

# Improved functionality and potency of next generation BinMLV viral vectors toward safer gene therapy

Dominique Van Looveren,<sup>1</sup> Giorgia Giacomazzi,<sup>2</sup> Irina Thiry,<sup>1</sup> Maurilio Sampaolesi,<sup>2</sup> and Rik Gijsbers<sup>1</sup>

<sup>1</sup>Laboratory for Viral Vector Technology and Gene Therapy, Department of Pharmacological and Pharmaceutical Sciences, KU Leuven, 3000 Leuven, Belgium; <sup>2</sup>Laboratory of Translational Cardiology, Department of Development and Regeneration, Stem Cell Research Institute, KU Leuven, 3000 Leuven, Belgium

**To develop safer retroviral murine leukemia virus (MLV)-based vectors, we previously mutated and re-engineered the MLV integrase: the W390A mutation abolished the interaction with its cellular tethering factors, BET proteins, and a retargeting peptide (the chromodomain of the CBX1 protein) was fused C-terminally. The resulting BET-independent MLV<sup>W390A-CBX</sup> was shown to integrate efficiently and more randomly, away from typical retroviral markers. In this study, we assessed the functionality and stability of expression of the redistributed MLV<sup>W390A-CBX</sup> vector in more depth, and evaluated safety using a clinically more relevant vector design encompassing a self-inactivated (SIN) LTR and a weak internal elongation factor 1 $\alpha$  short (EFS) promoter. MLV<sup>W390A-CBX</sup>-EFS produced like MLV<sup>WT</sup> and efficiently transduced laboratory cells and primary human CD34<sup>+</sup> hematopoietic stem cells (HSC) without transgene silencing over time, while displaying a more preferred, redistributed, and safer integration pattern. In a human mesoangioblast (MAB) stem cell model, the myogenic fusion capacity was hindered following MLV<sup>WT</sup> transduction, while this remained unaffected when applying MLV<sup>W390A-CBX</sup>. Likewise, smooth muscle cell differentiation of MABs was unaltered by MLV<sup>W390A-CBX</sup>-EFS. Taken together, our results underscore the potential of MLV<sup>W390A-CBX</sup>-EFS as a clinically relevant viral vector for *ex-vivo* gene therapy, combining efficient production with a preferable integration site distribution profile and stable expression over time.**

## INTRODUCTION

Integrating retroviral vectors have proven to be a powerful tool for long-term correction of genetic defects in a variety of severe blood and immune disorders.<sup>1–5</sup> However, in the initial clinical trials, therapeutic benefits were compromised by severe adverse events in a small subset of patients in the form of acute lymphoblastic leukemia and myelodysplastic syndromes.<sup>6–11</sup> These adverse events were directly attributed to the integration profile and the design of the integrated proviral genome, resulting in insertional mutagenesis.<sup>12</sup> Each retroviral family displays a specific integration profile that is orchestrated via the interaction of the retroviral integrase (IN) with specific host-cell co-factors that are co-opted by the viral PIC for integration. For instance, gammaretroviruses ( $\gamma$ RV; prototype murine leukemia

virus [MLV]) and their derived viral vectors integrate near strong enhancer and promoter regions,<sup>13–16</sup> while lentiviruses, such as HIV and HIV-based viral vectors, prefer integration in the body of actively transcribed gene bodies.<sup>17</sup> Hence, insertional mutagenesis is intrinsically linked to the integration preference of the proviral vector in the proximity of proto-oncogenes and the subsequent upregulation by the strong viral promoter and enhancer elements in the flanking long-terminal repeat (LTR) elements of the retroviral vectors, normally used to drive transgene expression, leading to aberrant expression of the proto-oncogene.<sup>9,18,19</sup>

Several efforts have been made to improve both safety and efficacy of integrating retroviral vectors by alterations in the vector design such as abolition of the strong enhancer elements from the U3 part of the LTR, resulting in self-inactivating (SIN) viral vectors.<sup>20–22</sup> The deletion of promoter/enhancer activity was counterbalanced by introduction of a heterologous promoter to drive transgene expression.<sup>23,24</sup> Nonetheless, SIN vector integration profiles remain unaltered and are still targeted to gene regulatory regions where they have the potential to disrupt or deregulate the transcription of nearby genes by other mechanisms.<sup>25</sup> Indeed, SIN viral vector designs still induced oncogene activation in *in-vitro* safety assays and resulted in a dominant myeloid clone in a  $\beta$ -thalassemia clinical trial.<sup>6,26,27</sup> However, replacing the potent viral promoters for a weaker cellular version, like elongation factor 1 $\alpha$  (EF1 $\alpha$ ) and phosphoglycerate kinase (PGK) promoters, in SIN-viral vectors greatly decreased their risk on insertional transformation.<sup>28,29</sup>

The specific integration preference of  $\gamma$ RV and their derived viral vectors for strong enhancers and promoter regions is dictated by interaction of the MLV pre-integration complex (PIC) with its cellular cofactors, the bromodomain and extra-terminal domain (BET)-containing family of proteins (BRD2, BRD3, and BRD4). We and others showed

Received 23 February 2021; accepted 16 July 2021;  
<https://doi.org/10.1016/j.omtm.2021.07.003>.

**Correspondence:** Rik Gijsbers, Laboratory for Viral Vector Technology and Gene Therapy, Department of Pharmacological and Pharmaceutical Sciences, KU Leuven, 3000 Leuven, Belgium.

**E-mail:** rik.gijsbers@kuleuven.be



that BET proteins serve as bimodal anchors on the host chromatin by binding to epigenetic chromatin marks associated with strong enhancers and promoter regions via their bromodomain on the one hand and to the MLV IN via the ET domain on the other.<sup>30–32</sup> In our previous work to improve the safety of retroviral vectors, we first uncoupled the interaction between the MLV IN and its cellular BET tethering factor, by a single substitution in the MLV IN (IN<sup>W390A</sup>) to create BET-independent (Bin) MLV vectors. BinMLV vectors produced and transduced efficiently, while displaying an integration pattern that associates less with traditional markers of MLV integration.<sup>33</sup> In a next step, we re-engineered the MLV IN by fusing small chromatin-binding peptides to the C-terminal end of the uncoupled IN<sup>W390A</sup> protein, such as a chromodomain peptide of CBX1 (referred to as IN<sup>W390A-CBX</sup>). The resulting MLV vectors (referred to as MLV<sup>W390A-CBX</sup>) in turn displayed an even more promising integration profile, with an integration site preference that was more random, detargeted away from the traditional markers of MLV integration.<sup>34</sup> Interestingly, MLV<sup>W390A-CBX</sup> produced well and transduced cells as efficiently as vectors carrying IN<sup>WT</sup> (MLV<sup>WT</sup>). In addition, this vector showed a reduced transformational potential compared to MLV<sup>WT</sup> in a serial colony-forming assay, supporting an improved safety profile.<sup>34</sup> Still, the vector used in these studies did not constitute the ideal clinical vector design, either being flanked by full MLV LTRs or carrying a potent viral spleen focus forming virus (SFFV) LTR promoter, known to deregulate genes up to 50 kb away, to drive transgene expression.<sup>27,28,34–36</sup>

In the present study, we assessed the functionality and the safety of the redistributed MLV<sup>W390A-CBX</sup> vector configuration compared to MLV<sup>WT</sup> using a clinically more relevant vector design combining a SIN LTR and a weak internal EF1 $\alpha$  short (EFS) promoter driving eGFP reporter gene expression. First, we analyzed transgene expression levels and expression stability over time in laboratory cells comparing MLV SIN vectors containing the EFS promoter as internal promoter with those using an SFFV promoter. The integration profile for MLV<sup>W390A-CBX</sup> was unaffected by the internal promoter and in line with our earlier data.<sup>34</sup> In a more detailed analysis, cells were sorted for eGFP expression prior to integration site sequencing to determine whether the more random integration pattern is retained in the active (eGFP<sup>+</sup>) population of cells. Our data showed that even when integration happens more randomly and thus also in regions with markers of silent chromatin, reporter gene expression is still supported, also when using the weaker EFS promoter. In line, MLV<sup>W390A-CBX</sup> transduction of clinically relevant human CD34<sup>+</sup> hematopoietic stem cell (HSC) corroborated stable transgene expression over time, with a more preferable and safer integration site distribution compared to MLV<sup>WT</sup>. In a final experiment, we questioned whether MLV<sup>W390A-CBX</sup>, displaying a potentially safer integration profile, remained active following differentiation of transduced progenitor cells. In a mesoangioblast (MAB) stem cell model, MLV<sup>W390A-CBX</sup> supported stable transgene expression over time, also following differentiation into skeletal and smooth muscle cells. In addition, the altered integration profile of MLV<sup>W390A-CBX</sup> resulted in a normal myogenic capacity of transduced mesoangioblasts compared to those using MLV<sup>WT</sup>, suggesting reduced genotoxicity of

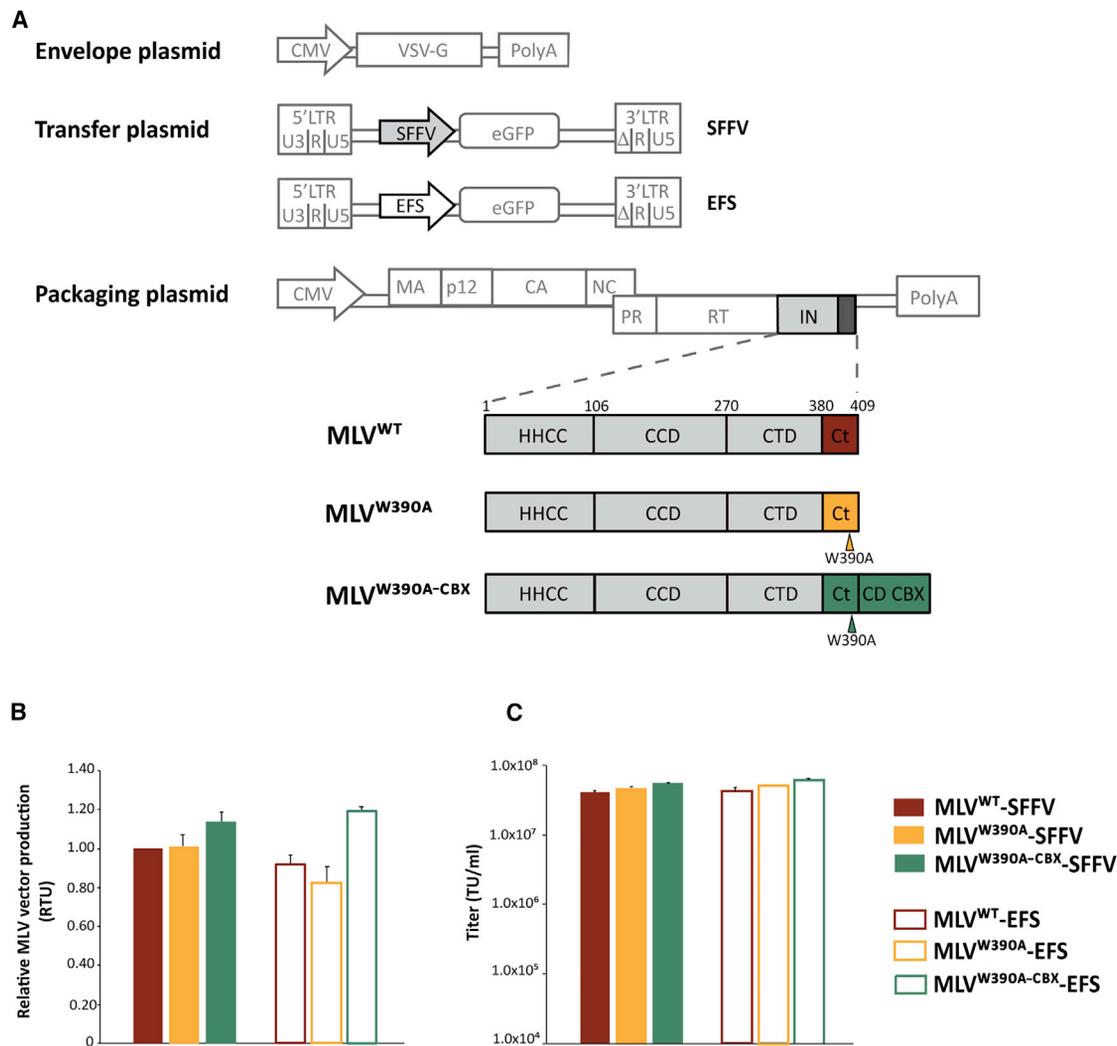
MLV<sup>W390A-CBX</sup> and demonstrating its value as viral vector for *ex vivo* gene therapy.

