene and minimal pretransplant conditioning. Mol. Ther. *19*, 1867–1877.

28. Sharma, A., Larue, R.C., Plumb, M.R., Malani, N., Male, F., Slaughter, A., Kessl, J.J., Shkriabai, N., Coward, E., Aiyer, S.S., et al. (2013). BET proteins promote efficient murine leukemia virus integration at transcription start sites. Proc. Natl. Acad. Sci. USA *110*, 12036–12041.

29. De Rijck, J., de Kogel, C., Demeulemeester, J., Vets, S., El Ashkar, S., Malani, N., Bushman, F.D., Landuyt, B., Husson, S.J., Busschots, K., et al. (2013). The BET family of proteins targets Moloney murine leukemia virus integration near transcription start sites. Cell Rep. *5*, 886–894.

30. Gupta, S.S., Maetzig, T., Maertens, G.N., Sharif, A., Rothe, M., Weidner-Glunde, M., Galla, M., Schambach, A., Cherepanov, P., and Schulz, T.F. (2013). Bromo- and extraterminal domain chromatin regulators serve as cofactors for murine leukemia virus integration. J. Virol. *87*, 12721–12736.

31. El Ashkar, S., De Rijck, J., Demeulemeester, J., Vets, S., Madlala, P., Cermakova, K., Debyser, Z., and Gijsbers, R. (2014). BET-independent MLV-based vectors target away from promoters and regulatory elements. Mol. Ther. Nucleic Acids *3*, e179.

32. Kaustov, L., Ouyang, H., Amaya, M., Lemak, A., Nady, N., Duan, S., Wasney, G.A., Li, Z., Vedadi, M., Schapira, M., et al. (2011). Recognition and specificity determinants of the human Cbx chromodomains. J. Biol. Chem. *286*, 521–529.

33. Fischle, W., Franz, H., Jacobs, S.A., Allis, C.D., and Khorasanizadeh, S. (2008). Specificity of the chromodomain Y chromosome family of chromodomains for lysine-methylated ARK(S/T) motifs. J. Biol. Chem. *283*, 19626–19635.

34. Sekhar, V., Reed, S.C., and McBride, A.A. (2010). Interaction of the betapapillomavirus E2 tethering protein with mitotic chromosomes. J. Virol. *84*, 543–557.

35. Barbera, A.J., Chodaparambil, J.V., Kelley-Clarke, B., Luger, K., and Kaye, K.M. (2006). Kaposi's sarcoma-associated herpesvirus LANA hitches a ride on the chromosome. Cell Cycle *5*, 1048–1052.

36. Trobridge, G.D., Miller, D.G., Jacobs, M.A., Allen, J.M., Kiem, H.P., Kaul, R., and Russell, D.W. (2006). Foamy virus vector integration sites in normal human cells. Proc. Natl. Acad. Sci. USA *103*, 1498–1503.

37. Mitchell, R.S., Beitzel, B.F., Schroder, A.R.W., Shinn, P., Chen, H., Berry, C.C., Ecker, J.R., and Bushman, F.D. (2004). Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. PLoS Biol. *2*, E234.

38. De Ravin, S.S., Su, L., Theobald, N., Choi, U., Macpherson, J.L., Poidinger, M., Symonds, G., Pond, S.M., Ferris, A.L., Hughes, S.H., et al. (2014). Enhancers are major targets for murine leukemia virus vector integration. J. Virol. *88*, 4504–4513.

39. Mukherjee, S., and Thrasher, A.J. (2013). Gene therapy for PIDs: progress, pitfalls and prospects. Gene *525*, 174–181.

40. Wu, X., Li, Y., Crise, B., Burgess, S.M., and Munroe, D.J. (2005). Weak palindromic consensus sequences are a common feature found at the integration target sites of many retroviruses. J. Virol. *79*, 5211–5214.

41. Holman, A.G., and Coffin, J.M. (2005). Symmetrical base preferences surrounding HIV-1, avian sarcoma/leukosis virus, and murine leukemia virus integration sites. Proc. Natl. Acad. Sci. USA *102*, 6103–6107.

42. Maertens, G.N., Hare, S., and Cherepanov, P. (2010). The mechanism of retroviral integration from X-ray structures of its key intermediates. Nature *468*, 326–329.

43. Sadelain, M., Papapetrou, E.P., and Bushman, F.D. (2011). Safe harbours for the integration of new DNA in the human genome. Nat. Rev. Cancer *12*, 51–58.

44. Papapetrou, E.P., Lee, G., Malani, N., Setty, M., Riviere, I., Tirunagari, L.M.S., Kadota, K., Roth, S.L., Giardina, P., Viale, A., et al. (2011). Genomic safe harbors permit high β-globin transgene expression in thalassemia induced pluripotent stem cells. Nat. Biotechnol. *29*, 73–78.

45. Modlich, U., Schambach, A., Brugman, M.H., Wicke, D.C., Knoess, S., Li, Z., Maetzig, T., Rudolph, C., Schlegelberger, B., and Baum, C. (2008). Leukemia induction after a single retroviral vector insertion in Evi1 or Prdm16. Leukemia *22*, 1519–1528.

