




# BET-Independent Murine Leukemia Virus Integration Is Retargeted *In Vivo* and Selects Distinct Genomic Elements for Lymphomagenesis

Ivan Nombela,<sup>a</sup> Martine Michiels,<sup>a</sup> Dominique Van Looveren,<sup>a</sup> Lukas Marcelis,<sup>b,c</sup> Sara el Ashkar,<sup>a</sup> Siska Van Belle,<sup>a</sup>  
 Anne Bruggemans,<sup>a</sup> Thomas Tousseyn,<sup>b,c</sup> Jürg Schwaller,<sup>d</sup> Frauke Christ,<sup>a</sup> Rik Gijssbers,<sup>a</sup> Jan De Rijck,<sup>a</sup>  Zeger Debyser<sup>a</sup>

<sup>a</sup>Laboratory for Molecular Virology and Gene Therapy, Department of Pharmaceutical and Pharmacological Sciences, KU Leuven, Leuven, Flanders, Belgium

<sup>b</sup>Department of Imaging and Pathology, Translational Cell and Tissue Research Lab, KU Leuven, Flanders, Belgium

<sup>c</sup>UZ Leuven, University Hospitals, Department of Pathology, KU Leuven, Flanders, Belgium

<sup>d</sup>Department of Biomedicine, University Children's Hospital (UKBB), University of Basel, Basel, Switzerland

**ABSTRACT** Moloney murine leukemia virus (MLV) infects BALB/c mice and induces T-cell lymphoma in mice. Retroviral integration is mediated by the interaction of the MLV integrase (IN) with members of the bromodomain and extraterminal motif (BET) protein family (BRD2, BRD3, and BRD4). The introduction of the W390A mutation into MLV IN abolishes the BET interaction. Here, we compared the replication of W390A MLV to that of wild-type (WT) MLV in adult BALB/c mice to study the role of BET proteins in replication, integration, and tumorigenesis *in vivo*. Comparing WT and W390A MLV infections revealed similar viral loads in the blood, thymus, and spleen cells. Interestingly, W390A MLV integration was retargeted away from GC-enriched genomic regions. However, both WT MLV- and W390A MLV-infected mice developed T-cell lymphoma after similar latencies represented by an enlarged thymus and spleen and multiorgan tumor infiltration. Integration site sequencing from splenic tumor cells revealed clonal expansion in all WT MLV- and W390A MLV-infected mice. However, the integration profiles of W390A MLV and WT MLV differed significantly. Integrations were enriched in enhancers and promoters, but compared to the WT, W390A MLV integrated less frequently into enhancers and more frequently into oncogene bodies such as *Notch1* and *Ppp1r16b*. We conclude that host factors direct MLV *in vivo* integration site selection. Although BET proteins target WT MLV integration preferentially toward enhancers and promoters, insertional lymphomagenesis can occur independently from BET, likely due to the intrinsically strong enhancer/promoter of the MLV long terminal repeat (LTR).

**IMPORTANCE** In this study, we have shown that the *in vivo* replication of murine leukemia virus happens independently of BET proteins, which are key host determinants involved in retroviral integration site selection. This finding opens a new research line in the discovery of alternative viral or host factors that may complement the dominant host factor. In addition, our results show that BET-independent murine leukemia virus uncouples insertional mutagenesis from gene enhancers, although lymphomagenesis still occurs despite the lack of an interaction with BET proteins. Our findings also have implications for the engineering of BET-independent MLV-based vectors for gene therapy, which may not be a safe alternative.

**KEYWORDS** BALB/c mouse, BET proteins, BRD4, integrase, integration site selection, lymphomagenesis, murine leukemia virus, retroviral integration

The integration of the retroviral genome into the host genome is a key step in the replication cycle of retroviruses. Several studies have pointed out that integration site selection of retroviruses is targeted and not a random event (1). This event is specific to distinct

**Editor** Peter Pelka, University of Manitoba

**Copyright** © 2022 Nombela et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Zeger Debyser, zeger.debyser@kuleuven.be.

The authors declare no conflict of interest.

**Received** 13 June 2022

**Accepted** 4 July 2022

**Published** 19 July 2022

retroviral families (2). For instance, human immunodeficiency virus type 1 (HIV-1) integration is driven toward active transcriptional units, while murine leukemia virus (MLV) integration is directed toward active enhancers and promoters (3–6). The main viral determinant coordinating integration is the integrase (IN). The biochemical reactions of retroviral integration are well understood. After reverse transcription, IN cleaves specific phosphodiester bonds near the viral DNA ends during the 3'-end processing reaction. Next, IN uses the resulting viral DNA 3'-OH groups during strand transfer to cut target DNA. Simultaneously, IN joins each viral DNA end to target DNA 5'-end phosphates. Both reactions proceed via the direct transesterification of phosphodiester bonds (7).

Lens epithelium-derived growth factor (LEDGF)/p75 is a direct binding partner of HIV-1 IN and specific for lentivirus integration (8). It tethers the viral preintegration complex (PIC) to the chromatin (9–12). The tethering mechanism of LEDGF/p75 requires the direct interaction of the integrase binding domain (IBD) at the C terminus with HIV-1 IN and the recognition of the H3K36me2/3 mark on nucleosomes through the N-terminal PWWP domain (13). Small-molecule inhibitors of the interaction between HIV-1 integrase and LEDGF/p75 have been developed and termed LEDGINs (14). The addition of LEDGINs during HIV infection not only reduces integration but also targets residual provirus away from H3K36me2/3 (13, 15). The residual provirus, targeted away from its preferential chromatin landscape, is transcriptionally less active, even after reactivation. It appears that retroviruses evolved to adopt epigenetic readers to find preferential integration sites associated with active transcription and/or latency (16).

In the case of MLV, proteins of the bromodomain and extraterminal domain (BET) family, which include BRD2, BRD3, and BRD4, were identified as binding partners of MLV IN and proposed to tether the MLV PIC to transcription start sites (TSSs) and enhancers (17–20). BET proteins are composed of two conserved N-terminal bromodomains (bromodomain 1 [BD1] and BD2), an extraterminal (ET) domain, and a C-terminal domain (CTD), which is present only in BRD4 (21). BD1 and BD2 are regions with hydrophobic amino acids able to recognize acetylated H3 and H4 tails (22) and act as chromatin readers. The ET domain associates with a variety of viral and cellular proteins, including transcription factors, chromatin-modifying factors, and histone-modifying enzymes (23). The CTD is necessary for the recruitment of positive transcription elongation factor b (P-TEFb) by the transcriptional complex (24). The MLV IN-BET interaction is mediated through the ET domain of BET proteins and the C-terminal domain of MLV IN (25) and results in the tethering of MLV IN to nucleosomes (26). Amino acids located between positions 390 and 405 of the MLV IN sequence define a conserved domain in gammaretroviruses involved in the interaction with BRD4 (25). A single substitution of tryptophan for alanine at position 390 (W390A) in this domain is sufficient to prevent the interaction with BRD2, BRD3, and BRD4 (19). BET proteins promote the efficient integration of MLV into the host genome (20). However, BET-independent MLV replicates with viral titers similar to those of wild-type (WT) MLV in cell culture (27), although its integration profile differs from that of the WT toward a more random distribution away from transcriptionally active genomic regions and transcriptional start sites (28). Therefore, the extent to which BET proteins are key factors for the replication and integration of MLV is still a matter of debate. In addition to this, retroviral p12 may also play a role in tethering MLV PICs to mitotic chromosomes to facilitate integration (29).

MLV has been studied as one of the prototype oncogenic animal retroviruses since the 1950s when it was noticed that leukemia could be transmitted to newborn mice by an unknown agent (30). This horizontal transmission occurs primarily through milk ingestion, while transmission by the venereal route is less common (31). MLV induces either lymphoblastic leukemia (31) or lymphoma in mice, and an enlarged thymus, spleen, or lymph node is present in infected adults (32). Since MLV can infect several B or T cells, either leukemia or lymphoma of these cell lineages or their corresponding immature lineages can be induced (33).

Upon integration, MLV can modify mouse gene expression by insertional mutagenesis (34). This phenomenon is characterized by the deregulation of gene expression at the transcriptional or posttranscriptional level (35, 36) and requires viral promoter and

enhancer elements located in the U3 region of the long terminal repeats (LTRs) of the MLV genome. Therefore, provirus insertion at the start of the gene may enhance the transcription of the target gene from the retroviral promoter. Alternatively, viral integration into the central part or close to the 3' end of a gene may produce a truncated form of the native protein, which can abort its regulatory control. Screening for MLV insertional mutagenesis in mouse models has been instrumental in identifying human oncogenes (37–39). Genes identified as targets for retroviral integration that also act as oncogenes are *Notch1*, *Pim1*, *Myc*, *Myb*, *Gfi*, and *Pvt1*, among others (40).

Many recent studies on the integration site selection of MLV were performed in the context of MLV-based vectors for human gene therapy (41–43). MLV-based vectors have been successfully used in clinical trials for patients with adenosine deaminase deficiency (44). However, serious safety concerns were raised when patients with X-linked severe combined immunodeficiency disease (X-SCID) developed T-cell leukemia after treatment with MLV-based viral vectors by the activation of the LMO2 oncogene (45–47). Several generations of viral vectors have been designed in order to increase the biosafety of retroviral therapy. A “second” generation is characterized by a self-inactivating (SIN) vector ensuring a deletion in the 3' LTR after reverse transcription that abolishes the LTR promoter activity (48). Most recently, a “third” generation of retroviral vectors was proposed. This generation is characterized by a mutation in MLV IN abolishing the interaction with BET proteins to achieve an integration profile targeted away from the TSSs of oncogenes under *in vitro* conditions (27, 28, 49). Still, our knowledge of the impact of integration site selection on oncogenesis *in vivo* is poor.

In fact, the role of retroviral host factors such as LEDGF/p75 and BET in retroviral pathogenesis in their natural host has not been unambiguously demonstrated. HIV replication in humans in the absence of LEDGF/p75 has not been investigated. Upcoming clinical trials with LEDGINs as HIV treatments or cures may provide this information in the future (50). Loyola et al. studied the pathogenesis of a C-terminally deleted MLV clone in a tumorigenesis model (51). That previous study was associated with some recombination events with endogenous retroviruses.

Here, we explored the role of the BET-MLV IN interaction in retroviral replication, integration site selection, and lymphomagenesis *in vivo*. In contrast to a C-terminal truncation, the site-specific W390A mutant replicated to the same levels as those of wild-type MLV. Mice infected with W390A MLV developed T-cell lymphoma, to an extent similar to that in mice infected with WT MLV. The W390A mutation targeted integration away from enhancers but increased integration into the bodies of known oncogenes. Our observations indicate that the loss of the BET-MLV IN interaction redirects viral integration sites but does not abrogate MLV-mediated lymphomagenesis. In addition, our work also clarifies the relative contribution of viral integration at enhancers versus promoters or gene bodies to MLV-induced insertional mutagenesis.

## RESULTS

**Efficient BET-independent *in vitro* replication of the W390A MLV mutant.** To investigate the role of BET proteins during MLV replication *in vivo*, we generated two BET-independent Moloney MLV molecular clones: W390A and a C-terminal truncation mutant. Previous studies revealed that IN W390A is essential for the interaction of IN with BET proteins and, as such, affects the integration site pattern of MLV (27, 28). In addition, it was shown previously that this mutation does not hamper the transduction efficiency of MLV-derived viral particles (46). However, its effect on multiple-round viral replication was not assessed previously. First, a codon-optimized sequence duplicating the overlapping sequence at the protein level was inserted in front of the *env* start codon in molecular clone p63.2. We created WT p63.2 to introduce the W390A mutation a single time into MLV IN. Otherwise, the introduction of this mutation into parental p63.2 would have caused selective pressure to revert this mutation due to the presence of the wild-type codon in the overlapping *Env* gene (see Fig. S1A in the supplemental material). We also

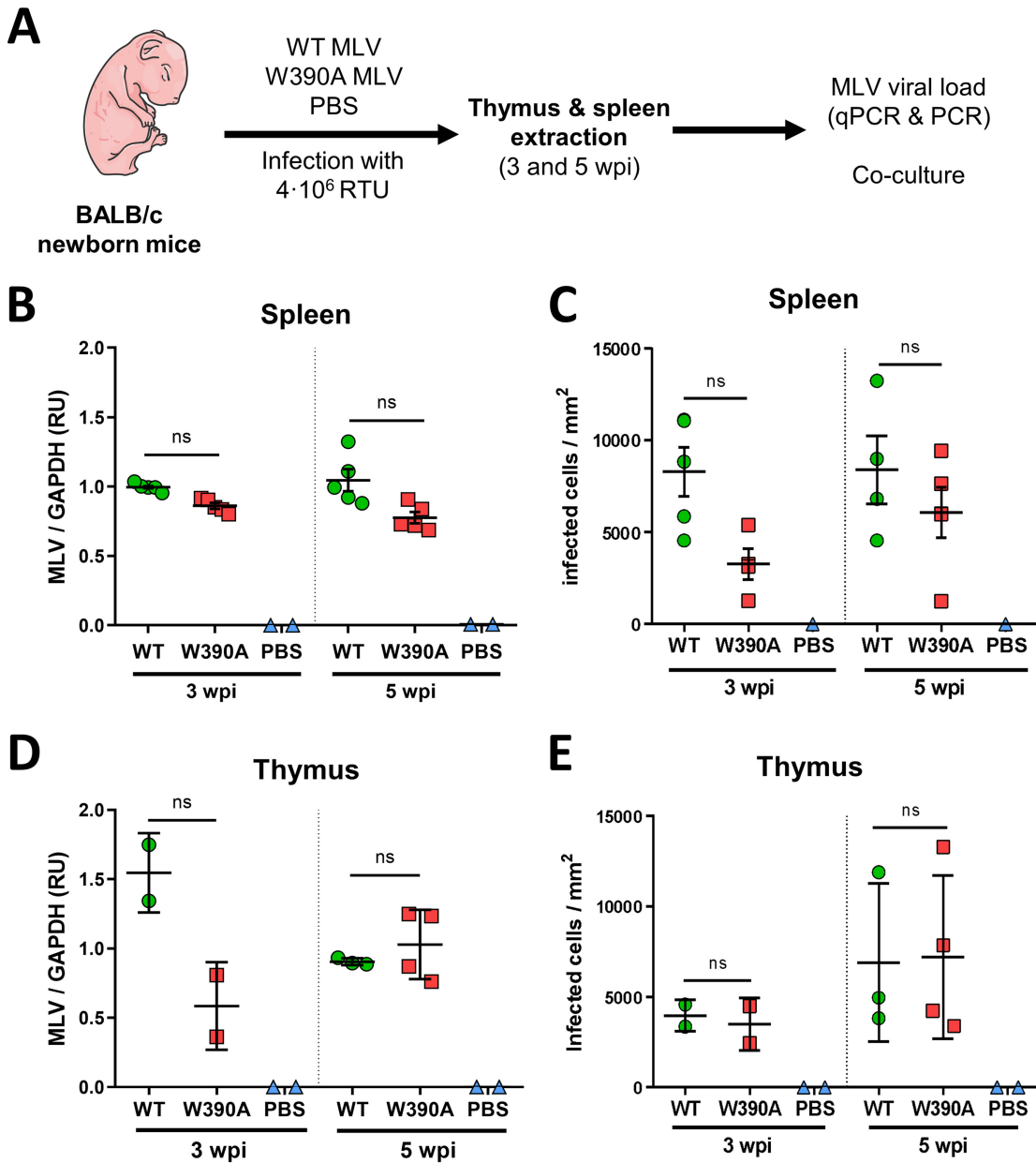
generated  $\Delta$ C p63.2, carrying a deletion in the C-terminal domain of MLV IN between positions 382 and 408 (Fig. S2) (52).