## RESULTS

### Vector design and production of next generation Bin MLV vectors

In a next step toward safer retroviral vectors for gene therapy, we opted to make a comprehensive analysis of the previously generated next-generation BinMLV IN<sup>W390A-CBX</sup> vector, further referred to as MLV<sup>W390A-CBX</sup> (Figure 1A; dark green), by implementing a clinically more relevant vector design.<sup>34</sup> In parallel, we used wild-type MLV (Figure 1A; MLV<sup>WT</sup>, dark red) and the Bin MLV vector only containing the point-mutation in IN (Figure 1A; MLV<sup>W390A</sup>, yellow) as controls.<sup>33</sup> We combined the respective packaging constructs with a self-inactivating LTR design and an internal EFS promoter, a short intron-less version derived from human EF1 $\alpha$ , to drive eGFP expression (Figure 1A; EFS, open bar).<sup>28,37</sup> As a reference, we used the same vectors carrying the potent SFFV promoter (Figure 1A; SFFV, filled bar). Vector particles pseudotyped with vesicular stomatitis virus glycoprotein (VSV-G) were produced, generating six different viral vectors: MLV<sup>WT</sup>-SFFV, MLV<sup>W390A</sup>-SFFV, MLV<sup>W390A-CBX</sup>-SFFV, MLV<sup>WT</sup>-EFS, MLV<sup>W390A</sup>-EFS, and MLV<sup>W390A-CBX</sup>-EFS, respectively. All viral vectors produced efficiently, as corroborated by comparable reverse transcriptase (RT) activities (Figure 1B; SFFV-driven filled bars, EFS-driven open bars) and transducing titers, with 293T titers reaching  $5 \times 10^7$  transducing units (TU)/mL (Figure 1C).

Following efficient MLV vector production, we assessed transduction of laboratory cell lines, SupT1 cells and K562 cells. Cells were transduced at three different multiplicities of infection (MOI) for the different vector configurations following normalization for RT activity. Transduction efficiencies were monitored over time by flow cytometry as %eGFP-positive (eGFP<sup>+</sup>) cells. The respective vectors reached similar transduction efficiencies for different MOIs in both cell lines (Figures S1A and S1B). When monitoring the stability of expression over time, %eGFP<sup>+</sup> cells slightly increased between days 3 and day 6 and was thereafter stable over time for both SFFV- and EFS-driven vector configurations (full and dashed lines, respectively) in SupT1 cells (Figure 2A; between 30.6% and 34.5% for SFFV vectors and between 37.9% and 41.6% for EFS vectors) as well as in K562 cells (Figure 2B; between 34.6% and 36.5% for SFFV vectors and between 38.8% and 43.0% for EFS vectors). Although transduction efficiencies were comparable for the respective vectors, mean fluorescence intensity (MFI) levels, serving as an indicator for eGFP transgene expression, showed to be different. In SupT1 cells an overall 1.6-fold difference was observed between SFFV-driven viral vectors and their EFS-driven counterparts (Figure 2C; Figure S1C). When comparing the respective vector configurations, MFI levels were 1.5-fold lower for MLV<sup>W390A</sup> and MLV<sup>W390A-CBX</sup> compared to MLV<sup>WT</sup> equipped with the same promoter (Figure 2C; Figure S1C), which may be explained by a retargeted integration profile. While differences in MFI were not that different in SupT1 cells, MFI levels in K562 cells of

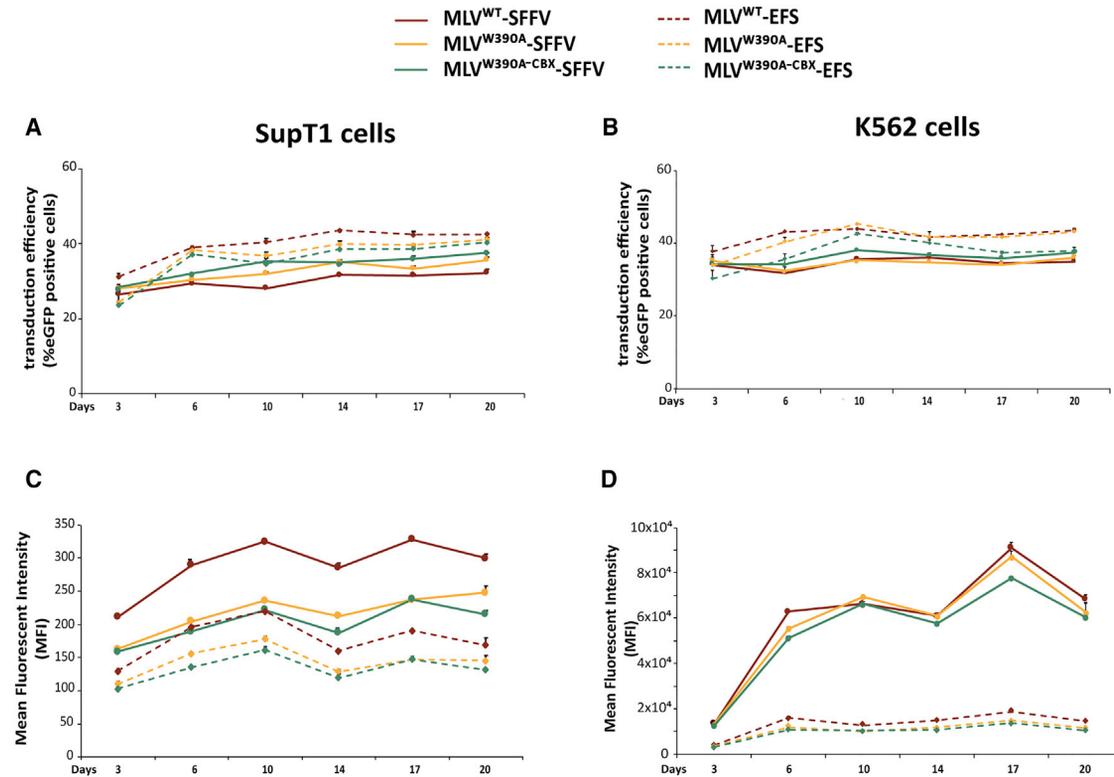


**Figure 1. Schematic overview of clinically relevant next-generation MLV vector designs**

(A) Schematic representation of the respective plasmids to produce MLV vectors, including different transfer plasmids (vector designs) with self-inactivating (SIN) architecture, carrying spleen focus forming virus promoter (SFFV) or elongation factor 1 $\alpha$  short promoter (EFS) as an internal promoter to drive eGFP expression (SFFV and EFS) and different packaging constructs, which upon production will result in different vector configurations (MLV<sup>WT</sup>, MLV<sup>W390A</sup>, and MLV<sup>W390A</sup>-CBX). Different MLV integrase (IN) configurations are depicted in more detail, composed by the N-terminal HHCC zinc binding domain, the catalytic core domain (CCD) and the C-terminal domain (CTD), containing the W390A point mutation in case of MLV<sup>W390A</sup> and the chromodomain peptide of CBX1 (CD CBX) fused to IN<sup>W390A</sup> in case of MLV<sup>W390A</sup>-CBX. (B) Relative MLV vector production determined by reverse transcriptase (RT) activity of the different MLV vectors as measured by SYBRGreen-I product-enhanced RT assay (SG-PERT), shown as RT units (RTUs). (C) Mean titers of the respective next-generation MLV vectors shown as transducing units (TU)/mL. Average values and standard deviations of triplicate measurements are shown. CMV, cytomegalovirus promoter; VSV-G, vesicular stomatitis virus glycoprotein G; poly(A), polyadenylation signal; LTR, long terminal repeat;  $\psi$ , packaging signal; eGFP, enhanced green fluorescent protein; MA, matrix; CA, capsid; NC, nucleocapsid; PR, protease; IN, integrase; MLV, murine leukemia virus.

EFS-containing vectors were on average 5-fold lower compared to the same vectors equipped with SFFV as an internal promoter (Figure 2D; Figure S1D). Interestingly, considering conditions with the same percentage of eGFP<sup>+</sup> cells, eGFP expression levels (MFI) of SFFV- and EFS-containing viral vectors were substantially higher in K562 cells compared to SupT1 cells (25-fold and 8-fold, respectively), indicating that both promoters are more functional in K562 cells.

Taken together, we assessed three different vector configurations (MLV<sup>WT</sup>, MLV<sup>W390A</sup>, and MLV<sup>W390A</sup>-CBX) that are all expected to distribute differently in the target cell genome. All three vectors efficiently transduced both SupT1 and K562 cells and resulted in stable reporter gene expression over time. Even though vectors equipped with SFFV displayed higher eGFP expression levels compared to those carrying a weaker EFS promoter, still expression over time was constant for both the EFS and SFFV vector designs.



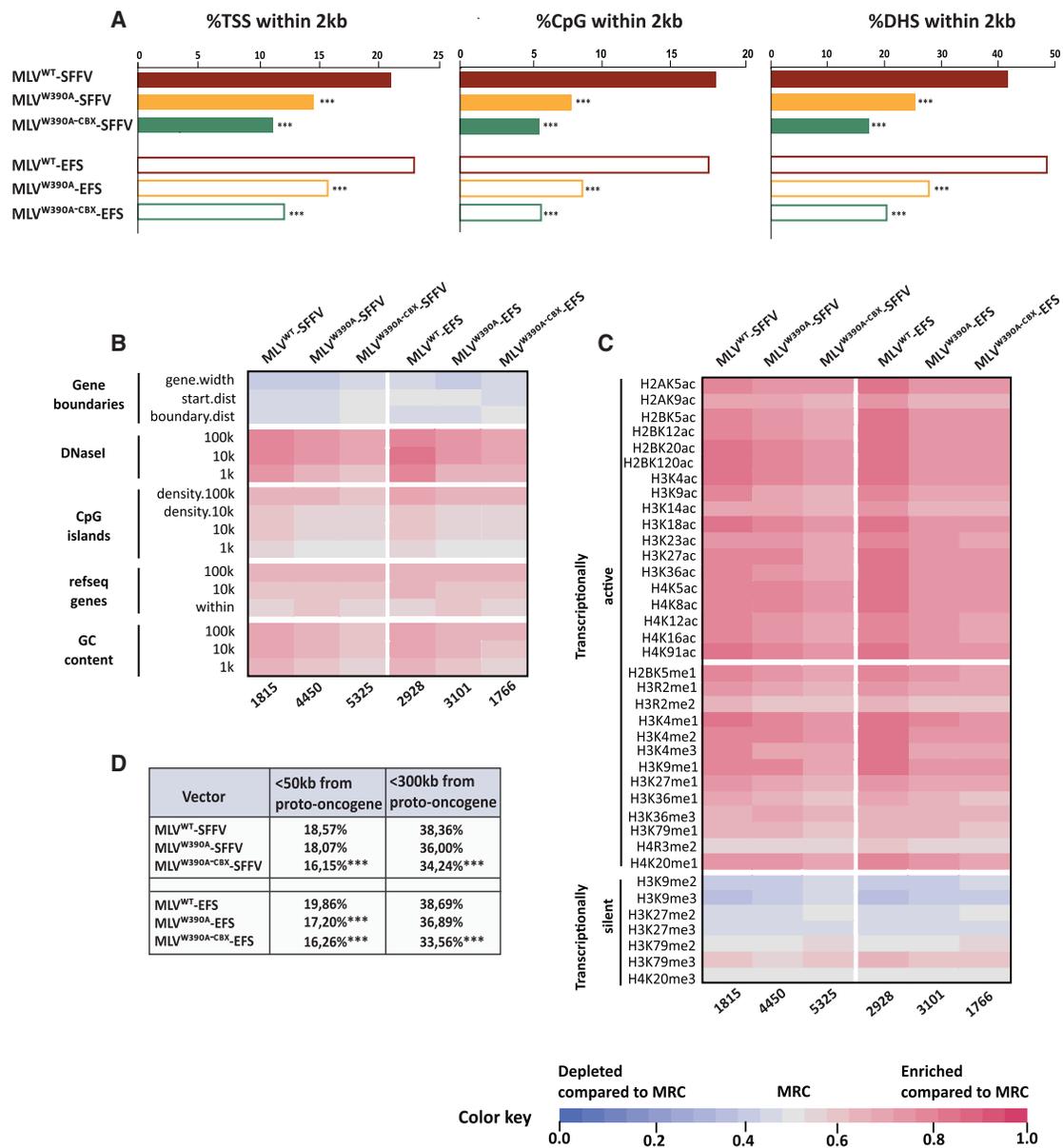
**Figure 2. Transduction efficiencies of clinically relevant next-generation MLV vectors in SupT1 and K562 cells**

FACS analysis of SupT1 cells (A and C) and K562 cells (B and D) transduced with equal RT-units of the indicated vectors at MOI of 1. Percentage of eGFP<sup>+</sup> cells (A and B) and mean fluorescence intensities (MFIs) (C and D) were measured by flow cytometry at different time points post transduction indicated in the figures (days). Data represent measurements from representative experiments out of three independent trials.