46. Modlich, U., Bohne, J., Schmidt, M., von Kalle, C., Knöss, S., Schambach, A., and Baum, C. (2006). Cell-culture assays reveal the importance of retroviral vector design for insertional genotoxicity. Blood *108*, 2545–2553.

47. McCulloch, E.A. (2003). Stem cells and diversity. Leukemia *17*, 1042–1048.

48. Borowicz, S., Van Scoyk, M., Avasarala, S., Karuppusamy Rathinam, M.K., Tauler, J., Bikkavilli, R.K., and Winn, R.A. (2014). The soft agar colony formation assay. J. Vis. Exp. *92*, e51998.

49. Kustikova, O.S., Wahlers, A., Kuhlcke, K., Stahle, B., Zander, A.R., Baum, C., and Fehse, B. (2003). Dose finding with retroviral vectors: correlation of retroviral vector copy numbers in single cells with gene transfer efficiency in a cell population. Blood *102*, 3934–3937.

50. Kang, E.M., Choi, U., Theobald, N., Linton, G., Long Priel, D.A., Kuhns, D., and Malech, H.L. (2010). Retrovirus gene therapy for X-linked chronic granulomatous disease can achieve stable long-term correction of oxidase activity in peripheral blood neutrophils. Blood *115*, 783–791.

51. Booth, C., Gaspar, H.B., and Thrasher, A.J. (2016). Treating immunodeficiency through HSC gene therapy. Trends Mol. Med. *22*, 317–327.

52. Newrzela, S., Cornils, K., Li, Z., Baum, C., Brugman, M.H., Hartmann, M., Meyer, J., Hartmann, S., Hansmann, M.L., Fehse, B., and von Laer, D. (2008). Resistance of mature T cells to oncogene transformation. Blood *112*, 2278–2286.

53. Maruggi, G., Porcellini, S., Facchini, G., Perna, S.K., Cattoglio, C., Sartori, D., Ambrosi, A., Schambach, A., Baum, C., Bonini, C., et al. (2009). Transcriptional enhancers induce insertional gene deregulation independently from the vector type and design. Mol. Ther. *17*, 851–856.

54. Montini, E., Cesana, D., Schmidt, M., Sanvito, F., Ponzoni, M., Bartholomae, C., Sergi Sergi, L., Benedicenti, F., Ambrosi, A., Di Serio, C., et al. (2006). Hematopoietic stem cell gene transfer in a tumor-prone mouse model uncovers low genotoxicity of lentiviral vector integration. Nat. Biotechnol. *24*, 687–696.

55. Modlich, U., Navarro, S., Zychlinski, D., Maetzig, T., Knoess, S., Brugman, M.H., Schambach, A., Charrier, S., Galy, A., Thrasher, A.J., et al. (2009). Insertional transformation of hematopoietic cells by self-inactivating lentiviral and gammaretroviral vectors. Mol. Ther. *17*, 1919–1928.

56. Cornils, K., Bartholomae, C.C., Thielecke, L., Lange, C., Arens, A., Glauche, I., Mock, U., Riecken, K., Gerdes, S., von Kalle, C., et al. (2013). Comparative clonal analysis of reconstitution kinetics after transplantation of hematopoietic stem cells gene marked with a lentiviral SIN or a γ-retroviral LTR vector. Exp. Hematol. *41*, 28–38.

57. Emery, D.W. (2011). The use of chromatin insulators to improve the expression and safety of integrating gene transfer vectors. Hum. Gene Ther. *22*, 761–774.

58. Antoniou, M.N., Skipper, K.A., and Anakok, O. (2013). Optimizing retroviral gene expression for effective therapies. Hum. Gene Ther. *24*, 363–374.

59. Moiani, A., Miccio, A., Rizzi, E., Severgnini, M., Pellin, D., Suerth, J.D., Baum, C., De Bellis, G., and Mavilio, F. (2013). Deletion of the LTR enhancer/promoter has no impact on the integration profile of MLV vectors in human hematopoietic progenitors. PLoS ONE *8*, e55721.

60. Trobridge, G.D. (2011). Genotoxicity of retroviral hematopoietic stem cell gene therapy. Expert Opin. Biol. Ther. *11*, 581–593.

61. Suerth, J.D., Maetzig, T., Brugman, M.H., Heinz, N., Appelt, J.-U., Kaufmann, K.B., Schmidt, M., Grez, M., Modlich, U., Baum, C., and Schambach, A. (2012). Alpharetroviral self-inactivating vectors: long-term transgene expression in murine hematopoietic cells and low genotoxicity. Mol. Ther. *20*, 1022–1032.

62. Kaufmann, K.B., Brendel, C., Suerth, J.D., Mueller-Kuller, U., Chen-Wichmann, L., Schwäble, J., Pahujani, S., Kunkel, H., Schambach, A., Baum, C., and Grez, M. (2013). Alpharetroviral vector-mediated gene therapy for X-CGD: functional correction and lack of aberrant splicing. Mol. Ther. *21*, 648–661.