HEK 293T cells were transfected with the plasmids p63.2, WT p63.2, W390A p63.2, and  $\Delta$ C p63.2 to compare the levels of production of viral particles (Fig. S1B). The reverse transcriptase (RT) activity in the supernatant was measured at 48 h posttransfection. Both WT p63.2 and W390A p63.2 were at least as efficient as the parental p63.2 viral clone in the production of viral particles (Fig. S1C). However, the truncation of the C-terminal tail (p63.2  $\Delta$ C) resulted in a 3-fold decrease in RT activity in comparison with that of the parental p63.2 clone (Fig. S1C). Next, NIH 3T3 cells were infected with equal RT units (RTU) of MLV 63.2, WT 63.2, W390A MLV, and  $\Delta$ C MLV molecular clones (Fig. S1B). In contrast to the other variants, the  $\Delta$ C p63.2 clone did not efficiently replicate in NIH 3T3 cells (Fig. S1D). Since the C-terminal deletion hampers viral production and replication, this viral clone was omitted from further experiments.

**W390A MLV replicates at wild-type levels in BALB/c mice.** Considering that WT MLV and W390A MLV can replicate to the same extent *in vitro*, we evaluated their infectivity *in vivo* next. Newborn BALB/c mice were injected with  $4 \times 10^6$  RTU of WT or W390A MLV or 50  $\mu$ L of phosphate-buffered saline (PBS), as indicated in Fig. 1A. As a measure of MLV replication, MLV IN RNA expression was analyzed in spleen and thymus samples by reverse transcription-quantitative PCR (RT-qPCR). MLV IN RNA levels in spleen and thymus samples were similar after infection with WT MLV and W390A MLV, although a trend toward a lower viral load was seen with W390A MLV in the thymus at 3 weeks postinfection (wpi) (Fig. 1B and D). Additionally, *in vivo* infectivity was assessed by coculture of NIH 3T3 cells with spleen or thymus cells extracted at 3 or 5 wpi from mice infected with WT or W390A MLV. No statistically significant difference was found in the numbers of infected cells in coculture with the spleen or thymus of either WT MLV- or W390A MLV-infected mice, although a trend toward lower infection of spleen cells infected with W390A MLV was seen at 3 wpi (Fig. 1C and E). To exclude revertants of W390A MLV, we sequenced part of the MLV IN gene in 6 mice infected with parental 63.2 MLV, WT MLV, or W390A MLV at 12 weeks postinfection (Fig. S3). Although some synonymous and nonsynonymous base changes were detected, the alanine codon GCA remained intact.

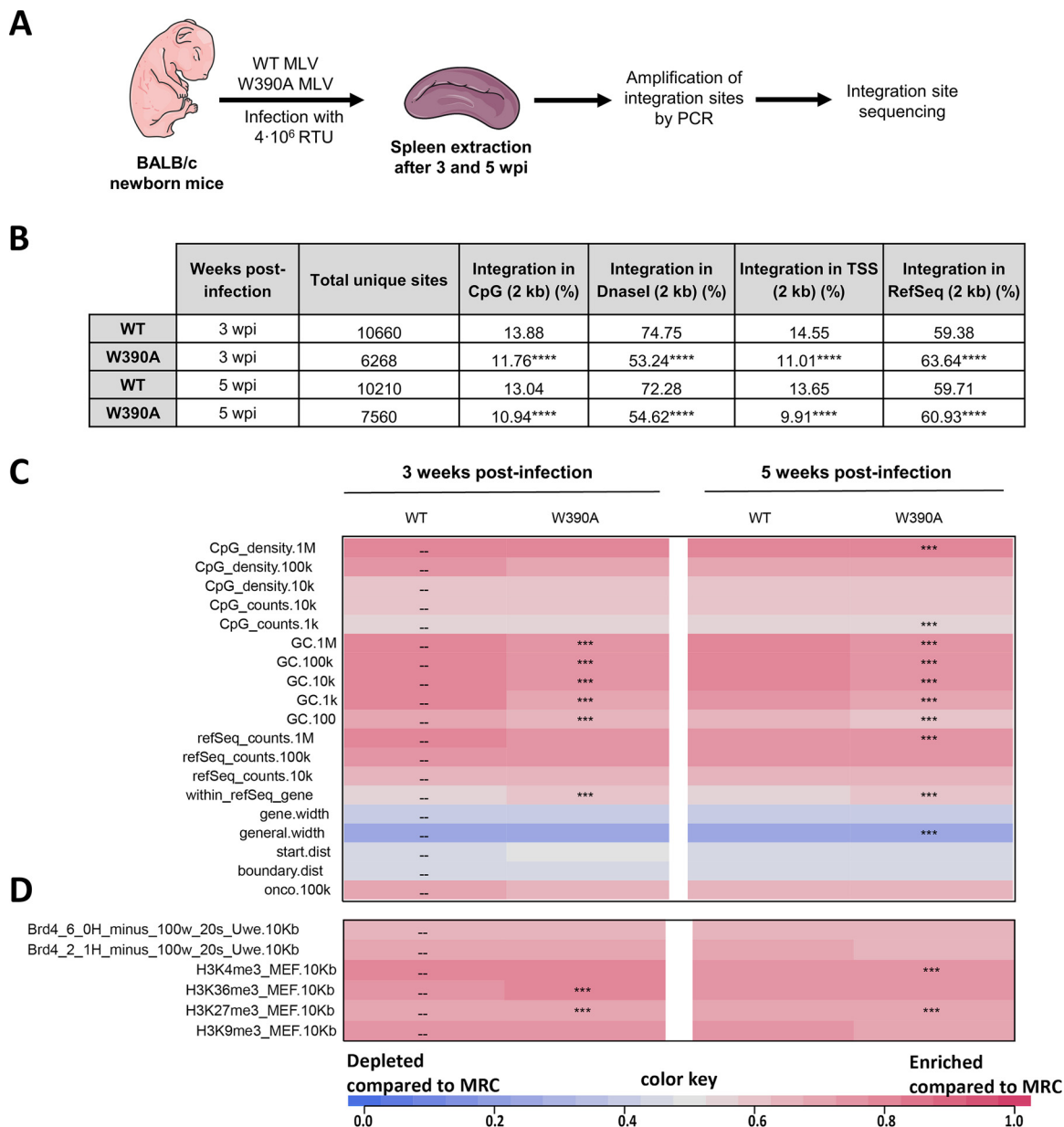
**W390A MLV retargets viral integration.** Next, we analyzed integration preferences relative to a set of genomic and epigenetic features. W390A IN is known to shift the integration site preference in cell culture (27, 28). In fact, previous *in vitro* results showed a BET-independent integration profile characterized by lower levels of integration into transcription start sites (TSSs) of RefSeq genes, CpG islands, and DNase I-hypersensitive sites (DHSs) (28). Integration sites in spleen samples of mice infected with WT or W390A MLV were analyzed at 3 and 5 weeks postinfection (Fig. 2A). Fig. 2B summarizes the distribution of integration sites for several genomic features in a 2-kb window. A total of 6,000 to 10,000 integration sites were obtained under each condition by pooling genomic DNA from spleen cells of three mice infected with either WT or W390A MLV. The presence of W390A significantly shifted integration away from regions enriched in CpG islands, DNase I-hypersensitive sites, GC-enriched regions, and TSSs (Fig. 2C and D), and integration occurred more frequently in positions located within RefSeq genes at both tested times postinfection (Fig. 2B). With respect to epigenetic features, W390A MLV integration shifted toward histone marks associated with transcriptionally active and open chromatin, such as H3K36me3 (Fig. 2D), often related with gene bodies (53). In general, these data demonstrate that the *in vivo* integration site preference of BET-independent MLV in mouse spleen cells differs from that of WT MLV, although the extent of retargeting of the virus away from CpG and TSSs (but not from DHSs) was less than that observed with retroviral vectors in culture (27, 28); BET-independent MLV was retargeted away from all these genomic features, pointing to the role of BET in determining MLV integration sites *in vivo*. Moreover, MLV replication *in vivo* may negatively select for some integration sites due to cytopathogenicity, in contrast to single-round MLV vectors.

**Both WT MLV and W390A MLV induce lymphoma in infected mice.** Virus-induced hematological malignancies are characterized by increased white blood cell (WBC) proliferation (54) and tumor infiltration into several organs and tissues, including the lymph nodes,



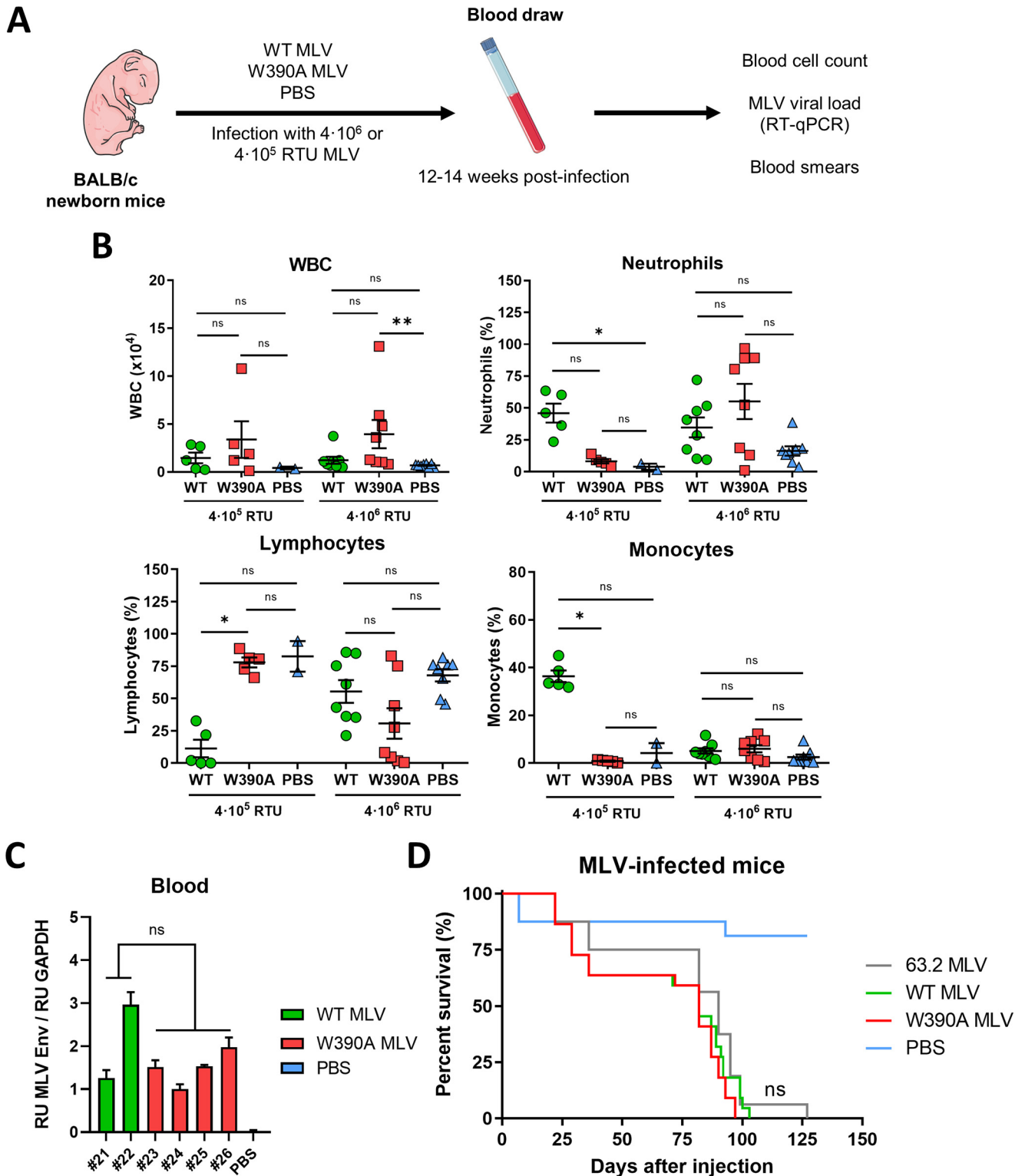
**FIG 1** *In vivo* replication of WT MLV and W390A MLV in murine tissue. (A) Schematic representation of the workflow to evaluate the *in vivo* infectivity of WT MLV and W390A MLV. Newborn mice were infected by intraperitoneal injection with  $4 \times 10^5$  RTU of WT MLV or W390A MLV 1 day after birth. Blood was drawn between 80 and 90 days after injection. Extracts of the spleen and thymus were made at 3 and 5 weeks postinfection. (B) MLV loads in spleen cells from mice infected with WT MLV ( $n = 5$ ) or W390A MLV ( $n = 5$ ) or injected with PBS ( $n = 2$ ) at 3 and 5 weeks postinfection. The viral load was measured by RT-qPCR of MLV IN relative to GAPDH levels and represented as relative units (RU). (C) Number of infected NIH 3T3 cells after 1 day of coculture with cells from spleens of mice infected with WT MLV ( $n = 4$ ) or W390A MLV ( $n = 4$ ) or injected with PBS ( $n = 2$ ) at 3 and 5 weeks postinfection. (D) MLV loads in thymus cells from mice infected with WT MLV (3 wpi,  $n = 2$ ; 5 wpi,  $n = 3$ ) or W390A MLV (3 wpi,  $n = 2$ ; 5 wpi,  $n = 4$ ) or injected with PBS ( $n = 2$ ) at 3 and 5 weeks postinfection. The viral load was measured by RT-qPCR of MLV IN relative to GAPDH levels. (E) Number of infected NIH 3T3 cells after 1 day of coculture with cells from spleens of mice infected with WT MLV (3 wpi,  $n = 2$ ; 5 wpi,  $n = 3$ ) or W390A MLV (3 wpi,  $n = 2$ ; 5 wpi,  $n = 4$ ) or injected with PBS ( $n = 2$ ) at 3 and 5 weeks postinfection. No statistically significant difference was found using a Kruskal-Wallis test (B to E). ns, not significant.