### Integration site profile of next-generation BinMLV vectors equipped with SFFV or EFS promoter is indistinguishable

In a next step, we determined the integration profile of the MLV vectors in SupT1 cells. Integration sites were recovered by ligation-mediated PCR (LM-PCR) followed by high-throughput Illumina sequencing and mapped to the human genome (hg18 assembly) by a previously described bioinformatic pipeline, yielding a total of 19,385 unique integration sites.<sup>38,39</sup> As an initial analysis, we examined the frequency of integration near typical markers associated with MLV integration (Figure 3A).  $\gamma$ RV integration is traditionally enriched near active enhancer and promoter regions.<sup>13,14</sup> In line, MLV<sup>WT</sup> integration was enriched near transcription start sites (TSS; 20%), CpG islands (CpG; 18%) and DNaseI hypersensitive sites (DHS; 40%) irrespective of the promoter driving the transgene expression (Figure 3A, dark red filled and open bars), whereas MLV<sup>W390A</sup> showed a profile that was significantly detargeted from these features (Figure 3A, yellow filled and open bars; \*\*\* $p < 0.001$ ; two-tailed  $\chi^2$  test compared to MLV<sup>WT</sup>), which is in line with earlier data.<sup>33,34</sup> Integration preferences for MLV<sup>W390A-CBX-SFFV</sup> and -EFS (dark green filled and open bars, respectively) near these features were shifted even more, as reported earlier.<sup>34</sup> In a more elaborate analysis, integration frequencies near a selection of genomic features and a collection of epigenetic marks known to associate with transcrip-

tionally active and silent chromatin were determined and represented as a genomic and an epigenetic heatmap, respectively (Figures 3B and 3C). Here, tile color depicts the correlation for an integration dataset with the respective genomic or epigenetic feature (indicated at the left side) relative to matched random controls (MRCs), as indicated by the colored receiver operating characteristic (ROC) curve area scale at the bottom of the panel. Pink tiles indicate that integration for a specific dataset is enriched in these features relative to MRC, while blue tiles indicate that integration is disfavored. A near random (MRC) distribution would result in a gray tile. Again, both MLV<sup>WT-SFFV</sup> and -EFS preferentially integrated in transcriptionally active chromatin (Figure 3C; tiles color dark pink) and disfavoring silent chromatin (Figure 3C; tiles color blue and gray). Uncoupling BET interaction (MLV<sup>W390A</sup>) resulted in an integration profile that associates less with the above-mentioned features (tiles color less pink), whereas MLV<sup>W390A-CBX</sup> viral vectors showed an integration profile with an even more pronounced shift toward random (Figure 3B). In addition, MLV<sup>W390A-CBX-SFFV</sup> and -EFS integration correlated less with histone modifications generally associated with active transcription, such as all acetylations and some histone methylations (Figure 3C; tile colors shift more toward gray), whereas integration occurred more frequently near H3K9me2 and H3K9me3, histone modifications associated with transcriptionally silent chromatin and

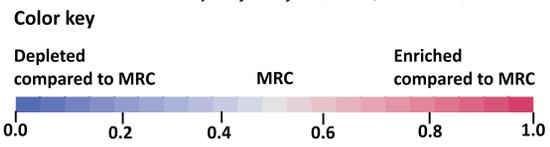
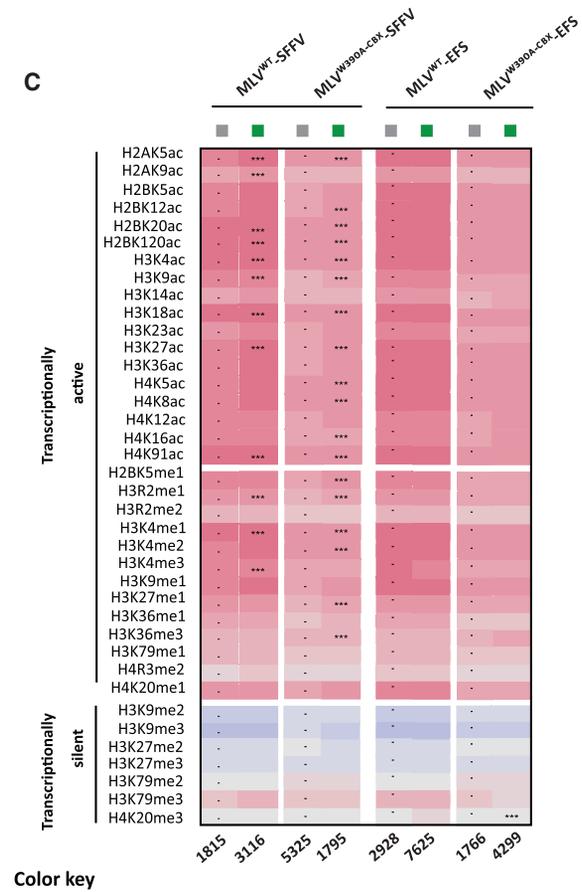
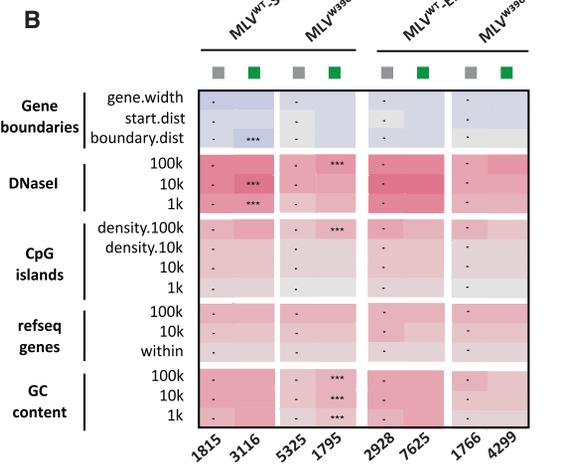
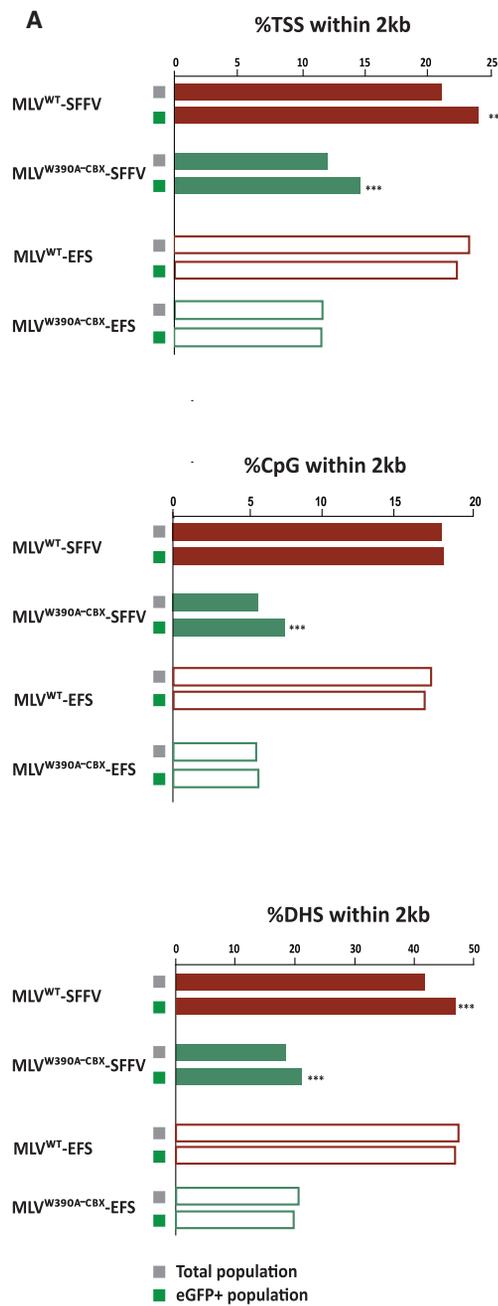


**Figure 3. Integration site distribution of clinically relevant next-generation MLV vectors**

Integration sites for the respective MLV-vector configurations and designs were obtained from SupT1 cells and their genomic distribution was determined. (A) Integration percentages in 2 kb windows around TSS, CpG, and DHS are listed. p values (\*) show significant departures from MLV<sup>WT</sup> (\*\*\*)  $p < 0.001$ , two-tailed  $\chi^2$  test). (B and C) Heatmaps summarizing the relation between vector integration site frequency and different genomic (B) or epigenetic features (C; 10 kb window) in SupT1 cells. Evaluated vector configurations are indicated above the columns. Features analyzed are shown to the left of the corresponding row of the heatmap. Tile color depicts the correlation for an integration dataset with the respective features relative to matched random controls (MRCs), as detailed in the colored receiver operating characteristic (ROC) area scale at the bottom of the panel. Number of unique integration sites is indicated below each column for the respective vector. (D) Integration frequencies for the respective viral vector designs near cancer-related oncogenes (<50 kb and <300 kb windows). p values show significance (\*\*\*)  $p < 0.001$ , two-tailed  $\chi^2$  test) compared to MLV<sup>WT</sup>. TSS, transcription start sites; CpG, CpG-rich island; DHS, DNase I-hypersensitive site.

described to be bound by CBX1 chromodomains<sup>40</sup> (Figure 3C; tile colors shift to gray for MLV<sup>W390A</sup> and MLV<sup>W390A-CBX</sup>). Asterisks indicate statistical significance of the integration site distributions of the respective MLV vectors relative to that of MLV<sup>WT</sup>-SFFV (Figure S2). Although differences were detected for some features, the

overall integration profile of EFS-containing MLV vector configurations was similar to that of their SFFV-driven counterparts, indicating that the endogenous promoter had no effect on the integration profile. Additionally, we determined integration site frequencies near cancer-related oncogenes in a 50 kb and 300 kb window. For both vector



(legend on next page)

designs, MLV<sup>W390A-CBX</sup> showed a significantly lower preference to integrate near oncogenes, corroborating an overall safer integration profile (Figure 3D, \*\*\*p < 0.001; two-tailed  $\chi^2$  test compared to MLV<sup>WT</sup> for each vector design separately).

#### Redistributed MLV W390A-CBX vector remains more active when carrying an EFS promoter

MLV<sup>W390A</sup> and MLV<sup>W390A-CBX</sup> vectors described a redistributed and more random integration profile compared to MLV<sup>WT</sup>, with less frequent integration near active chromatin markers and more near chromatin markers generally associated with transcriptionally silent chromatin compared to MLV<sup>WT</sup>. Nevertheless, transgene expression levels were comparable to that of MLV<sup>WT</sup> vectors (Figure 2). Since silenced integrations do not show in flow analysis but are picked up in integration site analysis, we wanted to corroborate that integrations that are shifted away from traditional markers of MLV integration do correlate with reporter gene expression. Indeed, the fact that the vector configuration leads to more random distribution for MLV<sup>W390A-CBX</sup>, and thus is potentially safer, does not imply that all redistributed vector integrations resulted in reporter gene expression, even though eGFP expression was in line for all vector designs. In fact, the observed eGFP expression of the MLV<sup>W390A-CBX</sup> vector might originate from integrations in transcriptionally active, and thus less favorable, regions with regard to viral vector safety. To investigate this, we transduced SupT1 cells at low transduction efficiency to ensure a single vector integration per cell and subsequently sorted them (eGFP<sup>+</sup> and eGFP<sup>-</sup> cells; Figures S3A and S3B). For simplicity and to reduce the number of samples to sequence, we opted to compare the vector configurations that showed the most profound difference in retargeting: MLV<sup>WT</sup> and MLV<sup>W390A-CBX</sup> (dark red and dark green, respectively). Integration sites were sequenced for the eGFP<sup>+</sup> fraction (Figure 4; green square), yielding a total of 16,835 unique integration sites and compared with the integration profile of the unsorted population (Figure 4; gray square). When considering the integration preferences of the SFFV-driven vectors, these showed significantly more integration near TSS and DHS for the eGFP<sup>+</sup> sorted cells (Figure 4A; 23.8% and 46.9% for MLV<sup>WT</sup> and 14.2% and 20.3% for MLV<sup>W390A-CBX</sup> near TSS and DHS, respectively) compared to the unsorted population for both MLV<sup>WT</sup> and MLV<sup>W390A-CBX</sup> vector configurations (20.7% and 40.1% for MLV<sup>WT</sup> and 11.7% and 18.3% for MLV<sup>W390A-CBX</sup> near TSS and DHS, respectively (Figure 4A; \*\*\*p < 0.001, two-tailed  $\chi^2$  test). Interestingly, for EFS-driven vectors integration preferences were not different when comparing integration preferences between the eGFP<sup>+</sup> cell population and the unsorted population, for both

the MLV<sup>WT</sup> and the MLV<sup>W390A-CBX</sup> vector configuration (Figure 4A), suggesting that EFS is less prone to positional effects.