63. Gijsbers, R., Ronen, K., Vets, S., Malani, N., De Rijck, J., McNeely, M., Bushman, F.D., and Debyser, Z. (2010). LEDGF hybrids efficiently retarget lentiviral integration into heterochromatin. Mol. Ther. *18*, 552–560.

64. Vets, S., De Rijck, J., Brendel, C., Grez, M., Bushman, F., Debyser, Z., and Gijsbers, R. (2013). Transient expression of an LEDGF/p75 chimera retargets lentivector integration and functionally rescues in a model for X-CGD. Mol. Ther. Nucleic Acids *2*, e77.

65. Vranckx, L.S., Demeulemeester, J., Debyser, Z., and Gijsbers, R. (2016). Towards a safer, more randomized lentiviral vector integration profile exploring artificial LEDGF chimeras. PLoS ONE *11*, e0164167.

66. Schneider, W.M., Brzezinski, J.D., Aiyer, S., Malani, N., Gyuricza, M., Bushman, F.D., and Roth, M.J. (2013). Viral DNA tethering domains complement replication-defective mutations in the p12 protein of MuLV Gag. Proc. Natl. Acad. Sci. U. S. A *110*, 9487–9492.

67. Elis, E., Ehrlich, M., Prizan-Ravid, A., Laham-Karam, N., and Bacharach, E. (2012). Tethers the murine leukemia virus pre-integration complex to mitotic chromosomes. PLoS Pathog. *8*, e1003103.

68. Rein, A. (2013). Murine leukemia virus p12 functions include hitchhiking into the nucleus. Proc. Natl. Acad. Sci. USA *110*, 9195–9196.

69. Hocum, J.D., Linde, I., Rae, D.T., Collins, C.P., Matern, L.K., and Trobridge, G.D. (2016). Retargeted foamy virus vectors integrate less frequently near proto-oncogenes. Sci. Rep. *6*, 36610.

70. Stein, S., Scholz, S., Schwäble, J., Sadat, M.A., Modlich, U., Schultze-Strasser, S., Diaz, M., Chen-Wichmann, L., Müller-Kuller, U., Brendel, C., et al. (2013). From bench to bedside: preclinical evaluation of a self-inactivating gammaretroviral vector for the gene therapy of X-linked chronic granulomatous disease. Hum. Gene Ther. Clin. Dev. *24*, 86–98.

71. Schambach, A., Bohne, J., Chandra, S., Will, E., Margison, G.P., Williams, D.A., and Baum, C. (2006). Equal potency of gammaretroviral and lentiviral SIN vectors for expression of O6-methylguanine-DNA methyltransferase in hematopoietic cells. Mol. Ther. *13*, 391–400.

72. Ibrahimi, A., Vande Velde, G., Reumers, V., Toelen, J., Thiry, I., Vandeputte, C., Vets, S., Deroose, C., Bormans, G., Baekelandt, V., et al. (2009). Highly efficient multicistronic lentiviral vectors with peptide 2A sequences. Hum. Gene Ther. *20*, 845–860.

73. Pizzato, M., Erlwein, O., Bonsall, D., Kaye, S., Muir, D., and McClure, M.O. (2009). A one-step SYBR Green I-based product-enhanced reverse transcriptase assay for the quantitation of retroviruses in cell culture supernatants. J. Virol. Methods *156*, 1–7.

74. Berry, C., Hannenhalli, S., Leipzig, J., and Bushman, F.D. (2006). Selection of target sites for mobile DNA integration in the human genome. PLoS Comput. Biol. *2*, e157.

75. Brady, T., Agosto, L.M., Malani, N., Berry, C.C., O'Doherty, U., and Bushman, F. (2009). HIV integration site distributions in resting and activated CD4+ T cells infected in culture. AIDS *23*, 1461–1471.

76. Ocwieja, K.E., Brady, T.L., Ronen, K., Huegel, A., Roth, S.L., Schaller, T., James, L.C., Towers, G.J., Young, J.A., Chanda, S.K., et al. (2011). HIV integration targeting: a pathway involving Transportin-3 and the nuclear pore protein RanBP2. PLoS Pathog. *7*, e1001313.

77. Marshall, H.M., Ronen, K., Berry, C., Llano, M., Sutherland, H., Saenz, D., Bickmore, W., Poeschla, E., and Bushman, F.D. (2007). Role of PSIP1/LEDGF/p75 in lentiviral infectivity and integration targeting. PLoS One *2*, e1340.

78. Brady, T., Lee, Y.N., Ronen, K., Malani, N., Berry, C.C., Bieniasz, P.D., and Bushman, F.D. (2009). Integration target site selection by a resurrected human endogenous retrovirus. Genes Dev. *23*, 633–642.

79. Kinsella, R.J., Kähäri, A., Haider, S., Zamora, J., Proctor, G., Spudich, G., Almeida-King, J., Staines, D., Derwent, P., Kerhornou, A., et al. (2011). Ensembl BioMarts: a hub for data retrieval across taxonomic space. Database (Oxford) *2011*, bar030.