spleen, and thymus (55). For this reason, we evaluated the pathology in blood, spleen, and thymus samples from mice infected with either WT or W390A MLV at 12 to 14 weeks postinfection. First, the proliferation of WBCs was assessed as described in the legend of Fig. 3A. Mice injected with  $4 \times 10^5$  RTU MLV (low dose) displayed an increased WBC count ( $1.46 \times 10^4 \pm 1.24 \times 10^4$  WBCs for WT MLV and  $3.38 \times 10^4 \pm 4.26 \times 10^4$  WBCs for W390A MLV) compared to those in mice injected with PBS ( $0.42 \times 10^4 \pm 0.12 \times 10^4$  WBCs),



**FIG 2** Integration site analysis of WT MLV and W390A MLV in infected mice. (A) Schematic representation of the workflow to determine the integration sites and the epigenetic features of WT MLV and W390A MLV integrants *in vivo*. (B) Genomic distribution of MLV integration sites obtained from spleen cells of mice infected with either WT MLV or W390A MLV at 3 and 5 wpi. Integration percentages for 2-kb windows around CpG-rich island midpoints, DNase I-hypersensitive sites (DHSs), transcription start sites (TSSs), and RefSeq genes are given. Statistical analysis was done using a chi-squared test comparing integrations from WT MLV-infected mice with integrations from W390A MLV-infected mice ( $P \leq 0.0001$ ). (C and D) Heatmaps depicting the relationship between integration site frequency and different genomic (C) or epigenetic (D) features within a 10-kb interval in splenic cells from mice infected with either WT MLV or W390A MLV. The data shown are based on pooled samples from 3 mice infected with either WT MLV or W390A MLV. Features analyzed are shown to the left of the corresponding row of the heatmap. Colors indicate whether a particular feature is disfavored (blue) or favored (red) for the integration of the respective data sets relative to their computer-generated matched random controls (MRCs), as detailed in the color scale at the bottom. Significant departures from WT MLV integration sites in splenic cells are based on a Wald test, referred to as the  $\chi^2$  distribution.

although this difference was not statistically significant (Fig. 3B). Similarly, no statistically significant difference was found between mice infected with WT MLV and those infected with W390A MLV. Mice infected with  $4 \times 10^6$  RTU MLV (high dose) displayed an increased WBC count ( $1.22 \times 10^4 \pm 1.04 \times 10^4$  WBCs for WT MLV and  $3.94 \times 10^4 \pm 4.18 \times 10^4$  WBCs for W390A MLV). Here, the increase in WBCs was statistically significant in W390A MLV-infected mice ( $P = 0.0044$ ). After low-dose infection, WT MLV induced an increase in the relative number of neutrophils and monocytes and a decrease in the number of



**FIG 3** Blood cell counts in WT MLV- and W390A MLV-infected mice. (A) Schematic representation of the workflow to evaluate the long-term effect of WT MLV and W390A MLV replication on blood cell numbers. Newborn mice were infected by intraperitoneal injection with  $4 \times 10^5$  RTU or  $4 \times 10^6$  RTU of either WT MLV or W390A MLV 1 day after birth. Blood was drawn between 80 and 90 days after injection. PBS was used as a negative control. (B) Number of WBCs in blood samples of mice infected with WT MLV ( $4 \times 10^5$  RTU,  $n = 5$ ;  $4 \times 10^6$  RTU,  $n = 8$ ) or W390A MLV ( $4 \times 10^5$  RTU,  $n = 5$ ;  $4 \times 10^6$  RTU,  $n = 8$ ) or injected with PBS ( $n = 2$  and  $8$  for comparisons with  $4 \times 10^5$  RTU and  $4 \times 10^6$  RTU, respectively). Data correspond to the number of WBCs in  $1 \mu\text{L}$  of blood. Percentages of neutrophils, lymphocytes, and monocytes in WBC samples are given for mice infected with WT MLV or W390A MLV. PBS injection was used as a negative control. Horizontal lines indicate means  $\pm$  standard deviations (SD), while dots represent individual values. Statistical significance was analyzed using a Kruskal-Wallis test. (C) MLV loads measured by RT-quantitative PCR of the MLV Env gene normalized to (Continued on next page)

lymphocytes, which was not observed with W390A MLV. This phenotype was not seen at the high dose (Fig. 3B). At the high dose, the percentages of monocytes, lymphocytes, and neutrophils were similar among mice infected with WT MLV and W390A MLV and those injected with PBS (Fig. 3B). A summary of the blood cell counts and statistics can be found in Table S2.

The viral load in whole blood of mice infected with WT MLV ( $n = 2$ ) or W390A MLV ( $n = 4$ ) was measured by RT-qPCR to detect Env gene mRNA at 21 days postinfection. The statistical analysis did not reveal any significant difference (Fig. 3C).

We next compared the rates of survival of mice infected with  $4 \times 10^6$  RTU WT MLV and W390A MLV. In this experiment, newborn mice were also infected with parental 63.2 MLV. An overview of the different infection experiments is shown in Table S3, while Fig. 3D shows the pooled data. Mice infected with the distinct MLV clones showed similar survival rates (Fig. 3D). Although the median survival time was slightly higher for mice infected with MLV 63.2 (90 days) than for those infected with either WT MLV or W390A MLV (82 days), no statistically significant difference was found between these MLV clones. A detailed heatmap of mouse survival, including statistical analysis, can be found in Table S3. Here, mice infected with W390A MLV display a trend toward reduced survival.

**Histopathology of mouse spleen and thymus after MLV infection.** Next, we analyzed the pathology of the spleen and thymus of mice infected with either WT or W390A MLV (Fig. 4A). At 3 and 5 weeks postinfection, prior to the development of lymphoma, spleens from mice infected with WT or W390A MLV were not enlarged (Fig. S4). In contrast, at late stages of the disease (12 to 14 weeks postinfection), the spleens of mice infected with either WT or W390A MLV were enlarged 20-fold compared to spleens from healthy mice injected with PBS (Fig. 4B and C). No statistically significant difference was found between WT MLV and W390A MLV.

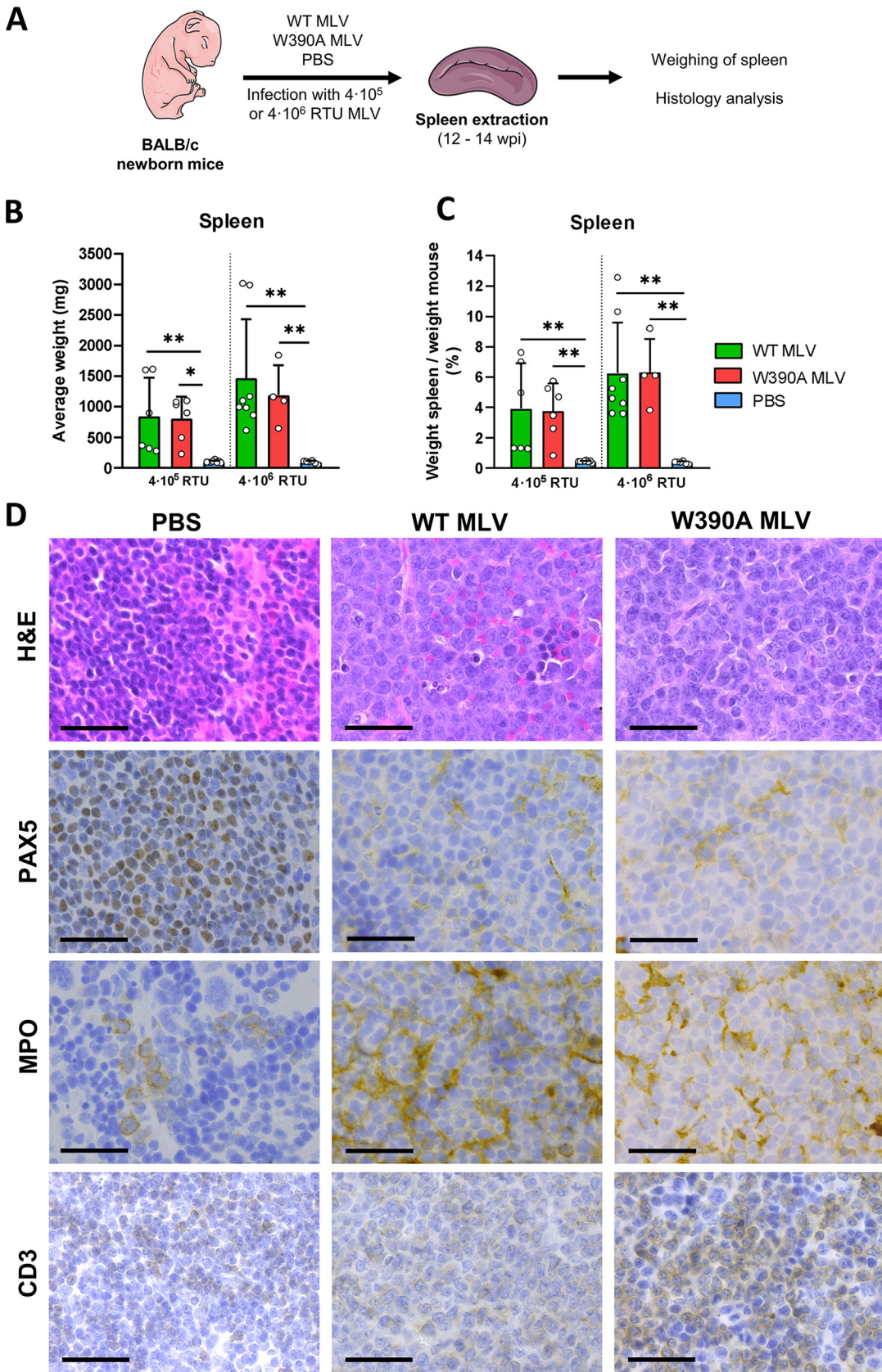
Hematoxylin and eosin (H&E) staining of spleens from mice infected with WT MLV or W390A MLV revealed an expansive white pulp with a malignant lymphoid population characterized by large, irregular nuclei and scant cytoplasm (Fig. 4D). The irregular nuclei displayed open chromatin and variably prominent nucleoli. Moreover, frequent mitotic figures were present in the spleens of mice infected with WT or W390A MLV. As opposed to these findings, spleens from mice injected with PBS displayed normal white and red pulp. In these healthy mice, the white pulp consists of small lymphocytes with rounded nuclei (Fig. 4D). We also determined the lineage contribution of the tumor cells by staining for PAX-5 (B cells), myeloperoxidase (MPO) (myeloid cells), and CD3 (T cells) (56–58). PAX-5 staining of samples from mice injected with PBS depicted normal nuclear staining of B cells in a lymphoid follicle (Fig. 4D), while MPO staining identified mainly myeloid precursors in the red pulp (Fig. 4D). PAX-5 and MPO staining of diseased mice showed nuclear and cytoplasmic staining, respectively, and the malignant population scored negative for either marker. In contrast to the markers MPO and PAX-5, CD3 staining of spleen samples from WT MLV-infected mice revealed relatively weak but extended expression across the samples, while in the spleen samples from mice infected with W390A MLV, staining was stronger. Pathology results were compatible with the development of T-cell lymphoma in mice infected with either WT MLV or W390A MLV.

Tumor infiltration was also observed in lung, liver, and kidney tissues from mice infected with WT MLV or W390A MLV (Fig. S5). Overall, these results demonstrate that mutant W390A MLV still induces T-cell lymphoma. No overt pathological differences

### FIG 3 Legend (Continued)

GAPDH levels in samples of whole blood from mice infected with WT MLV or W390A MLV, taken at 21 days postinfection. Each bar represents an individual mouse. Standard deviations were calculated from technical duplicates. (D) Survival rates of mice infected with 63.2 MLV, WT MLV, or W390A MLV. Newborn mice were infected with  $4 \times 10^6$  RTU of each molecular clone diluted in a total volume of 50  $\mu$ L PBS. Kaplan-Meier plots display the survival of mice infected with 63.2 MLV ( $n = 16$ ), WT MLV ( $n = 22$ ), or W390A MLV ( $n = 22$ ). PBS was used as a negative control ( $n = 16$ ). Each vertical step in the curve indicates one or more deaths. No statistically significant difference was found among 63.2 MLV, WT MLV, and W390A MLV using a Gehan-Breslow-Wilcoxon test.





**FIG 4** Spleen pathology in WT MLV- and W390A MLV-infected mice. (A) Schematic representation of the workflow to analyze the pathology in spleens from WT MLV- and W390A MLV-infected mice. Spleens were sampled between 80 and 140 days post-infection. (Continued on next page)

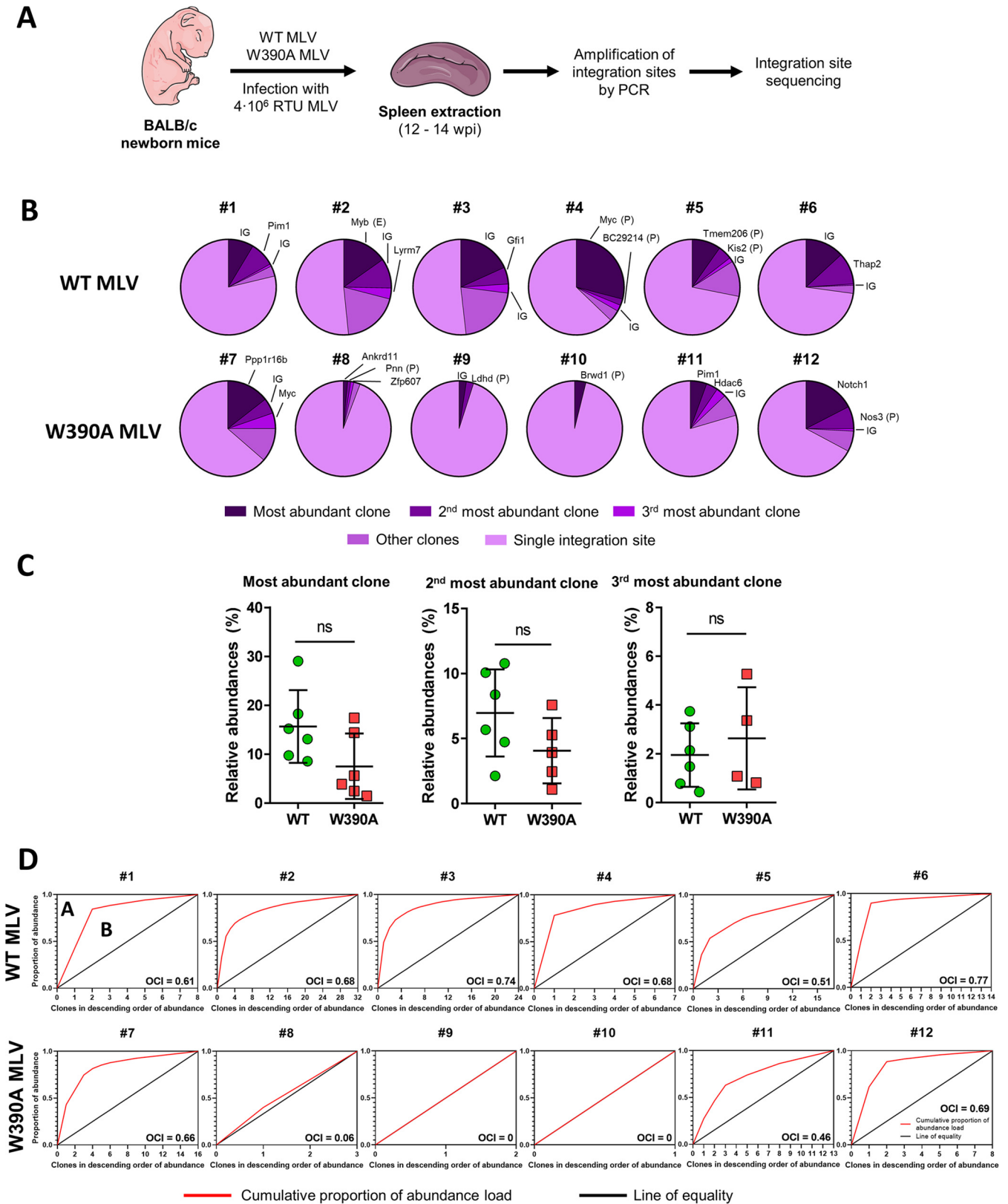
were evidenced between BET-dependent (WT) MLV and BET-independent (W390A) MLV, although stronger CD3 staining was evidenced with BET-independent MLV.