These findings were further corroborated in the genomic and epigenetic heatmaps (Figures 4B and 4C). Asterisks indicate statistical significance of the integration site distributions of the eGFP<sup>+</sup> integration profile relative to that in the bulk population (dashed) for every vector separately. Although the integration profile of eGFP<sup>+</sup> MLV-SFFV vectors demonstrated significant differences for several features relative to that of the non-sorted population, integration profiles of the sorted and non-sorted cells were not different for MLV-EFS vectors (Figures 4B and 4C). Hence, the MLV<sup>W390A-CBX</sup>-EFS integration profile is redistributed out of gene regulatory elements without affecting transcriptional activity. In addition, integration near epigenetic markers associated with transcriptionally silent chromatin, like H3K9me2/3 and H3K27me2/3, didn't differ in the active population of cells for MLV<sup>W390A-CBX</sup>, both for the SFFV and the EFS-driven vector designs, suggesting that integration near these features was not different when sorting for eGFP<sup>+</sup> cells (Figure 4C).

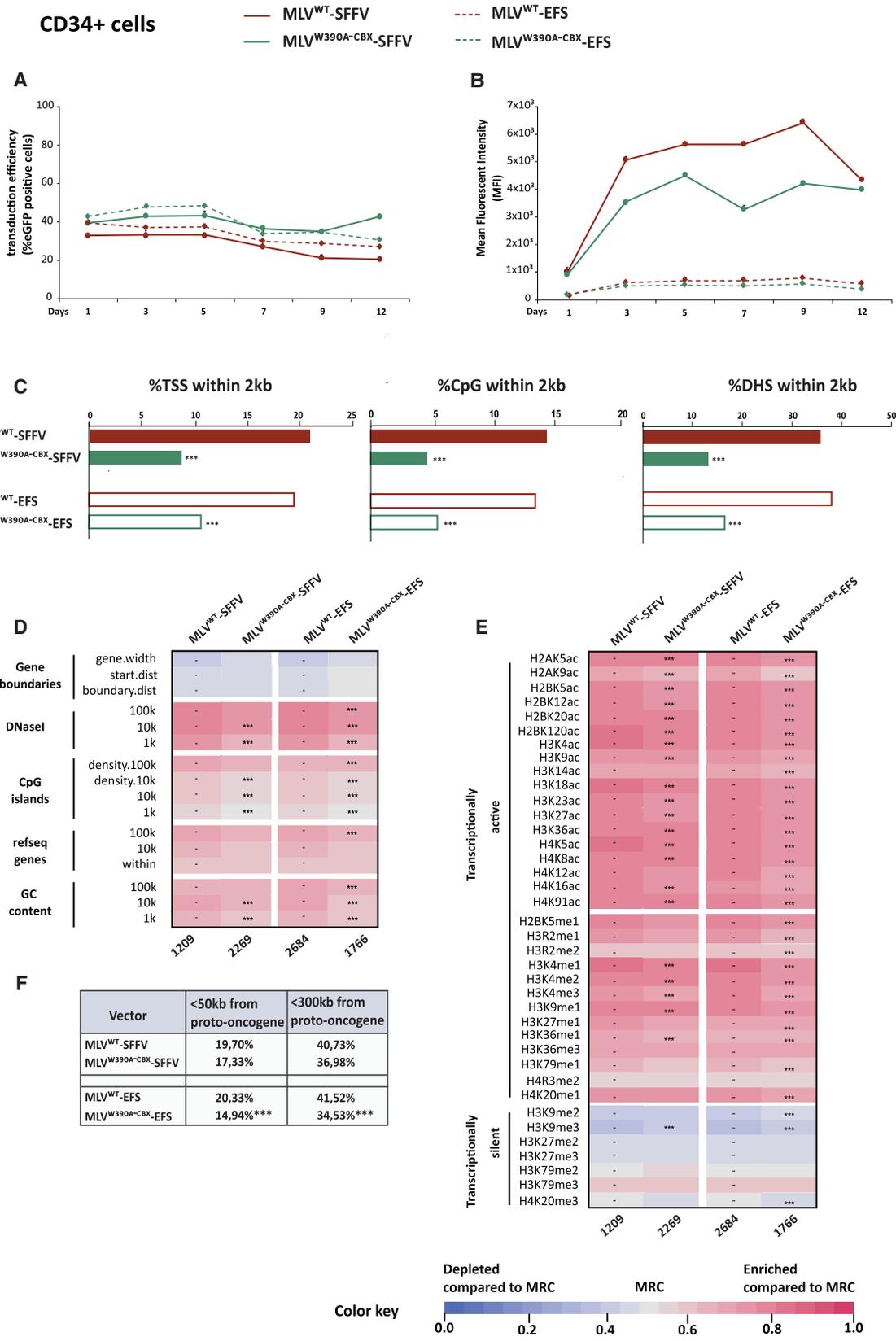
In conclusion, integration sites of eGFP<sup>+</sup> cells for SFFV-driven vectors were overall enriched near markers of active chromatin compared to MRC, whereas unsorted cells and eGFP<sup>+</sup> cells showed similar integration site distributions for EFS-driven vectors in all vector designs, suggesting that EFS driven viral vectors may be more interesting to use for gene therapeutic purposes. MLV<sup>W390A-CBX</sup>-EFS distributed more randomly, away from traditional markers of MLV integration, and supported active transcription at those positions.

#### Next-generation BinMLV vector shows a safer integration profile in hematopoietic stem cells

Stable transgene expression following integration in the host cell genome, especially when integration site choice is redistributed, is an important feature in the scope of long-term gene therapeutic correction. In a next step, we set out to evaluate transfer of next-generation BinMLV vectors to clinically more relevant cells. Primary human hematopoietic CD34<sup>+</sup> cells were transduced with either SFFV- and EFS-driven MLV<sup>WT</sup> or MLV<sup>W390A-CBX</sup> vector configurations at an MOI of 10. Therefore, transduction efficiency and expression level (eGFP fluorescence) were monitored at various time points up to 12 days after transduction. Although comparable percentages of eGFP<sup>+</sup> cells were obtained for all different MLV vectors (Figure 5A), MFI levels were on average 8-fold lower for EFS-containing vectors

#### Figure 4. Integration site distribution of clinically relevant next-generation MLV vectors in the active cell population

Murine leukemia virus (MLV)-based vector integration sites obtained from unsorted (gray square) and eGFP-sorted (green square, active) SupT1 cells and their genomic distribution. (A) Integration frequencies in 2 kb windows around TSS, CpG island midpoints and DHS are listed. p values (\*) show significant departures (\*\*\*p < 0.001, two-tailed  $\chi^2$  test) by comparing integration sites defined in the bulk population of cells with the integration profile in the eGFP<sup>+</sup> fraction for every vector separately. (B and C) Heatmaps summarizing the relation between vector integration site frequency from the eGFP-expressing SupT1 cells and different genomic (B) or epigenetic features within 10 kb interval (C). Evaluated vectors are indicated above the columns. Features analyzed are shown to the left of the corresponding row of the heatmap. Tile color depicts the correlation for an integration dataset with the respective features relative to matched random controls, 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. Statistical significance (asterisks, \*\*\*p < 0.001, Wald statistics referred to  $\chi^2$  distribution) is shown relative to the bulk population of cells for every vector separately (double dash). Grey square, unsorted cells; green square, eGFP<sup>+</sup> cells.



(legend on next page)

compared to SFFV-driven vectors and remained stable over time for both promoter set-ups (Figure 5B). Subsequently, we determined the integration profile of the MLV vectors in CD34<sup>+</sup> HSC. Integration sites were recovered by LM-PCR, followed by high-throughput Illumina sequencing and mapped to the human genome (hg18 assembly), yielding a total of 1,000–3,000 unique integration sites for each individual vector.<sup>38,39</sup> In line with the results obtained in laboratory cell lines, MLV<sup>W390A-CBX</sup> showed a significant reduction of integration near TSS, CpG islands, and DHS sites compared to MLV<sup>WT</sup>, underscoring redistribution of integration sites (Figure 5C, \*\*\*p < 0.001; two-tailed  $\chi^2$  test). Also, a more thorough analysis of integration frequency near genomic regions (Figure 5D) and epigenetic marks (Figure 5E) corroborated the effects seen in SupT1 cells before. MLV<sup>W390A-CBX</sup> integration displayed a more random integration site distribution compared to MLV<sup>WT</sup> for both promoter configurations in HSCs (Figures 5D and 5E; tile colors shift toward gray). For both SFFV and EFS vector designs, asterisks indicate significant integration site redistributions of the MLV<sup>W390A-CBX</sup> relative to MLV<sup>WT</sup> (double dash). Additionally, we analyzed integration frequencies within a 50 kb and 300 kb window surrounding known (proto-)oncogenes, confirming that also in CD34<sup>+</sup> HSCs MLV<sup>W390A-CBX</sup> integrated less frequently near oncogenes compared to MLV<sup>WT</sup> (Figure 5F, \*\*\*p < 0.001; two-tailed  $\chi^2$  test compared to MLV<sup>WT</sup>). Taken together, these findings demonstrated the ability of the next-generation MLV<sup>W390A-CBX</sup> to transduce clinically relevant HSCs with stable reporter expression over time, while distributing a more random integration profile that associates less with oncogenes.

### Functionality of retargeted next-gen MLV W390A-CBX in MAB model

Several gene therapeutic applications require transduction of stem cells that in turn mature to give rise to genetically corrected, differentiated cells. Vector integration occurs in the stem cell stage, and upon differentiation, the chromatin of the cell is reorganized to allow transcription of cell-specific programs. An ideal integrating viral vector (i) integrates efficiently in safe positions, but in addition (ii) does not affect the differentiation potential of stem cells and (iii) remains transcriptionally active upon differentiation. In order to test this, we further validated our next-gen MLV vectors in a human MAB stem cell model. MABs are vessel-associated multipotent stem cells that