**Both WT MLV- and W390A MLV-induced lymphomas display clonal expansion yet different integration profiles.** To assess if integration site selection was altered by the W390A mutation, six mice were infected with either WT MLV or W390A MLV and sacrificed 12 to 14 weeks after infection. In contrast to the above-described integration site analysis, at this time point, the mice had already developed lymphoma due to insertional oncogenesis. The integration profiles of both viruses were analyzed as described in the legend of Fig. 5A. Briefly, spleen cells from infected mice were isolated, and genomic DNA was extracted. Integration sites were determined by ligation-mediated PCR (LM-PCR), followed by high-throughput Illumina sequencing. A summary of the integration sites recovered by this analysis is provided in Table S4. An integration site detected  $\geq 2$  times is indicative of an MLV-infected parental cell that proliferates and generates a daughter cell with the MLV genome integrated into the same position of its genome (59). We refer to these groups of cells as “clones.” Fig. 5B displays the relative abundances of the clones. Mice infected with WT MLV presented one dominant clone that represented 10 to 30% of the total identified integration sites detected in our analysis. Overall, clones represented 20 to 50% of the integration sites detected. Other integration sites corresponded to single hits. In the case of W390A MLV, half of the mice showed a different pattern, with only a few clones being present (mice 8, 9, and 10) (Fig. 5B), while the other mice displayed a clonal pattern that resembled that of mice infected with WT MLV (mice 7, 11, and 12). No statistically significant difference was identified among the relative abundances of the three most dominant clones in mice infected with WT MLV or W390A MLV (Fig. 5C). Analysis of mice using the Bangham oligoclonality index (OCI) (60) revealed that clones generated in mice infected with WT MLV were associated with a more oligoclonal profile than clones generated in mice infected with W390A MLV, where three mice depicted a monoclonal profile defined by OCI values closer or equal to 0 (mice 8, 9, and 10). The distributions of clonal abundance statistically differed between clones generated by WT MLV and W390A MLV integration (Fig. S6A). WT MLV integration into intergenic positions generated three groups of clones in the tumors: low abundance (values of 2 to 3), medium abundance (values of between 4 and 6), and high abundance (values higher than 10) (Fig. S6B). In contrast, mutant virus infection generated a high number of clones with a low abundance together and few clones with a high abundance, regardless of integration into intergenic positions, gene bodies, or promoters. These results indicate that W390A MLV still induces the clonal proliferation of lymphoma cells.

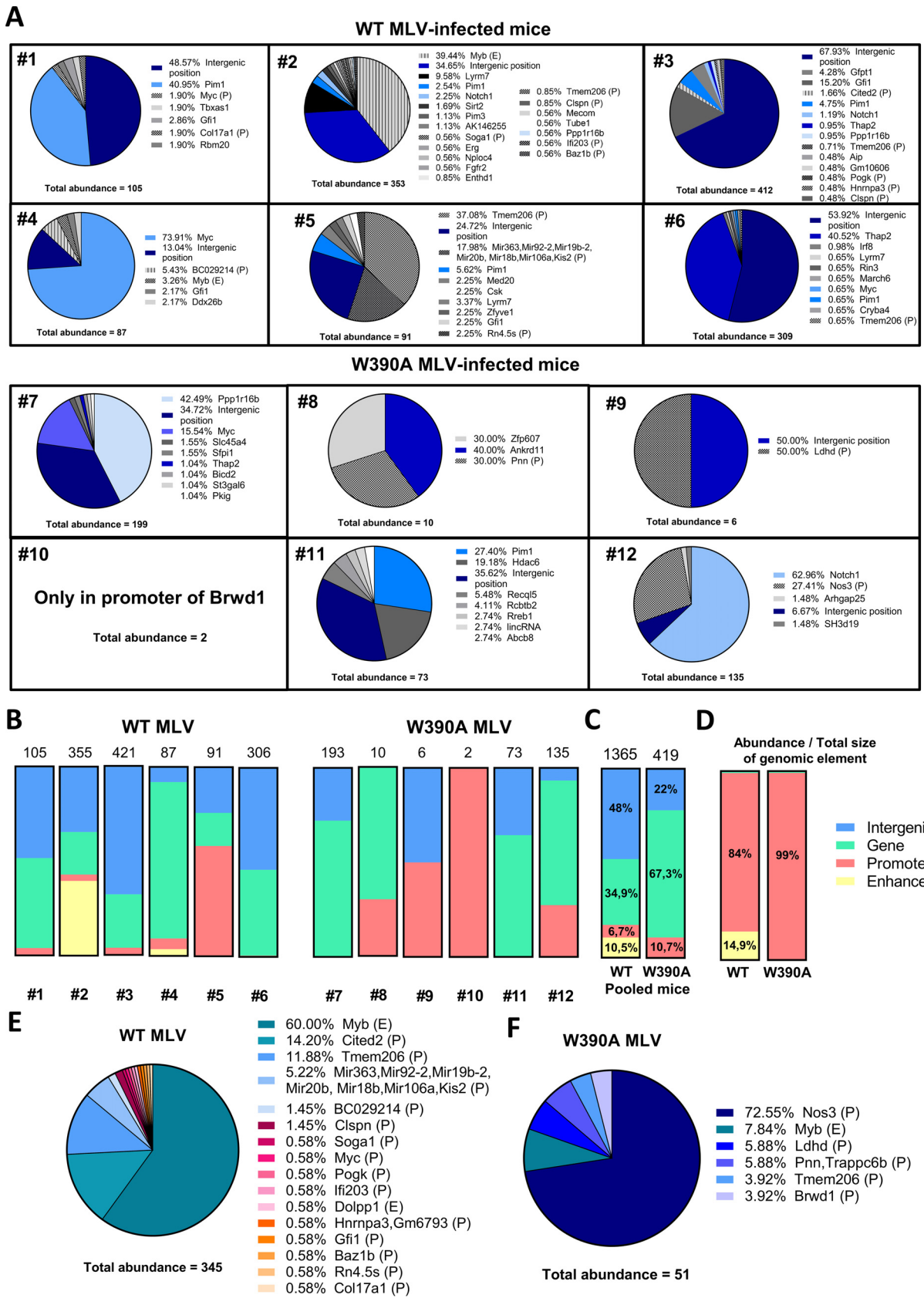
Next, we investigated the integration sites with a clonal origin in more depth. From the chromosome positions provided in Table S4, we identified genes targeted by MLV integration using an alignment to the NCBI37/mm9 mouse genome. WT MLV integration is characterized by a higher number of clonal integration sites (1,365 clones for WT MLV versus 284 clones for W390A MLV), (Fig. 6A). Both WT MLV and W390A MLV proviruses integrated into the *Myc*, *Notch1*, *Pim1*, *Ppp1r16b*, and *Thap2* oncogenes. In comparison with WT MLV, W390A MLV integrated less into intergenic positions but more frequently into the gene bodies of *Notch1* and *Ppp1r16b*. The integration of WT MLV and W390A MLV into these oncogene bodies had higher levels of clonal expansion and, therefore, abundances than those for other oncogene bodies found in this study. A chromosomal analysis of the MLV integration profile revealed that WT MLV preferentially integrated into chromosomes 5, 7, 10, and 12, while W390A MLV preferred chromosomes 2, 15, 17, and X (Fig. S7A). However, neither WT nor W390A MLV inte-

#### FIG 4 Legend (Continued)

90 days after infection. (B) Average weight of the spleens of mice infected with  $4 \times 10^6$  RTU of WT MLV ( $4 \times 10^5$  RTU,  $n = 6$ ;  $4 \times 10^6$  RTU,  $n = 9$ ) or W390A MLV ( $4 \times 10^5$  RTU,  $n = 6$ ;  $4 \times 10^6$  RTU,  $n = 4$ ) or injected with PBS ( $n = 8$  and 6 for comparisons with  $4 \times 10^5$  RTU and  $4 \times 10^6$  RTU, respectively) as negative controls. Bars show means  $\pm$  SD, while dots represent individual values. Statistically significant differences were analyzed using a Kruskal-Wallis test. (C) Percentage of the weight of the spleen in relation to the body mass of mice infected with WT MLV ( $4 \times 10^5$  RTU,  $n = 6$ ;  $4 \times 10^6$  RTU,  $n = 9$ ) or W390A MLV ( $4 \times 10^5$  RTU,  $n = 6$ ;  $4 \times 10^6$  RTU,  $n = 4$ ) or injected with PBS ( $n = 8$  and 6 for comparisons with  $4 \times 10^5$  RTU and  $4 \times 10^6$  RTU, respectively). Bars show means  $\pm$  SD, while dots represent individual values. Statistics were done using the Kruskal-Wallis test. (D) H&E, PAX-5, MPO, and CD3 staining of histological sections of spleens. Mice were infected with  $4 \times 10^5$  RTU of WT MLV or W390A MLV or injected with PBS as described above. Representative images are shown. Black arrows point to mitotic figures. PBS was used as a negative control. Bars, 20  $\mu$ m.



**FIG 5** Clonal expansion of tumor cells in the spleen induced by WT MLV or W390A MLV. (A) Schematic representation of the workflow to determine the integration site selection of WT MLV and MLV W90A in mouse tumor cells at 12 to 14 weeks postinfection. (B) Pie charts representing the relative abundances of clones and single integration sites. Each pie chart corresponds to a single mouse infected with WT MLV or W390 MLV. Intergenic site (IG), gene, enhancer (E), or promoter (P) names indicate the MLV integration sites of the three most dominant clones in each mouse. (C) Relative abundances of the top three integration sites in mice infected with WT MLV or W390A MLV. No statistically significant difference was found using a Mann-Whitney test. (D) Cumulative fraction of MLV clones from WT MLV- or W390A MLV-infected mice. The oligoclonality index (OCI) was calculated as  $OCI = A/(A + B)$ .



**FIG 6** Distribution of MLV abundances in clones generated by WT MLV or W390A MLV insertion. (A) Pie charts depicting individual gene insertional profiles of splenic cells with clonal expansion from mice infected with  $4 \times 10^6$  RTU of either WT MLV or W390A MLV. The total (Continued on next page)

**TABLE 1** Enrichment of integrations in enhancers and promoters for WT MLV and W390A MLV

MLV dose and construct	% integration/genomic size	
	Promoters	Enhancers
High dose		
WT	83.99	14.95
W390A	99.05	0
Low dose		
WT	49.85	48.79
W390A	76.11	17.96

gration into the host genome was driven by gene density (Fig. S7B). In summary, these results reveal that lymphoma cells induced by W390A MLV display a significantly different integration profile than that of lymphomas induced by WT MLV.

**W390A MLV is targeted away from enhancers *in vivo*.** We further analyzed if these integration sites corresponded to enhancer or promoter elements using EnhancerAtlas 2.0. Fig. 6B and C show MLV integration site preferences for gene bodies, intergenic positions, enhancers, and promoters. In comparison, WT MLV displayed a relatively high abundance for enhancers (as shown for individual mouse 2 and mouse 4) and intergenic positions (on pooled data, 10.5% and 48% of the total abundance, respectively), while the abundance of W390A MLV was higher in promoters (as shown in individual mice 9 and 10) or gene bodies (mice 7, 8, 11, and 12) (on average, 10.7% and 67.3% of the total abundance, respectively). Strikingly, integration into promoters and enhancers was particularly enriched considering their small size in the genome. Normalization to genomic size showed that WT MLV has an 84% abundance in promoters and a 14.95% integration preference for enhancers, whereas W390A MLV shows a 99% abundance in promoters and no integration events in enhancers (Fig. 6D) (Table 1). Next, we pooled all the abundance values for enhancers and promoters from mice infected with either WT MLV or W390A MLV in the same pie chart (Fig. 6E and F). WT MLV preferentially showed clonal expansion from integration into the *Myb* enhancer (60% of the total sites of integration into enhancer or promoter elements) and the *Cited2* (14.2%) and *Tmem206* (11.88%) promoters (Fig. 6A). Integration into the *Myb* enhancer and the *Tmem206* promoter was also detected in W390A MLV-infected cells but at lower frequencies (7.84% and 3.92%, respectively) ( $P < 0.0001$  by a  $\chi^2$  test comparing distributions).

The analysis of the integration profiles from individual mice revealed that each mouse has a unique integration profile driving the development of lymphoma (Fig. S8). In the case of W390A MLV, mice 7, 11, and 12 had a dominant clone in the tumor with integrations into the *Ppp1r16b*, *Pim1*, and *Notch1* oncogenes, respectively. In the case of WT MLV, mouse 2 displayed dominant integration at the *Myb* enhancer, while mouse 5 showed integrations into the *Tmem206* promoter. The lymphoma developed by mouse 4 was strongly driven by integration into the *Myc* oncogene.

**FIG 6** Legend (Continued)

abundance determined by NGS analysis is shown below each pie chart for both MLV constructs. The percentage before each gene name indicates the relative abundance of that gene. Genes with insertions of both WT MLV and W390A MLV are colored on a blue scale. (B) MLV abundances in intergenic positions (blue), gene bodies (green), promoters (red), and enhancers (yellow) in WT MLV- and W390A MLV-infected mice. The numeric mouse identifier is indicated below each bar. The total abundance of clonal integration sites in each mouse is indicated on top of the bar. (C) Pooled MLV abundances from intergenic positions (blue), gene bodies (green), promoters (red), and enhancers (yellow) in WT MLV ( $n = 6$ )- and W390A MLV ( $n = 6$ )-infected mice. The total abundance of pooled clonal integration sites from mice infected with either WT MLV or W390A MLV is indicated on top of the bar. The distribution of integration sites between WT MLV and W390A MLV was analyzed using a chi-squared test. The test reported a statistically significant difference between both distributions ( $P \leq 0.0001$ ). (D) MLV abundances from intergenic positions (blue), gene bodies (green), promoters (red), and enhancers (yellow) normalized by the total genomic size of each genomic feature. Data were normalized by dividing the percent integration site preference for each genomic feature by its own estimated total genomic size: intergenic positions (58.97%), genes (40.54%), promoters (0.05%), and enhancers (0.44%). Data represent results for six mice infected with WT MLV or W390A MLV. (E and F) Abundances in enhancer (E) or promoter (P) elements in spleen cells from mice infected with  $4 \times 10^6$  RTU of either WT MLV ( $n = 6$ ) or W390A MLV ( $n = 6$ ) isolated at 12 to 14 weeks postinfection. The total abundances of integration at enhancer and promoter sites are shown below each pie chart for both MLV constructs. The percentage before each gene name indicates the relative abundance of that gene. Integration into an enhancer or promoter element is indicated by E or P, respectively.

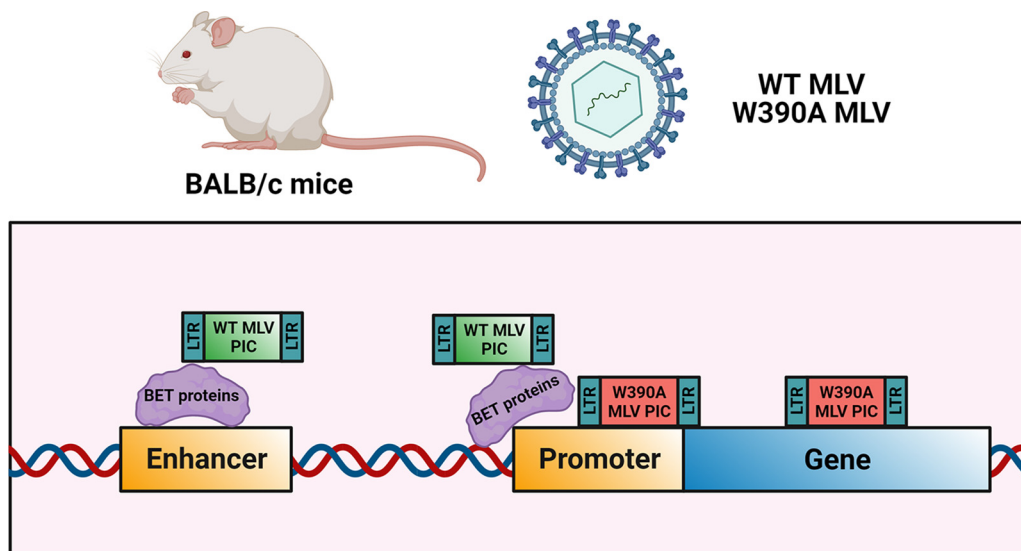
To determine the potential impact of the viral titer, we also analyzed the integration site selection of WT MLV and W390A MLV in two groups of three mice infected with a lower dose ( $5 \times 10^5$  RTU) (Fig. S9A). The results corroborate the findings in mice infected with a higher dose. Mice infected with W390A MLV at a lower dose displayed lower abundances for enhancers and promoters (Fig. S9B). When pooling the data from three mice, 3.1% of the total abundance was in enhancers, while 26.7% of the total abundance was found in promoters from WT MLV-infected mice (Fig. S9C). In contrast, mice infected with W390A MLV displayed only 1.3% of the total abundance in promoters and 2.7% in enhancers (Fig. S9C). Similar to mice infected with a higher dose of MLV, integration and clonal expansion from enhancers and promoters were enriched (for WT MLV, 49.9% in promoters and 48.9% in enhancers; for W390A MLV, 76.1% in promoters and 18% in enhancers) (Fig. S9D) (Table 1). All mice displayed a polyclonal integration profile, except for mouse 29, which displayed an oligoclonal integration profile (Fig. S9E). Integration into the *Myb* enhancer was found in WT MLV-infected mice (mice 29 and 30). In mice infected with W390A MLV, integration into *Notch1* and *Pvt1* oncogene bodies was evident (Fig. S9F).

Finally, we analyzed if MLV insertional mutagenesis resulted in truncated oncogenes using next-generation sequencing (NGS) analysis of WT MLV- or W390A MLV-infected mice. Integrations of both WT MLV and W390A MLV into *Thap2*, *Pim1*, and *Notch1* occurred toward the 3' end of the gene body (Fig. S10A). However, both WT and W390A MLV integrated into the central part of the *Ppp1r16b* gene. The integration preferences of WT MLV and W390A MLV differed with respect to the *Myc* gene; whereas WT MLV integrated near the 3' end, W390A MLV integrated near the 5' end. Previous studies on the mechanism of insertional mutagenesis reported that MLV insertion near gene bodies can increase gene expression due to intrinsic enhancer and promoter elements within the long terminal repeats (LTRs) of MLV (6, 61). Common targets for MLV integration are regions near the *Lmo2*, *Mecom2*, and *Prdm16* genes in humans (62) or mice (61). An intragenic insertion into these genes was not identified in our integration site analysis, but we analyzed the expression of these genes using RT-qPCR in spleen cells from a different set of MLV-infected mice, isolated at 12 to 14 weeks postinfection, by qPCR. The results show that MLV 63.2, WT MLV, and W390A MLV effectively increase the expression of *Mecom2* and *Prdm16* (Fig. S10B). However, 63.2 MLV and WT MLV were able to induce a stronger increase in *Mecom2* and *Prdm16* expression than W390A MLV.

Although most WT MLV integrations took place at intergenic positions and gene bodies, there was a relative enrichment for enhancers and promoters that represent only 0.49% of the mouse genome. The proportions of integration into enhancers and promoters in mice infected with WT MLV were 17.2% with the high dose and 29.8% with the low dose. BET-independent W390A MLV IN integrated away from enhancers in tumorigenic cells. At a high dose, only 10.7% of W390A MLV integrated into enhancers/promoters, and with a low dose, only 4% integrated into enhancers/promoters (Table 1).

## DISCUSSION

**BET proteins are not required for MLV *in vivo* integration and replication.** Our work unequivocally demonstrates that the interaction between BET proteins and MLV IN is not essential for MLV integration, replication, and lymphomagenesis *in vivo*. We reported previously that the W390A mutation abrogates the total interaction with BET proteins using an AlphaScreen assay, to a level similar to that with the MLV IN<sub>1-381</sub> construct that lacks the C-terminal tail of the MLV IN (19). In the present study, we demonstrate that W390A MLV can replicate under *in vivo* conditions, which strongly suggests that additional host or viral factors may be involved during integration (Fig. 7). We have shown that the *in vivo* integration of W390A was retargeted away from TSSs and CpG islands but to a lesser extent than with *in vitro* data (27, 28). This difference might be explained by the implication of different host cofactors considering that the cell types analyzed in the present study (splenic cells) and those analyzed in previous studies (CD34<sup>+</sup> hematopoietic cells, SupT1 cells, and HeLa cells, among others) are not the same. While the potential alternative cofactor remains unknown, these findings are in line with our previous observations for HIV



**FIG 7** Model of MLV integration *in vivo*. The MLV preintegration complex (PIC) is targeted to enhancers and promoters by the BET proteins, while the lack of an interaction between MLV PIC and BET proteins retargets integration toward promoters and gene bodies by the implication of an unknown host factor or chromatin feature or a direct implication of the MLV integrase. The strong MLV LTR enhancer/promoter can drive oncogenesis, even after retargeting.

integration. We and others reported previously that heparin binding growth factor (HDGF)-related protein 2 (HRP2) can replace LEDGF/p75 in LEDGF/p75 knockout cells (63, 64).

Our data show that the BET-independent MLV W390A mutant integrates and replicates in mice to an extent similar to that of the WT virus. This suggests that MLV IN either directly interacts with chromatin or uses alternative host factors or mechanisms. A likely candidate is p12, a cleavage product of MLV Gag that has been implicated in the early steps of the replication cycle of MLV (65, 66) as well as virion production (67). The MLV protein p12 can tether the MLV preintegration complex to host chromosomes (68, 69) by the phosphorylation of serine 61, the key event to tether the MLV preintegration complex to chromatin (49, 70). Yueh and Goff confirmed that the lack of p12 phosphorylation of the serine amino acids at positions 61 and 78 results in the impaired formation of LTR circles (71), but these mutants could still release normal levels of mature virions, suggesting that p12 may not be essential for the replication of MLV. On a similar trend, another study showed that alteration of p12-mediated tethering had no effect on MLV integration into TSSs (72). The relative roles of viral p12 and cellular BET proteins in tethering and targeting integration await further investigation. Possibly, tethering by p12 of MLV PICs to mitotic chromosomes is an essential feature while BET-mediated targeting is an optional feature that still may have evolutionary advantages since BET proteins remain bound to mitotic chromosomes and are associated with early gene transcription after the completion of mitosis (73).

**The MLV-BET interaction is not the essential driver of tumorigenesis.** BALB/c mice infected with either WT MLV or W390A MLV developed the same disease phenotype. Upon evaluation, we identified an enlargement of the thymus and spleen, both typical of lymphoma (74). Histopathology identified the same type of malignant population in both WT MLV- and W390A MLV-infected mice by H&E staining. These malignant cells are transcriptionally and mitotically more active than cells from mice injected with PBS. MPO and PAX-5 staining revealed the lack of B cells, granulocytes, and monocytes in spleen lymphomas. Histopathology and marker analyses revealed strong CD3 staining in mice infected with WT or W390A MLV, in contrast to mice injected with PBS, indicating that these mice developed lymphomas of T-cell precursors (75). Moreover, the lymphomas were characterized by polyclonal expansion. Lymphomas induced by retroviruses often display a polyclonal phenotype (76) due to integration into multiple loci in the same cell, which eventually leads to oncogenesis by the disruption of gene networks that

control proliferation (77). In summary, all tumors induced by WT MLV infection were characterized by a polyclonal profile. On the other hand, tumors induced by W390A MLV infection displayed a mixed profile consisting of polyclonal, oligoclonal, and monoclonal tumors (in the latter case for 1 out of 6 mice, corresponding to mouse 10).

**The lack of BET-MLV IN interaction still yields lymphomas through insertional mutagenesis in oncogene bodies.** In the context of gene transcription, BRD4 plays an important role in superenhancers (SEs) (78). SEs have been defined as clusters of enhancers that are occupied by a high density of multiple transcription factors (79). In this sense, BRD4 and MED1 can act as coactivators of the SEs, and their presence defines an SE (80). In cancer cells, the acquired cancerous phenotype relies on abnormal transcription promoted by superenhancers (78). Although BRD4 may acquire this function at later stages when a lymphoma develops in MLV-infected mice, the tethering mechanism whereby MLV IN binds to BRD4 may point to a role during the retroviral integration step. We show that MLV preferentially integrates into promoters and enhancers under *in vivo* conditions, which is in line with results reported previously by other groups using cell lines and primary cells (5, 6, 81). For WT MLV, integration into enhancers was strongly enriched, while W390A MLV integration was enriched in promoters and gene bodies. Normalization of the relative abundance in Fig. 6C to genomic size resulted in an integration preference of >98% (WT MLV, 98.9%; W390A MLV, 99%). Whereas WT MLV preferentially integrates into promoters and enhancers, BET-independent MLV preferentially integrates into promoters and gene bodies. The role of BRD4 is to target integration into enhancers known to interact with BRD4. W390A MLV integration into oncogene bodies has major implications for the development of lymphomas in infected mice. As shown in Fig. 6A, 4 out of 6 mice infected with W390A MLV had a dominant clone corresponding to integration into an oncogene body. Opposite this, only 1 out of 6 mice infected with WT MLV had a dominant clone as a result of integration into an oncogene body. Therefore, not only are integrations into gene bodies more frequent (as shown in Fig. S8 in the supplemental material), but also, once W390A MLV integrates into oncogene bodies, these genes are selected for driving the development of lymphomas in W390A MLV-infected mice. Still, W390A MLV targets a smaller set of oncogenes. Perhaps this is due to less efficient BET-independent integration or the use of a redundant targeting mechanism with another (host) factor.

WT MLV integration into the enhancer of *Myb* displayed strong clonal expansion (82), while W390A MLV was retargeted away from *Myb*, a common target for MLV integration. Also, WT MLV integrated into *Tmem206* and *Cited2* promoters, which induced strong clonal expansion. Both *Tmem206* and *Cited2* were recently described as oncogenes involved in lymphoma (83, 84) in human cells. Overexpression of *Myb* has been detected in both acute myeloid leukemia and non-Hodgkin lymphoma (85, 86), while *Tmem206* and *Cited2* overexpression was detected in colon cancer and prostate cancer, respectively. The molecular mechanism implicated in oncogenesis driven by *Tmem206* is the activation of the AKT signaling pathway (87). Moreover, both parental 63.2 MLV and WT MLV enhanced the expression of *Mecom2* and *Prdm16* (Fig. S10), while the increase in expression induced by W390A MLV was 2-fold lower, compatible with a lower integration frequency of W390A MLV near enhancers and/or promoter elements. The W390A mutation in MLV IN significantly retargeted viral integration away from enhancers toward oncogene bodies such as *Notch1* and *Ppp1r16b*. *Notch1* has been widely described as one of the driver genes in T-cell acute lymphoblastic leukemia/lymphoma (T-ALL) (88–90), while *Ppp1r16b* can promote aggressive lymphomas in mice (91). The truncation of *Notch1* can generate lymphomas by the loss of the PEST domain located in the 3' end of the *Notch1* gene, which leads to an enhanced protein half-life and, therefore, longer activity as a transcription factor in the nucleus (92). Insertional mutagenesis of *Notch1* can also result in deletions in the 5' end that activate intragenic promoters driving the expression of truncated transcripts that lack the negative regulatory region (NRR) (93). In the case of *Ppp1r16b*, a previous study reported a modest increase in the expression of *Ppp1r16b* in several MLV-induced lymphomas (91). Such an increase was more relevant when an MLV insertion happened at the *Ppp1r16b* locus.