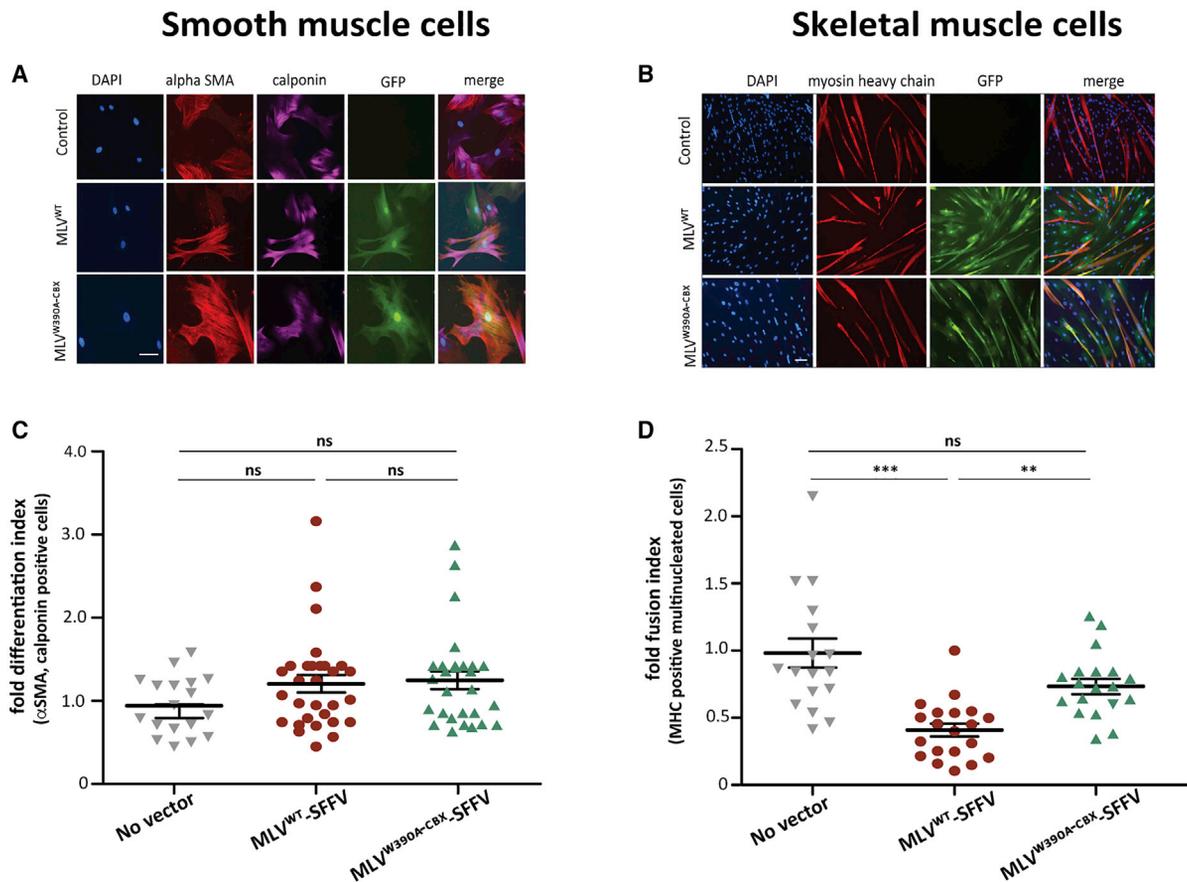
can differentiate into mesodermal lineages, including skeletal and smooth muscle cells. Hence, MABs have been extensively explored within the scope of treating muscular dystrophies.<sup>41,42</sup> In the current work, human MABs were purified from skeletal striated muscle from healthy subjects and transduced with the respective MLV vectors, resulting in comparable transduction efficiencies (Figure S4A) and eGFP expression levels (MFI) for both the SFFV and EFS vector designs (Figure S4B).

Subsequently, we evaluated the potential of MABs to differentiate into smooth and skeletal muscle cells following transduction. We used two MAB cell lines isolated in house from 2 different donors (referred to as MAB donor 1 and MAB donor 2) that each were transduced at an MOI of 3 with MLV<sup>WT</sup> or MLV<sup>W390A-CBX</sup> vector configuration containing the SFFV design (Figure 6). 3 days post transduction, transduced and non-transduced (NT) control cells were seeded in differentiation medium to differentiate toward smooth muscle or skeletal muscle cells over the course of 8 days (see cartoon Figure S5A). Smooth muscle cell differentiation was evaluated by immunofluorescent staining for smooth muscle cell markers, alpha smooth muscle actin (SMA) and calponin (Figure 6A), and the differentiation index was computed (Figures S6A and S6B). In parallel, skeletal muscle differentiation was scored by assessing their potential to fuse and form myotubes as shown by myosin heavy chain (MHC) and nuclei (DAPI) staining (Figure 6B). The fusion index was defined as the percentage of nuclei incorporated in the MHC<sup>+</sup> myotubes relative to the total number of nuclei (Figures S6C and S6D). Fold differentiation and fusion indices were calculated to allow pooling of the data obtained in both MAB cell lines for the SFFV vector design (Figures S6A and S6B for smooth muscle differentiation in Figure 6C and Figures S6C and S6D for skeletal muscle differentiation in Figure 6D). Whereas for smooth muscle differentiation, MLV<sup>WT</sup>-SFFV and MLV<sup>W390A-CBX</sup>-SFFV transduced MABs showed a fold differentiation index in line with untransduced MABs (Figure 6C; no significant [ns] difference to no vector control), differentiating the same transduced MABs to skeletal muscle cells was affected following transduction with MLV<sup>WT</sup> (Figure 6D; significantly lower fusion potential compared to untreated control MABs [no vector] and to MLV<sup>W390A-CBX</sup> (Figures 6C and 6D; Figures S6A–S6D; Kruskal-Wallis with uncorrected Dunn's test, \*\*p < 0.01, \*\*\*p < 0.001). The next-gen

### Figure 5. Integration site distribution analysis of next-generation MLV W390A-CBX vector in clinically relevant human CD34<sup>+</sup> HSCs

MLV-based vector integration sites were determined in primary human CD34<sup>+</sup> HSCs following transduction with the respective vectors, and their genomic distribution was assessed. MLV<sup>W390A-CBX</sup> and MLV<sup>WT</sup> vector designs were used, carrying an internal SFFV or EFS promoter to drive transgene expression. (A and B) CD34<sup>+</sup> HSC were transduced with the respective vectors at equal RT-units at an MOI of 10. Percentage of enhanced eGFP<sup>+</sup> cells (A) and MFIs (B) were measured by flow cytometry at different time points post transduction indicated in the figures (days). Data represent measurements from a representative experiment out of two independent trials. (C) Integration frequencies in 2 kb windows around TSS, CpG island midpoints, and DHS are listed. p values (\*) show significant departures (\*\*\*p < 0.001, two-tailed  $\chi^2$  test) from the respective MLV<sup>WT</sup> condition. (D and E) Heatmaps summarizing the relation between vector integration site frequency from the eGFP-expressing SupT1 cells and different genomic (D) or epigenetic features (E; 10 kb window). Evaluated vector configurations are indicated above each column. Features analyzed are shown to the left of the corresponding row of the heatmap. Tile color depicts the correlation for an integration dataset with the respective features relative to matched random controls, 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. Statistical significance (asterisks, \*\*\*p < 0.001, Wald statistics referred to  $\chi^2$  distribution) is shown relative to the MLV<sup>WT</sup> condition for each promoter configuration (double dash). Number of unique integration sites is indicated below each column for the respective vector. (F) Integration frequencies for the respective viral vector designs near cancer-related oncogenes (<50 kb and <300 kb windows). p values show significance (\*\*\*p < 0.001, two-tailed  $\chi^2$  test) compared to MLV<sup>WT</sup>.

## SFFV Design



**Figure 6. Assessment of next-generation BET-independent MLV-SFFV vectors in a mesoangioblast (MAB) model**

Smooth muscle differentiation (A and C) and myogenic differentiation (C and D) of human MAB that were transduced with MLV<sup>WT</sup> or MLV<sup>W390A-CBX</sup> vectors. MLV<sup>WT</sup> or MLV<sup>W390A-CBX</sup> vectors containing SFFV design are shown in (A)–(D). Differentiation in smooth muscle cells was determined by counting alpha smooth muscle actin (alpha SMA) and calponin<sup>+</sup> cells and myogenic fusion by counting myosin heavy chain (MHC) multinucleated myotubes. Fold differentiation indexes (B) and fold fusion indexes (D) show the combination of experiments performed in two different MAB cell clones (median  $\pm$  SD). White scale bars are 100  $\mu$ m. Statistical significance between groups was calculated by a non-parametric ANOVA, Kruskal-Wallis with uncorrected Dunn's test (ns, non-significant; \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ).

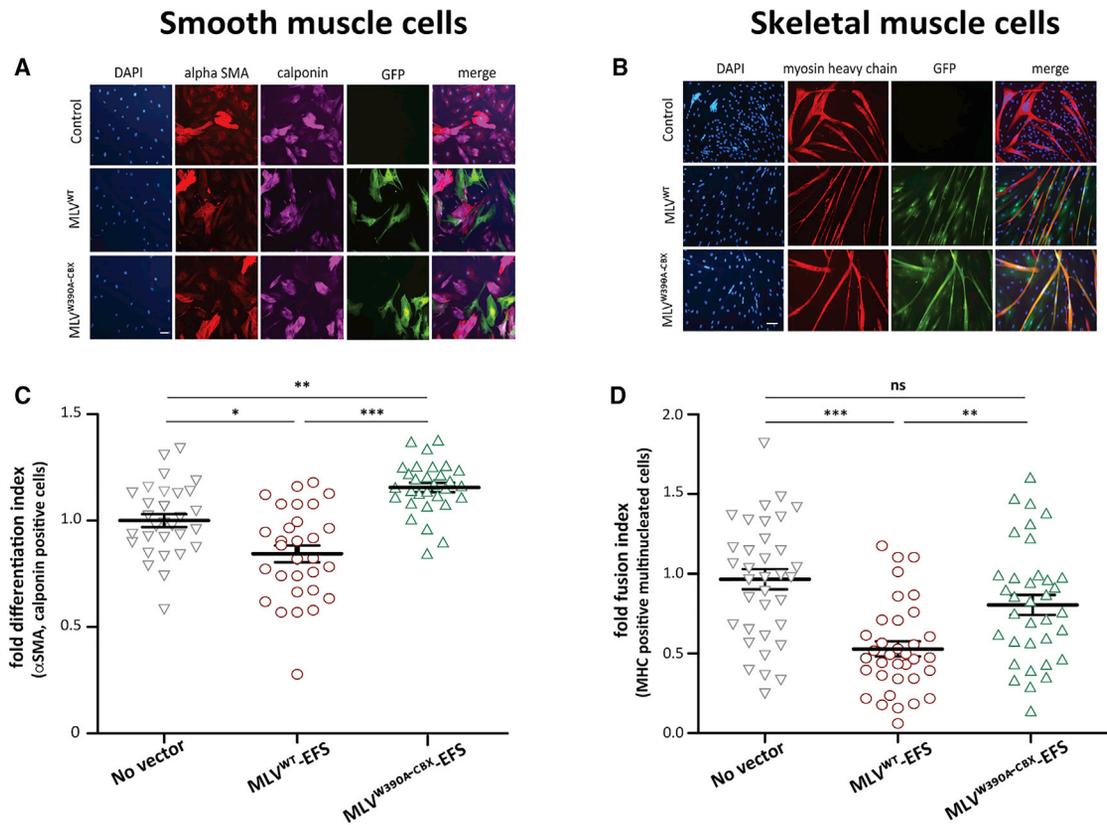
MLV<sup>W390A-CBX</sup>-SFFV configuration, on the other hand, did not disturb the myogenic capacity of MABs, with a fold fusion index that was not significantly different from NT, control MABs (Figure 6D; Kruskal-Wallis with uncorrected Dunn's test, ns).

Next, the same experiments were performed independently for vectors containing the EFS design (Figures S6E–S6H; Figures 7A–7D). MABs transduced with MLV<sup>WT</sup>-EFS showed a comparable fold differentiation index compared to no vector control MABs that was only weakly significant (Figures 7A and 7C; Kruskal-Wallis with uncorrected Dunn's test, \* $p < 0.05$ ). MLV<sup>W390A-CBX</sup>-EFS treated MABs showed a slightly higher differentiation index than no vector MABs (Figure 7C; Kruskal-Wallis with uncorrected Dunn's test, \*\* $p < 0.001$ ). For skeletal muscle differentiation, similar EFS-containing vector transduction results were obtained as with the SFFV promoter. Whereas MABs transduced with MLV<sup>WT</sup>-EFS showed a significantly

lower ability to differentiate into skeletal muscle cells compared to NT cells, the fold fusion index was not affected when MABs were transduced with MLV<sup>W390A-CBX</sup>-EFS (Figures 7B and 7D; Kruskal-Wallis with uncorrected Dunn's test, \*\* $p < 0.01$ , \*\*\* $p < 0.001$ ). Intriguingly, MLV<sup>W390A-CBX</sup>-EFS showed a higher fusion index in MABs than MLV<sup>WT</sup>-EFS, underscoring the potential of MLV<sup>W390A-CBX</sup>-EFS as viral vector (Figure 7D; Kruskal-Wallis with uncorrected Dunn's test, \*\* $p < 0.01$ , \*\*\* $p < 0.001$ ).

In addition, we assessed whether the integrated vectors supported transgene expression following differentiation and evaluated the contribution of vector transduced cells by counting the number of eGFP<sup>+</sup> cells in fully differentiated cells based on the cellular markers mentioned before. Both the myogenic and smooth muscle differentiation strategies resulted in the same number of vector-transduced cells (eGFP<sup>+</sup> cells) that contributed to the differentiated cell pool

## EFS Design



**Figure 7. Investigation of clinically relevant next-generation MLV-EFS vectors in a MAB model**

Smooth muscle differentiation (A and C) and myogenic differentiation (C and D) of human MAB that were transduced with MLV<sup>WT</sup> or MLV<sup>W390A-CBX</sup> vectors. MLV<sup>WT</sup> or MLV<sup>W390A-CBX</sup> vectors containing EFS design are shown in (A)–(D). Differentiation in smooth muscle cells was determined by counting alpha SMA and calponin<sup>+</sup> cells and myogenic fusion by counting MHC multinucleated myotubes. Fold differentiation indexes (B) and fold fusion indexes (D) show the combination of experiments performed in two different MAB cell clones (median ± SD). White scale bars are 100 μm. Statistical significance between groups was calculated by a non-parametric ANOVA, Kruskal-Wallis with uncorrected Dunn's test (ns, non-significant; \*p < 0.05; \*\*p < 0.01; \*\*\*p < 0.001).

for both SFFV and EFS vector designs (Figures S7A–S7D), which is in line with the original transduction efficiencies.