As such, by combining the data for all the mice used in this study infected with a higher dose of MLV (6 mice for WT MLV and W390A MLV) (Fig. 5 and 6) and a lower



dose of MLV (3 mice for WT MLV and W390A MLV) (Fig. S9), our results indicate that tumors induced by WT MLV (44.4%, or 4 out of 9 mice, corresponding to mice 2, 5, 27, and 28) could be caused by the overexpression of oncogenes (such as *Myb*) resulting from integration into an enhancer or promoter near such an oncogene, while the tumors induced by W390A MLV (77.7%, or 7 out of 9, corresponding to mice 7, 8, 11, 12, and 30 to 32) could be generated frequently after insertional mutagenesis into an oncogene body.

**Importance of the C-terminal tail of MLV IN in viral replication.** The C-terminal deletion of MLV IN has a major impact on the replication of MLV since  $\Delta$ C MLV was associated with reduced viral replication in comparison with W390A MLV. Although the reason for this is unknown at the moment, we speculate that MLV IN might have an effect on the assembly and release of newly formed MLV particles. The role of the MLV IN-BET protein interaction in an *in vivo* context was previously studied by Loyola et al. (51). In contrast to our present study with the W390A mutant, distinct BET-independent MLV constructs were used. First, a construct carrying a truncation of 23 amino acids from the C-terminal tail of MLV IN (MLV IN XN) (94) deficient for BRD4 interactions was generated. Next, an MLV IN construct lacking the terminal peptide (TP<sup>-</sup>) was made, including stop codons in the non-*env* reading frames to prevent the restoration of WT IN. The *MYC/Runx2* tumorigenesis mouse model was used by Loyola et al. (51), which is characterized by the spontaneous development of lymphomas (95). As such, it does not reflect primary insertional lymphomagenesis upon MLV infection. We used healthy BALB/c mice instead, which develop tumors only upon MLV infection. Both the MLV IN XN and MLV IN TP<sup>-</sup> constructs showed a 10-fold decrease in *in vitro* infectivity, which corresponds more or less to the 25-fold decrease in infectivity seen with  $\Delta$ C p63.2 MLV, lacking 27 C-terminal amino acids, compared to WT MLV and W390A MLV 24 h after infection of NIH 3T3 cells (Fig. S1D). While W390A MLV replicated to WT levels in our mouse model, no evaluation of infectivity under *in vivo* conditions was performed by Loyola and colleagues (51), but it can be deduced that the replication of MLV with IN C-terminal deletions was hampered.

Whereas W390A MLV-induced lymphomagenesis showed a trend toward reduced mouse survival, the TP-deleted MLV clones described previously by Loyola and colleagues resulted in the increased survival of *MYC/Runx2* mice in 15 days (51). This result may, however, reflect reduced virus replication *in vivo*. In the latter study, recombination with endogenous retrovirus *Pmv20* was able to revert the expression of the terminal peptide, confounding the analysis (51). We used BALB/c mice, characterized for the presence of polytropic and xenotropic endogenous retroviruses (96). Although mice infected with W390A MLV displayed a higher number of single mutations in MLV IN, suggesting selective pressure, no revertants due to mutagenesis or recombination were observed (Fig. S2). As for the integration site analysis, both studies report less integration into CpG islands and TSSs with BET-independent virus. In conclusion, both studies report BET-independent MLV integration in mice. In both studies, BET-independent MLV is retargeted but still associated with lymphoma. In the previous study, a poor replication capacity and recombination hampered the analysis of the role of BET interactions in lymphomagenesis, while our setup allowed a more detailed pathological and genomic analysis.

**Genotoxicity of BET-independent MLV vectors.** One important application of MLV integration studies is the development of safer MLV-based vectors for gene therapy. Previously, it was shown that the safety of SIN MLV vectors varies in different tissues (64). In this study, we have used the W390A mutation in MLV IN to obtain a better understanding of the relevance of BET proteins in the development of lymphoma. Overall, our results indicate that the prevention of the MLV IN-BET interaction was sufficient to retarget integration away from enhancers but not enough to prevent the development of lymphomas by insertional mutagenesis. Our analysis points to the additional role of the enhancer/promoter in the MLV LTR in driving oncogenesis. Second-generation SIN MLV vectors take this feature into account and are associated with an increased safety profile (48). An *in vivo* analysis of BET-independent SIN MLV vectors is warranted. Such MLV vectors may have a better genotoxic safety profile.

**Evolutionary advantage of the use of BET proteins for MLV integration.** If the interaction of MLV IN with BET is dispensable for MLV replication *in vivo* and lymphomagenesis, why did MLV evolve this mechanism? Although different explanations can be thought of, BET-mediated integration may provide a more efficient pathway. First, in the present study and the study reported previously by Loyola et al. (51), the number of integrations detected in mice infected with either W390A MLV or tail peptide-deleted MLV was lower than the number of integrations for WT MLV. In this sense, BET-dependent MLV may be able to establish a higher number of proviruses inside the host genome and a greater chance for productive infection. Second, we show how BET proteins direct MLV integration *in vivo* toward enhancers and promoters. Integration at these sites ensures high transcriptional activity. In fact, BRD4 can remain bound to several loci to remember the gene transcriptional profile after mitosis (73), a process that MLV relies on for completing nuclear entry. Integration at these sites may guarantee that the transcriptional machinery can initiate the transcription of the MLV genome and continue the replication cycle. The strong enhancer/promoter of the MLV LTR clearly contributes to the high transcriptional activity, which, in contrast to HIV, is independent of a transactivator such as Tat. Lymphomagenesis may not be an intrinsic feature of the MLV replication cycle but may be a bystander effect. Although this is not beneficial for a pathogen, in the case of mice, horizontal transmission to offspring is guaranteed before their death considering how early in their life span mice reach adulthood. The integration preference near enhancers and promoters and the strong LTR promoter all contribute to the risk for lymphomagenesis. Whether integration near oncogenes is stimulated by BRD4 interactions or whether BRD4-mediated integration near oncogenes is selected for clonal proliferation is not fully clear. Our data point toward the second scenario, but more detailed analyses are required. The multipronged mechanism to ensure proviral expression after transcription and, thus, the relative absence of latency may be associated with an intrinsic risk for insertional mutagenesis, clonal expansion, and tumorigenesis.

## MATERIALS AND METHODS

**Cell culture.** The human kidney HEK 293T/clone 17 (293T/17) cell line was acquired from the ATCC (ATCC CRL-11268). The NIH 3T3 cell line was a kind gift from Johan Van Lint from the Laboratory for Protein Phosphorylation and Proteomics (KU Leuven). Both cell lines were maintained in Dulbecco's modified Eagle's medium (DMEM)-GlutaMAX I (Gibco, Thermo Fisher Scientific) with 10% fetal bovine serum (FBS) (Gibco) at 37°C with 5% CO<sub>2</sub>.

**Animal tests.** BALB/c mice were acquired from Janvier Labs (Le Genest-Saint-Isle, France). All the experiments and procedures with animals were done according to European Directive 2010/63/EU for the protection of animals used for scientific purposes. The procedures were also reviewed by the Ethical Commission of Animal Tests (Ethische Commissie Dierproeven) of KU Leuven (internal approval number P210/2014, extended P116/2019). An overview of mice used in this study (including a reference to each figure and experiment) is shown in Table S1 in the supplemental material.

**MLV constructs.** Mutations in MLV IN of pNCS (71) were done using the site-directed ligase-independent mutagenesis (SLIM) cloning technique (97). Next, the SgrAI/Sall fragment was shuttled into the MLV molecular clone p63.2, which was obtained from Susan R. Ross (98). Plasmid p63.2 is a pBR322 plasmid containing the wild-type Moloney MLV provirus clone (99, 100). The C-terminal integrase sequences from the generated MLV clones are shown in Fig. S1.

**MLV production.** Approximately  $5.5 \times 10^6$  HEK 293T cells were plated into a petri dish. For *ex vivo* experiments, DMEM-GlutaMAX I (Gibco, Thermo Fisher Scientific) with 10% FBS (Gibco) and 50  $\mu$ g/mL gentamicin (Gibco) was used. For *in vivo* experiments, Opti-MEM I-GlutaMAX I (Gibco) supplemented with 2% FBS and 50  $\mu$ g/mL gentamicin (Gibco) was used. Twenty-five micrograms of the viral plasmid and 100  $\mu$ L of branched polyethyleneimine (PEI) (pH 7.4) (Sigma-Aldrich) were added to the cell culture. At 24 h posttransfection, fresh medium was added. Medium was harvested at 48 and 72 h posttransfection and filtered through a 0.45- $\mu$ m filter (Merck Millipore, Overijse, Belgium). MLV was concentrated by ultrafiltration using Vivaspin 50-kDa-molecular-weight-cutoff (MWCO) filter tubes (Merck Millipore).

**Semiquantitative PCR and RT-qPCR.** RNA extractions were done using the Aurum total RNA kit (Bio-Rad) according to the manufacturer's instructions. RNA was quantified using an SP062 nanophotometer (Implen, Munich, Germany), and 5  $\mu$ g of each sample was taken for reverse transcription. RT-PCR was done using the high-capacity cDNA reverse transcription kit (Applied Biosystems, Thermo Fisher Scientific, Brussels, Belgium).

Semiquantitative PCR was performed using iProof high-fidelity (HF) kit polymerase (Bio-Rad, Temse, Belgium) according to the manufacturer's instructions. PCR was done on the T personal Biometra thermocycler (Westburg). The cycling conditions were 98°C for 45 s and 40 cycles at 98°C for 20 s, 55°C for 20 s, and 72°C for 90 s, ending with 1 cycle at 72°C for 5 min. Semiquantitative PCR products were analyzed by gel electrophoresis with a 1% agarose gel (catalog number 15510-027; Invitrogen), using the Bio-Rad PowerPac basic electrophoresis power supply at 140 V for 40 min. A 1-kb GeneRuler (Thermo Fisher Scientific) was used as a ladder.

**TABLE 2** Primers used for SG-PERT, quantitative PCR, and semiquantitative PCR

Gene	Sequence		Probe	Technique
	Forward primer	Reverse primer		
MS2 phage PERT	TCCTGCTCAACTTCCTGTCGAG	CACAGGTCAAACCTCCTAGGAATG		SG-PERT
MLV Env	CCTACTACGAAGGGGTTG	CACATGGTACCTGTAGGGGC		qPCR
MLV IN	AAGCTCAGGCCAGGTAGAAAGAA	TCCCGATCTCCATTGGTTACCT		Revertant analysis
Lmo2	ACCCTACTTCCTGAAAGCC	GACACCCACAGAGGTCACAG		qPCR
Mecom2	CGACAGAGGAGTGGGAGAAG	TCTAAGATCTCCGATTCTACGGC		qPCR
Prdm16	CAGCACGGTGAAGCCATTC	GCGTGCATCCGCTTGTG		qPCR
Rodent GAPDH	TGTGTCCGTCGTGGATCTGA	CCTGCTTACCACCTTCTTGA	FAM-CCGCCTGGAGAAACCT GCCAAGTATG-TAMRA	qPCR

Quantitative PCR (qPCR) was performed using LightCycler 480 SYBR green I master mix (reference number 04707516001; Roche) or IQ Supermix (reference number 1708860; Bio-Rad) for the glyceraldehyde-3-phosphate dehydrogenase (GAPDH) gene with a 6-carboxyfluorescein (FAM)-6-carboxytetramethylrhodamine (TAMRA)-labeled probe. Primer concentrations were adjusted to a final concentration of 300 nM. Two-step qPCR was done using the LightCycler 480 instrument (Roche, Anderlecht, Belgium). The activation cycle was performed at 95°C for 10 min, followed by 40 cycles at 95°C for 10 s and then 60°C for 30 s. Cooling was done by 1 cycle at 37°C during 1 s. Primer sequences used for SYBR green product-enhanced reverse transcriptase (SG-PERT) assays, quantitative PCR, semiquantitative PCR, and the amplification of integration sites are listed in Tables 2 and 3.

**MLV titration by a SYBR green product-enhanced reverse transcriptase assay.** Titration was done as described previously (101). MLV was produced in NIH 3T3 cells. The supernatant was filtered through a 0.45-µm filter (Millipore). The MLV supernatants and standards were lysed using 2× lysis buffer (0.25% Triton X-100, 50 mM KCl, 100 mM Tris-HCl, 40% glycerol [pH 7.4]) with incubation at room temperature for 10 min. After lysis, 1× sample dilution buffer was added to each sample and standard [10× dilution buffer contains 50 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 200 mM KCl, and 200 mM Tris-HCl (pH 8.3)] at a final dilution of 1:10. Standards were diluted in a dilution series of 1:5. PCR was performed with a final concentration of 0.25 U/µL Fermentas TruStart Hot Start *Taq* DNA polymerase (Thermo-Fisher) in 100 µL of a solution containing 2× reaction buffer [5 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 20 mM KCl, 20 mM Tris-HCl (pH 8.3)], 10 mM MgCl<sub>2</sub>, 0.2 mg/mL bovine serum albumin (BSA), a 1:10,000 dilution of SYBR green I, 500 µM deoxynucleoside triphosphates (dNTPs), 4 µM forward primer MS2 phage PERT, 4 µM reverse primer MS2 phage PERT (both listed in Table 2), and 2 µg/mL MS2 RNA. The thermocycler used was the LightCycler 480 instrument (Roche). The cycling conditions were 37°C for 60 min, 95°C for 5 min, and 45 cycles at 95°C for 5 s, 55°C for 5 s, and 72°C for 15 s. The final cycle was done at 81°C for 10 s.

**Quantification of viral DNA in cells from the spleen, thymus, and blood.** Genomic DNA was extracted from mouse spleen, thymus, and blood samples using the commercial GenElute mammalian genomic DNA miniprep kit (Sigma-Aldrich). Two-step quantitative PCR was performed using MLV Env primers (102) (listed in Table 2) that amplify the region of the MLV genome located between bp 6546 and 6765 (NCBI reference sequence accession number [NC\\_001501.1](#)). Rodent GAPDH was used as a housekeeping gene using the LightCycler 480 instrument (Roche, Anderlecht, Belgium). The activation cycle was performed at 95°C for 10 min, followed by 40 cycles at 95°C for 10 s and then 60°C for 30 s. Cooling was done in 1 cycle at 37°C during 1 s.

**Quantification of MLV titers by coculture.** For the quantification of MLV titers by coculture, round, 12-mm-diameter BioCoat coverslips (Corning; Fisher Scientific, Merelbeke, Belgium) were first added to a 24-well plate (Nunc; Thermo Scientific, Asse, Belgium). Next, approximately 1.5 × 10<sup>5</sup> NIH 3T3 cells were seeded into each well. After 24 h, NIH 3T3 cells were cocultured with 10<sup>4</sup> cells from the spleen or thymus for 24 h. Next, NIH 3T3 cells were rinsed with PBS and fixed with 4% paraformaldehyde. For staining, a primary antibody against the MLV protein p12 was obtained from hybridoma alpha CA cells (ATCC CRL-1890). The primary antibody was incubated for 1 h. Goat anti-mouse biotin (Dako Denmark, Glostrup, Denmark) was used as a secondary antibody, and samples were incubated for 20 min, followed by an incubation step with streptavidin-horseradish peroxidase (HRP) (Dako Denmark) for 30 min. Next, 3,3'-diaminobenzidine (DAB) staining was done. Coverslips were mounted using Mowiol 4-88 (Calbiochem, Merck). Infected cells were counted with a Leica DMR microscope (×40 magnification).

**TABLE 3** Primers used for amplification of integration sites<sup>a</sup>

Amplification step	Sequence	
	Oligonucleotide 1	Oligonucleotide 2
PCR1	ATTCGAAGCGAACAACACTGAC	CCTCTTGCAAGTTCATCCGAC
PCR2	AATGATACGGCGACCACCGAGATCTACACCAGGACTGAC GCTATGGTAATTGTTGACCCCTCGGATTACCCG	CAAGCAGAAGACGGCATAACGAGATXXXXXXXXXXXXAGTCAG TCAGCCTGATTGACTACCCGTCAGCGG

<sup>a</sup>X corresponds to nucleotide positions used as barcodes for sequencing. The set 1 oligonucleotides from PCR1 and PCR2 bind to the linker, while the set 2 oligonucleotides bind to the MLV LTR.

**Mouse infection.** One-day-old mice were infected by intraperitoneal injection with 5 or 50  $\mu\text{L}$  of  $8 \cdot 10^4$  RTU/ $\mu\text{L}$  of MLV, as indicated for each experiment. Moribund mice were sacrificed when the disease reached a late stage, around 12 weeks postinfection, by cervical dislocation. Before sacrifice, mice were weighed. Samples of whole blood, thymus, and spleen were obtained for histological analysis.

**Blood sampling and cell counting.** Blood samples were taken from the submandibular vein by puncture, and blood was stored in MAP microtubes coated with 1 mg EDTA (Becton, Dickinson, Franklin Lakes, NJ). The blood was diluted 1:10 in PBS, and blood cell counting was performed using the Siemens (Munich, Germany) Advia 2120 hematology analyzer. For further confirmation, blood smears were performed, and blood cells were identified and counted microscopically to corroborate the results from the automatic cell counter.

**Histological sections.** Tissues were fixed in 10% neutral buffered formalin for 24 h and embedded in paraffin. Formalin-fixed and paraffin-embedded sections were cut at a 3- $\mu\text{m}$  thickness. Automated hematoxylin and eosin (H&E) staining was performed on a Leica (Wetzlar, Germany) ST5010 Autostainer XL instrument for all samples.

**Integration site sequencing.** Integration site sequencing was done as described previously (103). Briefly, mice were injected with WT MLV or W390A MLV. At 3 and 5 weeks postinfection or at a later stage of the disease (12 to 14 weeks postinfection), mice were sacrificed, and spleens were collected. Genomic DNA was extracted using the GenElute mammalian genomic DNA miniprep kit (Sigma). Sonication with the Covaris M220 instrument was used to shear genomic DNA randomly, and DNA linkers were ligated to the sheared DNA. Integration sites were amplified by nested PCR using iProof high-fidelity kit polymerase (Bio-Rad). Two sets of oligonucleotides that bind to the MLV LTR and the linker were used for PCR (103) (sequences are listed in Table 3), whereby Illumina sequencing adapters were linked to the second oligonucleotide. PCR products were purified using AMPure XP magnetic beads and sequenced using Illumina MiSeq (Illumina Inc., San Diego, CA, USA) paired-end reads with 300 cycles by the Leuven Genomics Core (KU Leuven). The relative abundance of each unique integration site was calculated according to the following formula: relative abundance = (abundance of integration sites/total abundance)  $\times$  100.

**Revertant analysis.** Genomic DNA was extracted using the GenElute mammalian genomic DNA miniprep kit (Sigma). PCR using iProof polymerase was performed with the following reaction mix: 10  $\mu\text{L}$  iProof HF buffer, 1  $\mu\text{L}$  dNTP mix (250  $\mu\text{M}$  each), 1  $\mu\text{L}$  forward primer (5  $\mu\text{M}$ ), 1  $\mu\text{L}$  reverse primer (5  $\mu\text{M}$ ), 1  $\mu\text{L}$  DNA template (20 ng), 35.5  $\mu\text{L}$  nuclease-free water, and 0.5  $\mu\text{L}$  iProof polymerase. The primers used for PCR (both listed in Table 3) amplify the region of the MLV genome located between bp 4988 and 5658 (NCBI reference sequence accession number [NC\\_001501.1](#)). The cycling conditions were 98°C for 30 s, followed by 30 cycles at 98°C for 10 s, 54°C for 20 s, and 72°C for 30 s. The final cycle was done at 72°C for 10 min. The PCR product was sequenced by LGC Biosearch Technologies by Sanger sequencing.

**Statistics and software.** *P* values associated with each graph are represented by asterisks in the figures (\*,  $P \leq 0.05$ ; \*\*,  $P \leq 0.01$ ; \*\*\*,  $P \leq 0.001$ ; \*\*\*\*,  $P \leq 0.0001$ ). The statistical tests used to analyze significant differences between conditions are indicated in each figure legend. INSPIRED software was used to analyze the integration sites (103). Integration site positions were aligned to the NCBI37/mm9 mouse genome to identify insertions in mouse genes using UCSC Genome Browser. The genomic sizes of enhancers in mouse spleens were obtained from EnhancerAtlas 2.0. The genomic sizes of genes were obtained from the UCSC gene database (NCBI37/mm9, July 2007 assembly). The genomic sizes of mouse promoters were obtained from the EPD new promoter database (GRC38/mm10, December 2011 assembly). EnhancerAtlas 2.0 was used to analyze integration into enhancer and promoter features (104). GraphPad Prism 8 was used to make graphs and statistical calculations.

## SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

**SUPPLEMENTAL FILE 1**, XLSX file, 0.02 MB.

**SUPPLEMENTAL FILE 2**, XLSX file, 0.01 MB.

**SUPPLEMENTAL FILE 3**, XLSX file, 0.01 MB.

**SUPPLEMENTAL FILE 4**, XLSX file, 0.2 MB.

**SUPPLEMENTAL FILE 5**, PDF file, 5.6 MB.

## ACKNOWLEDGMENTS

This work was supported by the Research Foundation Flanders (FWO) (G0A5316N and SBO-SAPHIR) and the KU Leuven Research Council (C14/17/095-3M170311). D.V.L., S.V.B., and A.B. all received doctoral fellowships from the FWO. F.C. is an Industrial Research Fund (IOF) fellow.

We thank Alexander Tzankov for preparation of histology samples of lung, liver, and kidney; Nancy Boeckx (KU Leuven) for the analysis of blood smears; as well as Susan R. Ross for critical review and reading and providing us the p63.2 MLV construct. We thank Servier Medical Art for allowing the free use of their pictures in scientific publications under the Creative Commons License.

J.D.R., I.N., and Z.D. conceptualized the experiments of this study; M.M., D.V.L., S.V.B., and S.E.A. performed the experiments and analyzed the data; I.N. and A.B. pooled and

analyzed genetic data; L.M., T.T., and J.S. performed the histology analysis; R.G., F.C., and Z.D. supervised experimental work; I.N. wrote the first draft of the manuscript; and Z.D. supervised the project and reviewed the final version of the manuscript. All authors read and approved the final manuscript.

We declare that we have no conflicts of interest.

## REFERENCES

- Lewinski MK, Yamashita M, Emerman M, Ciuffi A, Marshall H, Crawford G, Collins F, Shinn P, Leipzig J, Hannenhalli S, Berry CC, Ecker JR, Bushman FD. 2006. Retroviral DNA integration: viral and cellular determinants of target-site selection. *PLoS Pathog* 2:e60. <https://doi.org/10.1371/journal.ppat.0020060>.
- Mitchell RS, Beitzel BF, Schroder ARW, Shinn P, Chen H, Berry CC, Ecker JR, Bushman FD. 2004. Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. *PLoS Biol* 2:E234. <https://doi.org/10.1371/journal.pbio.0020234>.
- De Palma M, Montini E, Santoni de Sio FR, Benedicenti F, Gentile A, Medico E, Naldini L. 2005. Promoter trapping reveals significant differences in integration site selection between MLV and HIV vectors in primary hematopoietic cells. *Blood* 105:2307–2315. <https://doi.org/10.1182/blood-2004-03-0798>.
- Schroder AR, Shinn P, Chen H, Berry C, Ecker JR, Bushman F. 2002. HIV-1 integration in the human genome favors active genes and local hotspots. *Cell* 110:521–529. [https://doi.org/10.1016/s0092-8674\(02\)00864-4](https://doi.org/10.1016/s0092-8674(02)00864-4).
- LaFave MC, Varshney GK, Gildea DE, Wolfsberg TG, Baxevasis AD, Burgess SM. 2014. MLV integration site selection is driven by strong enhancers and active promoters. *Nucleic Acids Res* 42:4257–4269. <https://doi.org/10.1093/nar/gkt1399>.
- De Ravin SS, Su L, Theobald N, Choi U, Macpherson JL, Poidinger M, Symonds G, Pond SM, Ferris AL, Hughes SH, Malech HL, Wu X. 2014. Enhancers are major targets for murine leukemia virus vector integration. *J Virol* 88:4504–4513. <https://doi.org/10.1128/JVI.00011-14>.
- Engelman A, Cherepanov P. 2014. Retroviral integrase structure and DNA recombination mechanism. *Microbiol Spectr* 2(6):MDNA3-0024-2014. <https://doi.org/10.1128/microbiolspec.MDNA3-0024-2014>.
- Busschots K, Vercammen J, Emiliani S, Benarous R, Engelborghs Y, Christ F, Debyser Z. 2005. The interaction of LEDGF/p75 with integrase is lentivirus-specific and promotes DNA binding. *J Biol Chem* 280:17841–17847. <https://doi.org/10.1074/jbc.M411681200>.
- Llano M, Saenz DT, Meehan A, Wongthida P, Peretz M, Walker WH, Teo W, Poeschla EM. 2006. An essential role for LEDGF/p75 in HIV integration. *Science* 314:461–464. <https://doi.org/10.1126/science.1132319>.
- Ciuffi A, Llano M, Poeschla E, Hoffmann C, Leipzig J, Shinn P, Ecker JR, Bushman F. 2005. A role for LEDGF/p75 in targeting HIV DNA integration. *Nat Med* 11:1287–1289. <https://doi.org/10.1038/nm1329>.
- Shun M-C, Raghavendra NK, Vandegraaff N, Daigle JE, Hughes S, Kellam P, Cherepanov P, Engelman A. 2007. LEDGF/p75 functions downstream from preintegration complex formation to effect gene-specific HIV-1 integration. *Genes Dev* 21:1767–1778. <https://doi.org/10.1101/gad.1565107>.
- Cherepanov P, Maertens G, Proost P, Devreese B, Van Beeumen J, Engelborghs Y, De Clercq E, Debyser Z. 2003. HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. *J Biol Chem* 278:372–381. <https://doi.org/10.1074/jbc.M209278200>.
- Hendrix J, Gijbsbers R, De Rijck J, Voet A, Hotta J, McNeely M, Hofkens J, Debyser Z, Engelborghs Y. 2011. The transcriptional co-activator LEDGF/p75 displays a dynamic scan-and-lock mechanism for chromatin tethering. *Nucleic Acids Res* 39:1310–1325. <https://doi.org/10.1093/nar/gkq933>.
- Christ F, Shaw S, Demeulemeester J, Desimie BA, Marchand A, Butler S, Smets W, Chaltin P, Westby M, Debyser Z, Pickford C. 2012. Small-molecule inhibitors of the LEDGF/p75 binding site of integrase block HIV replication and modulate integrase multimerization. *Antimicrob Agents Chemother* 56:4365–4374. <https://doi.org/10.1128/AAC.00717-12>.
- Vranckx LS, Demeulemeester J, Saleh S, Boll A, Vansant G, Schrijvers R, Weydert C, Battivelli E, Verdin E, Cereseto A, Christ F, Gijbsbers R, Debyser Z. 2016. LEDGIN-mediated inhibition of integrase-LEDGF/p75 interaction reduces reactivation of residual latent HIV. *EBioMedicine* 8:248–264. <https://doi.org/10.1016/j.ebiom.2016.04.039>.
- Debyser Z, Christ F, De Rijck J, Gijbsbers R. 2015. Host factors for retroviral integration site selection. *Trends Biochem Sci* 40:108–116. <https://doi.org/10.1016/j.tibs.2014.12.001>.
- Studamire B, Goff SP. 2008. Host proteins interacting with the Moloney murine leukemia virus integrase: multiple transcriptional regulators and chromatin binding factors. *Retrovirology* 5:48. <https://doi.org/10.1186/1742-4690-5-48>.
- Gupta SS, Maetzig T, Maertens GN, Sharif A, Rothe M, Weidner-Glunde M, Galla M, Schambach A, Cherepanov P, Schulz TF. 2013. Bromo- and extra-terminal domain chromatin regulators serve as cofactors for murine leukemia virus integration. *J Virol* 87:12721–12736. <https://doi.org/10.1128/JVI.01942-13>.
- De Rijck J, de Kogel C, Demeulemeester J, Vets S, El Ashkar S, Malani N, Bushman FD, Landuyt B, Husson SJ, Busschots K, Gijbsbers R, Debyser Z. 2013. The BET family of proteins targets Moloney murine leukemia virus integration near transcription start sites. *Cell Rep* 5:886–894. <https://doi.org/10.1016/j.celrep.2013.09.040>.
- Sharma A, Larue RC, Plumb MR, Malani N, Male F, Slaughter A, Kessl JJ, Shkriabai N, Coward E, Aiyer SS, Green PL, Wu L, Roth MJ, Bushman FD, Kvaratskhelia M. 2013. BET proteins promote efficient murine leukemia virus integration at transcription start sites. *Proc Natl Acad Sci U S A* 110:12036–12041. <https://doi.org/10.1073/pnas.1307157110>.
- Taniguchi Y. 2016. The bromodomain and extra-terminal domain (BET) family: functional anatomy of BET paralogous proteins. *Int J Mol Sci* 17:1849. <https://doi.org/10.3390/ijms17111849>.
- Morinière J, Rousseaux S, Steuerwald U, Soler-López M, Curtet S, Vitte A-L, Govin J, Gaucher J, Sadoul K, Hart DJ, Krijgsveld J, Khochbin S, Müller CW, Petosa C. 2009. Cooperative binding of two acetylation marks on a histone tail by a single bromodomain. *Nature* 461:664–668. <https://doi.org/10.1038/nature08397>.
- Belkina AC, Denis GV. 2012. BET domain co-regulators in obesity, inflammation and cancer. *Nat Rev Cancer* 12:465–477. <https://doi.org/10.1038/nrc3256>.
- Jang MK, Mochizuki K, Zhou M, Jeong HS, Brady JN, Ozato K. 2005. The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. *Mol Cell* 19:523–534. <https://doi.org/10.1016/j.molcel.2005.06.027>.
- Crowe BL, Larue RC, Yuan C, Hess S, Kvaratskhelia M, Foster MP. 2016. Structure of the Brd4 ET domain bound to a C-terminal motif from gamma-retroviral integrases reveals a conserved mechanism of interaction. *Proc Natl Acad Sci U S A* 113:2086–2091. <https://doi.org/10.1073/pnas.1516813113>.
- Larue RC, Plumb MR, Crowe BL, Shkriabai N, Sharma A, DiFiore J, Malani N, Aiyer SS, Roth MJ, Bushman FD, Foster MP, Kvaratskhelia M. 2014. Bimodal high-affinity association of Brd4 with murine leukemia virus integrase and mononucleosomes. *Nucleic Acids Res* 42:4868–4881. <https://doi.org/10.1093/nar/gku135>.
- El Ashkar S, Van Looveren D, Schenk F, Vranckx LS, Demeulemeester J, De Rijck J, Debyser Z, Modlich U, Gijbsbers R. 2017. Engineering next-generation BET-independent MLV vectors for safer gene therapy. *Mol Ther Nucleic Acids* 7:231–245. <https://doi.org/10.1016/j.omtn.2017.04.002>.
- El Ashkar S, De Rijck J, Demeulemeester J, Vets S, Madlala P, Cermakova K, Debyser Z, Gijbsbers R. 2014. BET-independent MLV-based vectors target away from promoters and regulatory elements. *Mol Ther Nucleic Acids* 3:e179. <https://doi.org/10.1038/mtna.2014.33>.
- Wight DJ, Boucherit VC, Nader M, Allen DJ, Taylor IA, Bishop KN. 2012. The gammaretroviral p12 protein has multiple domains that function during the early stages of replication. *Retrovirology* 9:83. <https://doi.org/10.1186/1742-4690-9-83>.
- Gross L. 1951. "Spontaneous" leukemia developing in C3H mice following inoculation in infancy, with AK-leukemic extracts, or AK-embryos [sic]. *Proc Soc Exp Biol Med* 76:27–32. <https://doi.org/10.3181/00379727-76-18379>.
- MacLachlan NJ, Dubovi EJ, Barthold SW, Swayne DE, Winton JR (ed). 2017. *Fenner's veterinary virology*, 5th ed. Academic Press, San Diego, CA.