Taken together, our results showed that human mesoangioblasts transduced with the MLV<sup>W390A-CBX</sup> configuration have a comparable potential to differentiate into smooth and skeletal muscle cells as untreated MABs, whereas MABs transduced with MLV<sup>WT</sup> were hampered in fusing and formation of multinucleated myotubes. Moreover, equipping MLV<sup>W390A-CBX</sup> with the EFS design elicits an improved differentiation into both smooth and skeletal muscle cells compared to MLV<sup>WT</sup>, rendering MLV<sup>W390A-CBX</sup>-EFS as a fine and possibly even better alternative as MLV<sup>WT</sup> to use as viral vector for gene therapeutic applications.

## DISCUSSION

*Ex-vivo* gene therapy holds great promise to provide a cure for many disorders and has been particularly successful in treating diseases of the hematopoietic system. Due to their stable integration of the trans-

gene, retroviral-based vectors have shown their therapeutic benefit in the past decades in gene therapy studies for several primary immunodeficiencies (PID), such as X-linked severe combined immunodeficiency (X-SCID), adenosine-deaminase SCID (ADA-SCID), and Wiskott-Aldrich syndrome (WAS). MLV-based vectors were successfully applied in the first clinical trial for ADA-SCID, which led in 2016 to the approval by the European Medicines Agency (EMA) of the first retroviral vector-based gene therapy product in Europe (Strimvelis), to treat patients that lack a suitable HLA-matched stem cell donor.<sup>43,44</sup> However, in gene therapy trials to treat other PIDs, the same vector design resulted in severe adverse effects in a small subset of patients in the form of acute lymphoblastic leukemia and myelodysplastic syndromes in a subset of diseases, raising concerns on the safety of retroviral vectors for gene therapy.<sup>8–11</sup> The understanding of the mechanisms that contributed to tumorigenesis of the therapeutic vectors provided new insights to engineer vectors that transduce the target cells with the same efficacy as the first-generation retroviral vectors but that display a lower oncogenic potential.<sup>12</sup>

Approaches that have been implemented include the deletion of the strong enhancer elements from the U3 part of the LTR, resulting in SIN viral vectors,<sup>27,36,45</sup> the compensation of the loss of enhancer activity by heterologous endogenous promoters, or even cell-specific promoters and the inclusion of insulator elements that act as both enhancer-blocking and a boundary against potential silencing of the viral vectors.<sup>28,29,46–48</sup> Furthermore, alternative vector systems have been developed based on viral strains with a more favorable integration profile, such as foamy viral (FV) or alpharetroviral vectors<sup>49–52</sup> and also retrotransposon-based gene transfer has been improved with the Sleeping Beauty transposase displaying a near-to-random integration profile.<sup>53–55</sup>

In 2017, we reported on the fusion of small peptide sequences to the C-terminal tail of an MLV variant that was unable to interact with BET-proteins (IN<sup>W390A</sup>).<sup>34</sup> Fusion of the chromodomain of the heterochromatin binding protein 1 $\beta$  (CBX1) protein turned out to be the best option. The resulting Bin MLV<sup>W390A-CBX</sup> viral vector showed an improved safety profile with integration site preferences that are more random, detargeted away from the traditional markers of MLV integration and showed a reduced risk of hematopoietic cell transformation when tested in genotoxicity assays.<sup>34</sup> A comparable approach where the full CBX1 protein was fused to the FV IN was shown to be effective in FV vectors.<sup>56</sup>

In the present work, we validated the next-generation MLV<sup>W390A</sup> and MLV<sup>W390A-CBX</sup> vectors in more depth, using a clinically relevant vector design by implementing an endogenous shortened version of the EF1 $\alpha$  (EFS) promoter in a SIN vector design. Although EFS-driven vectors transduced cells as efficiently as SFFV-driven vectors (same %eGFP<sup>+</sup> cells), reporter expression levels (MFI) differed for both promoters depending on which cell type was used. No difference between SFFV and EFS promoters was found in MABs in terms of mean fluorescent intensity and also in SupT1 cells differences were rather small (1.6-fold). In K562 and primary human CD34<sup>+</sup> HSCs on the other hand, the effect was more pronounced with MFI levels that were 6- to 8-fold lower for EFS-driven vectors compared to the same vectors using the SFFV instead, indicating that promoter activity is substantially different in specific cells. These observations can partially be explained by the difference of lineage-dependent transcription factor binding motifs between the two promoters. Overall, these results corroborate the more subtle modulation of transgene expression when using the EFS promoter compared to SFFV, as reported already in earlier studies.<sup>28,57</sup>

Investigation of the integration profile revealed that the internal promoter used to drive expression of the transgene didn't affect the integration preference of the viral vector in laboratory cell lines (SupT1) and in clinically more relevant primary human CD34<sup>+</sup> HSCs. In line with earlier results, integration site distribution of MLV<sup>W390A-CBX</sup> was retargeted away from typical regions for MLV integration, like TSS, DHS sites, and CpG islands and increased in regions enriched in heterochromatin markers. Additionally, we corroborated that MLV<sup>W390A-CBX</sup> integrated significantly less near oncogenes compared to MLV<sup>WT</sup>, both in SupT1 cells and in primary CD34<sup>+</sup>

HSC, underscoring that MLV<sup>W390A-CBX</sup> describes a safer integration pattern. From a gene therapeutic perspective, it is important that integration sites near heterochromatin histone modifications support effective transgene expression. However, not all vector integrations will result in expression, especially when less potent promoters are used. Moreover, the transcriptional potential of the viral vector is believed to be associated with the chromatin environment in the neighborhood of the integration site that might result in vector silencing. Usually when integration sites are determined, there is no information on the expression level of the contributing integration sites, which occasionally may even be silenced. Hence, we set out to sort cells based on eGFP marker expression and determined the integration profile only in the eGFP<sup>+</sup> cell population. Since MLV<sup>W390A-CBX</sup> demonstrated the most prominent shift in integration site distribution, integration site preferences were analyzed for SupT1 cells transduced with WT MLV and MLV<sup>W390A-CBX</sup> that were subsequently sorted for eGFP. Interestingly, integration preferences in eGFP<sup>+</sup> cells were comparable to those in the bulk population when using the weaker EFS promoter for both MLV vector configurations. Nonetheless, the active integration profile (eGFP<sup>+</sup>) of SFFV-driven MLV vectors was different compared to that of the unsorted cell population, and more targeted near epigenetic marks associated with active transcription. The fact that the integration profile of EFS-driven vectors remained unaltered in the active population might be because this promoter is less prone to positional effects compared to the SFFV promoter. Especially in induced pluripotent stem cells (iPSCs) and human embryonic stem cells (hESCs) it was already shown before that the EFS promoter is more resistant to silencing effects compared to the SFFV promoter.<sup>57,58</sup> In addition, it was shown by other studies that virus-derived promoter sequences are more prone to methylation, leading to silencing, compared to human cellular promoters.<sup>59</sup> In our case, EFS integration sites in eGFP<sup>+</sup> cells associate significantly better with markers for active chromatin compared to the bulk cells for the SFFV design, whereas we did not observe this bias for EFS-driven MLV vectors. Importantly, MLV<sup>W390A-CBX</sup> integration sites demonstrated to be enriched near silent regions in the eGFP<sup>+</sup> population, both with EFS and SFFV internal promoters, indicating that integration in heterochromatin still supports effective eGFP expression. Alpharetroviral vectors in contrast, that display a near-random integration profile, are generally silenced and their transcriptionally active sites are rather found in close proximity to active TSS, as shown by single-cell sorting of transcriptionally active (eGFP<sup>+</sup>) cells by Miklik and colleagues.<sup>60,61</sup> Even though this strong silencing could be attributed to LTR silencing (which was used to drive the reporter gene), Hoffmann and colleagues<sup>57</sup> also showed substantial  $\alpha$ RV vector silencing using SIN-configurations in iPSC. Hence, our findings demonstrated the interesting integration profile of the MLV<sup>W390A-CBX</sup> vector in regions enriched in silent chromatin markers, while maintaining promoter activity. Furthermore, this is the first time that the viral vector integration profile was defined in transgene-expressing eGFP<sup>+</sup> cells for different internal promoters. Since we show that significant differences can be retrieved depending on which internal promoter is used, this kind of analysis would be interesting to perform for every

viral vector and cell type applied in clinical trials to decide which promoter is more optimal in a specific disease setting. At least in SupT1 cells, EFS promoter activity was less sensitive to positional effects linked to where integration occurs compared to the SFFV promoter.

As differentiated cells are less suitable for *ex-vivo* cell therapeutic approaches, most therapies target stem cells. This implies that integration takes place in a stem cell nuclear chromatin architecture, which is remodeled following differentiation. Activation and inactivation of different cellular programs will affect the neighborhood and thus the transcriptional potential of the integrated vector. We first wanted to evaluate whether the differentiation ability of stem cells is altered following transduction with our respective vector configurations and moreover whether transgene expression was still supported following differentiation. Next to disorders related to the hematopoietic system, also other diseases, like muscle dystrophies, can benefit from *ex-vivo* gene therapy using retroviral vectors. MABs isolated from adult skeletal muscles display pericyte markers and have the capacity to differentiate into mesodermal lineages including skeletal and smooth muscles, which makes MABs an interesting source to target to repair muscular dystrophies.<sup>41,42</sup> Of note, it was already shown before that a functional artificial skeletal muscle can be generated by implanting MABs on top of residual host muscle.<sup>62</sup> MAB differentiation toward smooth and skeletal muscle can be performed and monitored *in vitro*, and thus provides a good model to assess the effect of vector integration on cell differentiation for the MLV<sup>W390A-CBX</sup> vector configuration. MABs treated with MLV<sup>W390A-CBX</sup> showed the same differentiation ability compared to untransduced MABs when differentiating into smooth muscle cells, regardless of the promoter element used. Curiously, MABs transduced with MLV<sup>W390A-CBX</sup> equipped with the EFS vector design performed better in differentiating toward smooth muscle cells compared to MLV<sup>WT</sup>, illustrating the potential of the MLV<sup>W390A-CBX</sup>-EFS vector. Even more intriguing was the differentiation toward skeletal muscle cells: while MABs transduced with MLV<sup>W390A-CBX</sup> showed the same fusion potential as untransduced control MABs (no significant differences), MABs transduced with MLV<sup>WT</sup> were significantly affected in their capacity to fuse following differentiation to skeletal muscle cells both when using the SFFV or the EFS design. These results reveal that MABs transduced with MLV<sup>W390A-CBX</sup> perform as well as control MABs, in contrast to MABs treated with MLV<sup>WT</sup>, which perform worse. The fact that the myogenic and not the smooth muscle cell differentiation was affected by MLV<sup>WT</sup> transduction, starting from the same transduced MAB pool, might be explained by the different complexity of the differentiation processes toward smooth and skeletal muscle cells, respectively. Myogenic induction on MABs involves a multistep mechanism, where both fusion and differentiation of MABs are required and achieved by largely unknown mechanisms induced by serum starvation. Differentiation toward smooth muscle on the other hand, is less complex and is obtained very efficiently in response to TGF- $\beta$ .<sup>63,64</sup> Hence, integration of MLV<sup>WT</sup> close to promoter and enhancer sequences in MABs may easily dysregulate their myogenic differentiation, while smooth muscle differentiation remains unaltered.<sup>13,14</sup> The more randomly distributed integration profile of the

next-gen MLV<sup>W390A-CBX</sup> viral vector on the other hand didn't disturb MAB differentiation, hinting toward a reduced genotoxicity of MLV<sup>W390A-CBX</sup> compared to MLV<sup>WT</sup>. Since both MLV vector designs supported similar reporter expression levels following differentiation into smooth and skeletal muscle cells, this indicates that the redistribution of vector integration of MLV<sup>W390A-CBX</sup> did not negatively affect reporter gene expression. Together, our results show that MLV<sup>W390A-CBX</sup> allows MABs to differentiate efficiently toward smooth and skeletal muscle cells, underscoring the therapeutic potential of this viral vector.

In conclusion, we engineered clinically relevant MLV<sup>W390A-CBX</sup> vectors with a retargeted integration profile that was overall safer and still supported potent transgene expression. Expression remained stable over time, also after differentiation, leading to an improved differentiation potential of MABs compared to MLV<sup>WT</sup>. Combination of the endogenous EFS promoter to drive transgene expression together with the more desired integration profile of MLV<sup>W390A-CBX</sup> may form an ideal combination to prevent insertional mutagenesis in *ex-vivo* cell therapeutic applications.

## MATERIALS AND METHODS

### Developing plasmid constructs and vector production

MLV<sup>WT</sup>, MLV<sup>W390A</sup>, and MLV<sup>W390A-CBX</sup> packaging constructs were cloned as previously described in pcDNA3.MLV.gp packaging plasmid, a kind gift from Professor Axel Schambach.<sup>33,34,37</sup> pSRS11.SF.eGFP.pre (referred to here as SFFV) and pSRS11.EF-S.eGFP.pre (referred to here as EFS) were used as transfer plasmids, which were kindly provided by Axel Schambach and Michael Rothe (MHH, Hannover, Germany). The integrity of all plasmids was verified by DNA sequencing. MLV-based vectors were produced as previously described by a triple polyethylenimine (PEI, Polysciences) based transfection of HEK293T cells with pVSV-G envelope (pLP-VSV-G #646B, Thermo Fisher Scientific; Brussels; Belgium), pcDNA3.MLV.gp packaging plasmids and the respective transfer plasmids (SFFV and EFS).<sup>37,65</sup> After filtering using a 0.45  $\mu$ M pore-size syringe filter (Corning, Senefte; Belgium), produced vector supernatants were concentrated by tangential flow filtration using a Vivaspin (Vivascience) and stored at  $-80^{\circ}\text{C}$  until use. Subsequently, non-functional RT units (RTUs, non-functional titration) were measured by the SYBRGreen-I product-enhanced RT assay (SG-PERT) and titration was performed on HEK293T cells and analyzed 3 days after transduction via flow cytometry to determine functional transducing titers.<sup>66</sup>

### Cell culture and vector transduction

All cells were grown in a humidified atmosphere containing 5% CO<sub>2</sub> at 37°C. HEK293T cells (293T/17 ATCC CRL-11268 batch 3984732) were cultured in Dulbecco's modified Eagle's medium (DMEM; GIBCO-BRL, Merelbeke, Belgium) supplemented with 2% v/v heat inactivated fetal bovine serum (FBS; GIBCO-BRL) and gentamicin (50  $\mu$ g/mL, GIBCO-BRL). SupT1 cells (provided by the National Institutes of Health reagent program, NIH, Bethesda, MD) and K562 cells were cultured in Roswell Park Memorial Institutes medium (RPMI, GIBCO-BRL) supplemented with 10% heat-inactivated FBS

(GIBCO-BRL) and gentamicin (50 µg/mL, GIBCO-BRL). SupT1 cells and K562 cells were seeded in 96-well plates ( $1 \times 10^5$  cells/well) and transduced with a MOI of 1, 3, and 9 of the respective vectors, after normalization based on RT units. 72 h post-transduction, cells were harvested when 90% confluent and used for eGFP flow cytometry analysis. The remainder of the transduced cells was further cultivated for at least 15 days to eliminate non-integrated DNA and again a flow cytometry analysis was performed. Afterward, 2 million cells were pelleted for further genomic DNA isolation.

Human peripheral blood mononuclear cells (PBMCs) were purified from buffy coats, obtained from the Red Cross blood transfusion center, using density gradient centrifugation (Lymphoprep; Axis-Shield). Primary CD34<sup>+</sup> HSCs were positively selected with anti-CD34-conjugated microbeads according to the manufacturer's instructions (MACS, Mileny Biotec) and stimulated for 48 h in StemSpan SFEMII medium containing CC100 Cytokine Cocktail (STEMCELL Technologies). Prior to cell transduction, CD34<sup>+</sup> HSC were pre-stimulated for 2 days in StemSpan medium enriched with CC100 Cytokine Cocktail. An MOI of 10 of the different vectors was applied by spinoculation (2 h,  $1,200 \times g$ ). Cells were analyzed for eGFP expression by flow cytometry at different time points. 20 days later, cell pellets were generated for further genomic DNA isolation.

#### Flow cytometry analysis and cell sorting

eGFP fluorescence was monitored by flow cytometry analysis after cells were fixed (2% paraformaldehyde [PFA] final) for 15 min at room temperature (RT) using a Guava easyCyte Flow Cytometer system (Miltenyi Biotec). Living cells were selected based on the forward and side scatter channel (FSC-H/SSC-H) and doublets were excluded based on the FSC-A/FSC-H plot. A total of at least 15,000 single living cells were counted. Each sample was measured in triplicates. For sorting experiments, the cells were sorted based on eGFP expression by a S3e Cell Sorter Instrument (Bio-Rad) 20 days after transduction. Data analysis was performed by Guava Suite Software and FlowJo software.

#### Genomic DNA isolation and quantification of integrated copy number via qPCR

Genomic DNA (gDNA) isolation and qPCR were performed as previously described. Briefly, two million cells were pelleted and genomic DNA was extracted using a Mammalian Genomic DNA miniprep kit (Sigma-Aldrich). Samples corresponding to 100 ng genomic DNA were used for analysis. Each reaction contained 12.5 µL IQ Supermix (Bio-Rad), 40 nM forward and reverse WPRE primer, and 40 nM WPRE probe in a final volume of 25 µL. B-actin was quantified as endogenous control. 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.

#### Integration site amplification and analysis of the integration profile

Integration sites were determined as described by Sherman et al.<sup>38</sup> In brief, DNA linkers are ligated to genomic DNA that was randomly

sheared by sonication (Covaris M220 unit). Provirus/host genome junctions were amplified by nested PCR using primers complementary to the linker and primers complementary to the LTR end of the integrated MLV vector. A blocking oligonucleotide was used during both PCR reactions to reduce polymerase extension from the internal vector fragment. Illumina sequencing adapters are attached to the DNA primers of the second PCR. Products were purified by AMPure XP magnetic beads and sequenced using the Illumina Miseq platform. Reads were filtered based on perfect matching of the LTR linker, barcode, and flanking LTR. All sites were mapped to the human reference genome requiring a perfect match within 3 bp of the LTR end. MRCs were computationally generated and matched to experimental sites with respect to the distance to the nearest shearing site. Analysis was performed using the INSPIRED software.<sup>38,39</sup> Specific genomic features were retrieved from the UCSC database. The association with these 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) using Rstudio and the hiAnnotator package from Bioconductor. The Allonco list was derived from the Bushman Lab cancer gene list (<http://www.bushmanlab.org/links/genelists>).

#### MAB cell model

Human primary MAB lines were isolated as previously described from two different donors, referred to in the text as MAB donor 1 and MAB donor 2.<sup>67</sup> Written informed consent was obtained from the subjects who provided their samples for MAB harvest. MAB isolation and characterization was approved by the medical ethics committee of the University Hospital Leuven (n° S5732-388 ML11268). MABs were characterized by means of flow cytometry, mRNA analysis and *in vitro* differentiation assays (data not shown). For maintenance, MABs were cultured in Iscove's modified Dulbecco's medium (IMDM) supplemented with 15% FBS, 1% pen-strep, 1% L-glutamine, 1% sodium pyruvate, 1% non-essential amino acids, 1% insulin-transferrin-selenium, and 0.2% beta-mercaptoethanol (all reagents derived from GIBCO) in 5% CO<sub>2</sub> at 37°C. Two independent experiments were performed by using two different MAB cell lines (MAB donor 1 and MAB donor 2) for both SFFV and EFS vector designs (see cartoon [Figure S5A](#)). MAB donor 1 and MAB donor 2 were transduced with the respective MLV vectors, (normalized based on RTUs) at an MOI of 3. For each MAB cell line (MAB donor 1 and MAB donor 2) myogenic and smooth muscle differentiation was induced 8 days after transduction by incubating the cells with DMEM high glucose, supplemented with 2% of HS, 1% penicillin/streptomycin solution, 2 mM glutamine and 1 mM sodium pyruvate (all reagents from GIBCO) for 12 days (see cartoon [Figure S5](#)). For smooth muscle differentiation, extra  $1,000 \times$  TGF-β was added to the medium. Immunofluorescence staining was performed by 0.2% triton-based permeabilization containing 1% BSA, 5% donkey serum background blocking followed by overnight incubation with primary antibody at 4°C. For myogenic differentiation anti-MyHC (in house polyclonal anti-mouse; 1:20) was used, for smooth muscle differentiation both anti-calponin and anti-alpha SMA (Cy3 conjugated;

Merck) were used. The day after, 1 h incubation with 1:500 Alexa-Fluor-conjugated secondary antibodies and 1:2,000 DAPI staining was performed. Pictures were acquired on Nikon Ti2 automated fluorescent image scanner, pre-specified 4 × 4 fields per well centered on a 24-well format, captured using an automated x-y motorized stage.

### Statistical analysis

Transduction experiments are expressed as means ± standard deviation. Ranked Wald statistics was used to calculate the statistical significance (asterisks) for a given genomic feature between integration site datasets relative to MLV<sup>WT</sup> or MLV<sup>W390A</sup> (dashes). Significant deviation from MLV<sup>WT</sup> or MLV<sup>W390A</sup> for genomic features was calculated using a two-tailed  $\chi^2$  test. To compare samples in the MAB stem cell model, we used Kruskal-Wallis with uncorrected Dunn's test. Confidence intervals were fixed at 99.9% (\* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001). Data are reported as median ± standard deviation (SD).

### Data availability

The raw bioinformatics data required to reproduce these findings are available to download from Mendeley Data – Van Looveren et al. 20201222.<sup>68</sup>

### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.omtm.2021.07.003>.

### ACKNOWLEDGMENTS

We thank Hanne Grosemans for her technical assistance. We thank Jan De Rijck for critical reading of the manuscript. We thank the Institute of Experimental Hematology at Hannover Medical School for sharing the EFS sequence. We thank John Everett for his excellent help in bioinformatic analysis. Figures used for the graphical abstract and Figure S5 were adapted from vector objects derived from <https://www.freepik.com>. We thank and acknowledge <https://www.freepik.com> for the free usage of graphical images (Background vector created by brgfx, <https://www.freepik.com/free-photos-vectors/background>; Love vector created by pch.vector, <https://www.freepik.com/free-photos-vectors/love>). This work was supported by grants from the KU Leuven Research Council (OT/13/098-3M130157), FWO SB (3M140565 SB/151139), FWO Vlaanderen (G0B3516N and G0D4517N) and C1-KUL3DMUSYC (C14/17/111). D.V.L. received funding from the Research Foundation Flanders (FWO-SB). G.G. was a recipient of FWO (G088715N).

### AUTHOR CONTRIBUTIONS

D.V.L. and R.G. designed the experiments. D.V.L. and R.G. wrote the manuscript. D.V.L., G.G., and I.T. performed experiments and collected and analyzed data. D.V.L. performed bioinformatics for integration site analysis. R.G. and M.S. supervised the project.

### DECLARATION OF INTERESTS

The authors declare no competing interests.

### REFERENCES

1. Scott, C.T., and DeFrancesco, L. (2016). Gene therapy's out-of-body experience. *Nat. Biotechnol.* 34, 600–607.
2. Kaufmann, K.B., Büning, H., Galy, A., Schambach, A., and Grez, M. (2013). Gene therapy on the move. *EMBO Mol. Med.* 5, 1642–1661.
3. Cicalese, M.P., and Aiuti, A. (2015). Clinical applications of gene therapy for primary immunodeficiencies. *Hum. Gene Ther.* 26, 210–219.
4. Cavazzana-Calvo, M., Fischer, A., Hacein-Bey-Abina, S., and Aiuti, A. (2012). Gene therapy for primary immunodeficiencies: Part 1. *Curr. Opin. Immunol.* 24, 580–584.
5. Aiuti, A., Bacchetta, R., Seger, R., Villa, A., and Cavazzana-Calvo, M. (2012). Gene therapy for primary immunodeficiencies: Part 2. *Curr. Opin. Immunol.* 24, 585–591.
6. Cavazzana-Calvo, M., Payen, E., Negre, O., Wang, G., Hehir, K., Fusil, F., Down, J., Denaro, M., Brady, T., Westerman, K., et al. (2010). Transfusion independence and HMGA2 activation after gene therapy of human  $\beta$ -thalassaemia. *Nature* 467, 318–322.
7. Maldarelli, F., Wu, X., Su, L., Simonetti, F.R., Shao, W., Hill, S., Spindler, J., Ferris, A.L., Mellors, J.W., Kearney, M.F., et al. (2014). HIV latency. Specific HIV integration sites are linked to clonal expansion and persistence of infected cells. *Science* 345, 179–183.
8. 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.
9. Hacein-Bey-Abina, S., Von Kalle, C., Schmidt, M., McCormack, M.P., Wulffraat, N., Leboulch, P., Lim, A., Osborne, C.S., Pawliuk, R., Morillon, E., et al. (2003). LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science* 302, 415–419.
10. Ott, M.G., Schmidt, M., Schwarzwaelder, K., Stein, S., Siler, U., Koehl, U., Glimm, H., Köhlcke, K., Schilz, A., Kunkel, H., et al. (2006). Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EV11, PRDM16 or SETBP1. *Nat. Med.* 12, 401–409.
11. 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 EV11 activation after gene therapy for chronic granulomatous disease. *Nat. Med.* 16, 198–204.
12. Bushman, F.D. (2020). Retroviral Insertional Mutagenesis in Humans: Evidence for Four Genetic Mechanisms Promoting Expansion of Cell Clones. *Mol. Ther.* 28, 352–356.
13. LaFave, M.C., Varshney, G.K., Gildea, D.E., Wolfsberg, T.G., Baxevanis, A.D., and Burgess, S.M. (2014). MLV integration site selection is driven by strong enhancers and active promoters. *Nucleic Acids Res.* 42, 4257–4269.
14. 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.
15. Wu, X., Li, Y., Crise, B., and Burgess, S.M. (2003). Transcription start regions in the human genome are favored targets for MLV integration. *Science* 300, 1749–1751.
16. 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.
17. Schröder, A.R.W., Shinn, P., Chen, H., Berry, C., Ecker, J.R., and Bushman, F. (2002). HIV-1 integration in the human genome favors active genes and local hotspots. *Cell* 110, 521–529.
18. Persons, D.A., and Baum, C. (2011). Solving the problem of  $\gamma$ -retroviral vectors containing long terminal repeats. *Mol. Ther.* 19, 229–231.
19. Howe, S.J., Mansour, M.R., Schwarzwaelder, K., Bartholomae, C., Hubank, M., Kempinski, 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.
20. Maetzig, T., Galla, M., Baum, C., and Schambach, A. (2011). Gammaretroviral vectors: biology, technology and application. *Viruses* 3, 677–713.

21. Yu, S.F., von Rüden, T., Kantoff, P.W., Garber, C., Seiberg, M., Rütger, U., Anderson, W.F., Wagner, E.F., and Gilboa, E. (1986). Self-inactivating retroviral vectors designed for transfer of whole genes into mammalian cells. *Proc. Natl. Acad. Sci. USA* 83, 3194–3198.
22. Kuo, C.Y., and Kohn, D.B. (2016). Gene Therapy for the Treatment of Primary Immune Deficiencies. *Curr. Allergy Asthma Rep.* 16, 39.
23. Demaison, C., Parsley, K., Brouns, G., Scherr, M., Battmer, K., Kinnon, C., Grez, M., and Thrasher, A.J. (2002). High-level transduction and gene expression in hematopoietic repopulating cells using a human immunodeficiency [correction of immunodeficiency] virus type 1-based lentiviral vector containing an internal spleen focus forming virus promoter. *Hum. Gene Ther.* 13, 803–813.
24. Yam, P.Y., Li, S., Wu, J., Hu, J., Zaia, J.A., and Yee, J.-K. (2002). Design of HIV vectors for efficient gene delivery into human hematopoietic cells. *Mol. Ther.* 5, 479–484.
25. 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.
26. Bosticardo, M., Ghosh, A., Du, Y., Jenkins, N.A., Copeland, N.G., and Candotti, F. (2009). Self-inactivating retroviral vector-mediated gene transfer induces oncogene activation and immortalization of primary murine bone marrow cells. *Mol. Ther.* 17, 1910–1918.
27. 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.
28. 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.
29. 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  $\beta$ -globin locus control region in combination with the EF1 $\alpha$  short promoter allows enhanced lentiviral vector-mediated erythroid gene expression with conserved multilineage activity. *Mol. Ther.* 20, 1400–1409.
30. 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.
31. Sharma, A., Larue, R.C., Plumb, M.R., Malani, N., Male, F., Slaughter, A., Kessel, 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.
32. 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.
33. El Ashkar, S., De Rijck, J., Demeulemeester, J., Vets, S., Madlala, P., Cermakova, K., Debyser, Z., and Gijssbers, R. (2014). BET-independent MLV-based Vectors Target Away From Promoters and Regulatory Elements. *Mol. Ther. Nucleic Acids* 3, e179.
34. El Ashkar, S., Van Looveren, D., Schenk, F., Vranckx, L.S., Demeulemeester, J., De Rijck, J., Debyser, Z., Modlich, U., and Gijssbers, R. (2017). Engineering Next-Generation BET-Independent MLV Vectors for Safer Gene Therapy. *Mol. Ther. Nucleic Acids* 7, 231–245.
35. Zhou, S., Fatima, S., Ma, Z., Wang, Y.-D., Lu, T., Janke, L.J., Du, Y., and Sorrentino, B.P. (2016). Evaluating the Safety of Retroviral Vectors Based on Insertional Oncogene Activation and Blocked Differentiation in Cultured Thymocytes. *Mol. Ther.* 24, 1090–1099.
36. Montini, E., Cesana, D., Schmidt, M., Sanvito, F., Bartholomae, C.C., Ranzani, M., Benedicenti, F., Sergi, L.S., Ambrosi, A., Ponzoni, M., et al. (2009). The genotoxic potential of retroviral vectors is strongly modulated by vector design and integration site selection in a mouse model of HSC gene therapy. *J. Clin. Invest.* 119, 964–975.
37. 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.
38. Sherman, E., Nobles, C., Berry, C.C., Six, E., Wu, Y., Dryga, A., Malani, N., Male, F., Reddy, S., Bailey, A., et al. (2016). INSPIRED: A Pipeline for Quantitative Analysis of Sites of New DNA Integration in Cellular Genomes. *Mol. Ther. Methods Clin. Dev.* 4, 39–49.
39. Berry, C.C., Nobles, C., Six, E., Wu, Y., Malani, N., Sherman, E., Dryga, A., Everett, J.K., Male, F., Bailey, A., et al. (2016). INSPIRED: Quantification and Visualization Tools for Analyzing Integration Site Distributions. *Mol. Ther. Methods Clin. Dev.* 4, 17–26.
40. 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.
41. Minasi, M.G., Riminucci, M., De Angelis, L., Borello, U., Berarducci, B., Innocenzi, A., Caprioli, A., Sirabella, D., Baiocchi, M., De Maria, R., et al. (2002). The meso-angioblast: a multipotent, self-renewing cell that originates from the dorsal aorta and differentiates into most mesodermal tissues. *Development* 129, 2773–2783.
42. Dellavalle, A., Sampaolesi, M., Tonlorenzi, R., Tagliafico, E., Sacchetti, B., Perani, L., Innocenzi, A., Galvez, B.G., Messina, G., Morosetti, R., et al. (2007). Pericytes of human skeletal muscle are myogenic precursors distinct from satellite cells. *Nat. Cell Biol.* 9, 255–267.
43. South, E., Cox, E., Meader, N., Woolcott, N., and Griffin, S. (2019). Strimvelis® for Treating Severe Combined Immunodeficiency Caused by Adenosine Deaminase Deficiency: An Evidence Review Group Perspective of a NICE Highly Specialised Technology Evaluation. *Pharmacoeconom. Open* 3, 151–161.
44. 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.
45. Rothe, M., Schambach, A., and Biasco, L. (2014). Safety of gene therapy: new insights to a puzzling case. *Curr. Gene Ther.* 14, 429–436.
46. Browning, D.L., and Trobridge, G.D. (2016). Insulators to Improve the Safety of Retroviral Vectors for HIV Gene Therapy. *Biomedicines* 4, 4.
47. Emery, D.W., Yannaki, E., Tubb, J., and Stamatoyannopoulos, G. (2000). A chromatin insulator protects retrovirus vectors from chromosomal position effects. *Proc. Natl. Acad. Sci. USA* 97, 9150–9155.
48. 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.
49. 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.
50. Moiani, A., Suerth, J.D., Gandolfi, F., Rizzi, E., Severgnini, M., De Bellis, G., Schambach, A., and Mavilio, F. (2014). Genome-wide analysis of alpharetroviral integration in human hematopoietic stem/progenitor cells. *Genes (Basel)* 5, 415–429.
51. 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.
52. 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.
53. Hou, X., Du, Y., Deng, Y., Wu, J., and Cao, G. (2015). Sleeping Beauty transposon system for genetic etiological research and gene therapy of cancers. *Cancer Biol. Ther.* 16, 8–16.
54. Kebriaei, P., Izsvák, Z., Narayanavari, S.A., Singh, H., and Ivics, Z. (2017). Gene Therapy with the Sleeping Beauty Transposon System. *Trends Genet.* 33, 852–870.
55. Hudecek, M., Izsvák, Z., Johnen, S., Renner, M., Thumann, G., and Ivics, Z. (2017). Going non-viral: the Sleeping Beauty transposon system breaks on through to the clinical side. *Crit. Rev. Biochem. Mol. Biol.* 52, 355–380.

56. 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.
57. Hoffmann, D., Schott, J.W., Geis, F.K., Lange, L., Müller, F.-J., Lenz, D., Zychlinski, D., Steinemann, D., Morgan, M., Moritz, T., and Schambach, A. (2017). Detailed comparison of retroviral vectors and promoter configurations for stable and high transgene expression in human induced pluripotent stem cells. *Gene Ther.* 24, 298–307.
58. 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.
59. Herbst, F., Ball, C.R., Tuorto, F., Nowrouzi, A., Wang, W., Zavidij, O., Dieter, S.M., Fessler, S., van der Hoeven, F., Kloz, U., et al. (2012). Extensive methylation of promoter sequences silences lentiviral transgene expression during stem cell differentiation in vivo. *Mol. Ther.* 20, 1014–1021.
60. Miklík, D., Šenigl, F., and Hejnar, J. (2018). Proviruses with Long-Term Stable Expression Accumulate in Transcriptionally Active Chromatin Close to the Gene Regulatory Elements: Comparison of ASLV-, HIV- and MLV-Derived Vectors. *Viruses* 10, 116.
61. Šenigl, F., Miklík, D., Auxt, M., and Hejnar, J. (2017). Accumulation of long-term transcriptionally active integrated retroviral vectors in active promoters and enhancers. *Nucleic Acids Res.* 45, 12752–12765.
62. Cossu, G., Previtali, S.C., Napolitano, S., Cicalese, M.P., Tedesco, F.S., Nicastro, F., Noviello, M., Roostalu, U., Natali Sora, M.G., Scarlato, M., et al. (2015). Intra-arterial transplantation of HLA-matched donor mesoangioblasts in Duchenne muscular dystrophy. *EMBO Mol. Med.* 7, 1513–1528.
63. Tagliafico, E., Brunelli, S., Bergamaschi, A., De Angelis, L., Scardigli, R., Galli, D., Battini, R., Bianco, P., Ferrari, S., Cossu, G., et al. (2004). TGFβ/BMP activate the smooth muscle/bone differentiation programs in mesoangioblasts. *J. Cell Sci.* 117, 4377–4388.
64. Roobrouck, V.D., Clavel, C., Jacobs, S.A., Ulloa-Montoya, F., Crippa, S., Sohni, A., Roberts, S.J., Luyten, F.P., Van Gool, S.W., Sampaolesi, M., et al. (2011). Differentiation potential of human postnatal mesenchymal stem cells, mesoangioblasts, and multipotent adult progenitor cells reflected in their transcriptome and partially influenced by the culture conditions. *Stem Cells* 29, 871–882.
65. Ibrahim, 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.
66. 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.
67. Quattrocchi, M., Swinnen, M., Giacomazzi, G., Camps, J., Barthélemy, I., Ceccarelli, G., Caluwé, E., Grosemans, H., Thorrez, L., Pelizzo, G., et al. (2015). Mesodermal iPSC-derived progenitor cells functionally regenerate cardiac and skeletal muscle. *J. Clin. Invest.* 125, 4463–4482.
68. Dominique V.L. (2020), “VanLooveren et al. 20201222”, Mendeley Data, V1. <https://doi.org/10.17632/wx25yk73n7.1>.