32. Low A, Okeoma CM, Lovsin N, de las Heras M, Taylor TH, Peterlin BM, Ross SR, Fan H. 2009. Enhanced replication and pathogenesis of Moloney murine leukemia virus in mice defective in the murine APOBEC3 gene. *Virology* 385:455–463. <https://doi.org/10.1016/j.virol.2008.11.051>.
33. Tsuruyama T, Hiratsuka T, Yamada N. 2017. Hotspots of MLV integration in the hematopoietic tumor genome. *Oncogene* 36:1169–1175. <https://doi.org/10.1038/ncr.2016.285>.
34. Touw IP, Erkeland SJ. 2007. Retroviral insertion mutagenesis in mice as a comparative oncogenomics tool to identify disease genes in human leukemia. *Mol Ther* 15:13–19. <https://doi.org/10.1038/sj.mt.6300040>.
35. Dabrowska MJ, Ejegod D, Lassen LB, Johnsen HE, Wabl M, Pedersen FS, Dybkær K. 2013. Gene expression profiling of murine T-cell lymphoblastic lymphoma identifies deregulation of S-phase initiating genes. *Leuk Res* 37:1383–1390. <https://doi.org/10.1016/j.leukres.2013.04.012>.
36. Sokol M, Wabl M, Ruiz IR, Pedersen FS. 2014. Novel principles of gamma-retroviral insertional transcription activation in murine leukemia virus-induced end-stage tumors. *Retrovirology* 11:36. <https://doi.org/10.1186/1742-4690-11-36>.
37. Knight S, Collins M, Takeuchi Y. 2013. Insertional mutagenesis by retroviral vectors: current concepts and methods of analysis. *Curr Gene Ther* 13:211–227. <https://doi.org/10.2174/1566523211313030006>.
38. Uren AG, Kool J, Berns A, van Lohuizen M. 2005. Retroviral insertional mutagenesis: past, present and future. *Oncogene* 24:7656–7672. <https://doi.org/10.1038/sj.onc.1209043>.
39. Mikkers H, Berns A. 2003. Retroviral insertional mutagenesis: tagging cancer pathways. *Adv Cancer Res* 88:53–99. [https://doi.org/10.1016/S0065-230X\(03\)88304-5](https://doi.org/10.1016/S0065-230X(03)88304-5).
40. Coffin JM, Hughes SH, Varmus H (ed). 1997. *Retroviruses*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
41. Hu WS, Pathak VK. 2000. Design of retroviral vectors and helper cells for gene therapy. *Pharmacol Rev* 52:493–511.
42. Sinn PL, Sauter SL, McCray PB, Jr. 2005. Gene therapy progress and prospects: development of improved lentiviral and retroviral vectors—design, biosafety, and production. *Gene Ther* 12:1089–1098. <https://doi.org/10.1038/sj.gt.3302570>.
43. Persons DA. 2010. Lentiviral vector gene therapy: effective and safe? *Mol Ther* 18:861–862. <https://doi.org/10.1038/mt.2010.70>.
44. Aiuti A, Cassani B, Andolfi G, Mirolo M, Biasco L, Recchia A, Urbinati F, Valacca C, Scaramuzza S, Aker M, Slavin S, Cazzola M, Sartori D, Ambrosi A, Di Serio C, Roncarolo MG, Mavilio F, Bordignon C. 2007. Multilineage hematopoietic reconstitution without clonal selection in ADA-SCID patients treated with stem cell gene therapy. *J Clin Invest* 117:2233–2240. <https://doi.org/10.1172/JCI31666>.
45. Howe SJ, Mansour MR, Schwarzwaelder K, Bartholomae C, Hubank M, Kempinski H, Brugman MH, Pike-Overzet K, Chatters SJ, de Ridder D, Gilmour KC, Adams S, Thornhill SI, Parsley KL, Staal FJT, Gale RE, Linch DC, Bayford J, Brown L, Quayle M, Kinnon C, Ancliff P, Webb DK, Schmidt M, von Kalle C, Gaspar HB, Thrasher AJ. 2008. Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. *J Clin Invest* 118:3143–3150. <https://doi.org/10.1172/JCI35798>.
46. Hacein-Bey-Abina S, Garrigue A, Wang GP, Soulier J, Lim A, Morillon E, Clappier E, Caccavelli L, Delabesse E, Beldjord K, Asnafi V, MacIntyre E, Dal Cortivo L, Radford I, Brousse N, Sigaux F, Moshous D, Hauer J, Borkhardt A, Belohradsky BH, Wintergerst U, Velez MC, Leiva L, Sorensen R, Wulffraat N, Blanche S, Bushman FD, Fischer A, Cavazzana-Calvo M. 2008. Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. *J Clin Invest* 118:3132–3142. <https://doi.org/10.1172/JCI35700>.
47. Stein S, Ott MG, Schultze-Strasser S, Jauch A, Burwinkel B, Kinner A, Schmidt M, Krämer A, Schwäble J, Glimm H, Koehl U, Preiss C, Ball C, Martin H, Göhring G, Schwarzwaelder K, Hofmann W-K, Karakaya K, Tchatchou S, Yang R, Reinecke P, Kühlcke K, Schlegelberger B, Thrasher AJ, Hoelzer D, Seger R, von Kalle C, Grez M. 2010. Genomic instability and myelodysplasia with monosomy 7 consequent to EVI1 activation after gene therapy for chronic granulomatous disease. *Nat Med* 16:198–204. <https://doi.org/10.1038/nm.2088>.
48. Zufferey R, Dull T, Mandel RJ, Bukovsky A, Quiroz D, Naldini L, Trono D. 1998. Self-inactivating lentivirus vector for safe and efficient *in vivo* gene delivery. *J Virol* 72:9873–9880. <https://doi.org/10.1128/JVI.72.12.9873-9880.1998>.
49. Borrenberghs D, Zurnic I, De Wit F, Acke A, Dirix L, Cereseto A, Debysers Z, Hendrix J. 2019. Post-mitotic BET-induced reshaping of integrase quaternary structure supports wild-type MLV integration. *Nucleic Acids Res* 47:1195–1210. <https://doi.org/10.1093/nar/gky1157>.
50. Vansant G, Bruggemans A, Janssens J, Debysers Z. 2020. Block-and-lock strategies to cure HIV infection. *Viruses* 12:84. <https://doi.org/10.3390/v12010084>.
51. Loyola L, Achuthan V, Gilroy K, Borland G, Kilbey A, Mackay N, Bell M, Hay J, Aiyyer S, Fingerma D, Villanueva RA, Cameron E, Kozak CA, Engelman AN, Neil J, Roth MJ. 2019. Disrupting MLV integrase: BET protein interaction biases integration into quiescent chromatin and delays but does not eliminate tumor activation in a MYC/Runx2 mouse model. *PLoS Pathog* 15:e1008154. <https://doi.org/10.1371/journal.ppat.1008154>.
52. Aiyyer S, Rossi P, Malani N, Schneider WM, Chandar A, Bushman FD, Montelione GT, Roth MJ. 2015. Structural and sequencing analysis of local target DNA recognition by MLV integrase. *Nucleic Acids Res* 43:5647–5663. <https://doi.org/10.1093/nar/gkv410>.
53. Shilatifard A. 2006. Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. *Annu Rev Biochem* 75:243–269. <https://doi.org/10.1146/annurev.biochem.75.103004.142422>.
54. Blumenreich MS. 1990. The white blood cell and differential count. In Walker HK, Hall WD, Hurst JW (ed), *Clinical methods: the history, physical, and laboratory examinations*, 3rd ed. Butterworths, Boston, MA.
55. Modlich U, Kustikova OS, Schmidt M, Rudolph C, Meyer J, Li Z, Kamino K, von Neuhoff N, Schlegelberger B, Kuehlcke K, Bunting KD, Schmidt S, Deichmann A, von Kalle C, Fehse B, Baum C. 2005. Leukemias following retroviral transfer of multidrug resistance 1 (MDR1) are driven by combinatorial insertional mutagenesis. *Blood* 105:4235–4246. <https://doi.org/10.1182/blood-2004-11-4535>.
56. Pinkus GS, Pinkus JL. 1991. Myeloperoxidase: a specific marker for myeloid cells in paraffin sections. *Mod Pathol* 4:733–741.
57. Hu ZB, Ma W, Uphoff CC, Metge K, Gignac SM, Drexler HG. 1993. Myeloperoxidase: expression and modulation in a large panel of human leukemia-lymphoma cell lines. *Blood* 82:1599–1607. <https://doi.org/10.1182/blood.V82.5.1599.1599>.
58. Gu Z, Churchman ML, Roberts KG, Moore I, Zhou X, Nakitandwe J, Hagiwara K, Pelletier S, Gingras S, Berns H, Payne-Turner D, Hill A, Iacobucci I, Shi L, Pounds S, Cheng C, Pei D, Qu C, Newman S, Devidas M, Dai Y, Reshmi SC, Gastier-Foster J, Raetz EA, Borowitz MJ, Wood BL, Carroll WL, Zweidler-McKay PA, Rabin KR, Mattano LA, Maloney KW, Rambaldi A, Spinelli O, Radich JP, Minden MD, Rowe JM, Luger S, Litzow MR, Tallman MS, Racevskis J, Zhang Y, Bhatia R, Kohlschmidt J, Mrózek K, Bloomfield CD, Stock W, Kornblau S, Kantarjian HM, Konopleva M, Evans WE, et al. 2019. PAX5-driven subtypes of B-progenitor acute lymphoblastic leukemia. *Nat Genet* 51:296–307. <https://doi.org/10.1038/s41588-018-0315-5>.
59. Liu R, Simonetti FR, Ho YC. 2020. The forces driving clonal expansion of the HIV-1 latent reservoir. *Virol J* 17:4. <https://doi.org/10.1186/s12985-019-1276-8>.
60. Bingham CR, Cook LB, Melamed A. 2014. HTLV-1 clonality in adult T-cell leukaemia and non-malignant HTLV-1 infection. *Semin Cancer Biol* 26:89–98. <https://doi.org/10.1016/j.semcancer.2013.11.003>.
61. Modlich U, Schambach A, Brugman MH, Wicke DC, Knoess S, Li Z, Maetzgi T, Rudolph C, Schlegelberger B, Baum C. 2008. Leukemia induction after a single retroviral vector insertion in Evi1 or Prdm16. *Leukemia* 22:1519–1528. <https://doi.org/10.1038/leu.2008.118>.
62. Yamada K, Tsukahara T, Yoshino K, Kojima K, Agawa H, Yamashita Y, Amano Y, Hatta M, Matsuzaki Y, Kurotori N, Wakui K, Fukushima Y, Osada R, Shiozawa T, Sakashita K, Koike K, Kumaki S, Tanaka N, Takeshita T. 2009. Identification of a high incidence region for retroviral vector integration near exon 1 of the LMO2 locus. *Retrovirology* 6:79. <https://doi.org/10.1186/1742-4690-6-79>.
63. Schrijvers R, Vets S, De Rijck J, Malani N, Bushman FD, Debysers Z, Gijssbers R. 2012. HRP-2 determines HIV-1 integration site selection in LEDGF/p75 depleted cells. *Retrovirology* 9:84. <https://doi.org/10.1186/1742-4690-9-84>.
64. Wang H, Jurado KA, Wu X, Shun M-C, Li X, Ferris AL, Smith SJ, Patel PA, Fuchs JR, Cherepanov P, Kvaratskhelia M, Hughes SH, Engelman A. 2012. HRP2 determines the efficiency and specificity of HIV-1 integration in LEDGF/p75 knockout cells but does not contribute to the antiviral activity of a potent LEDGF/p75-binding site integrase inhibitor. *Nucleic Acids Res* 40:11518–11530. <https://doi.org/10.1093/nar/gks913>.
65. Crawford S, Goff SP. 1984. Mutations in gag proteins P12 and P15 of Moloney murine leukemia virus block early stages of infection. *J Virol* 49:909–917. <https://doi.org/10.1128/JVI.49.3.909-917.1984>.
66. Yuan B, Fassati A, Yueh A, Goff SP. 2002. Characterization of Moloney murine leukemia virus p12 mutants blocked during early events of infection. *J Virol* 76:10801–10810. <https://doi.org/10.1128/jvi.76.21.10801-10810.2002>.
67. Yuan B, Li X, Goff SP. 1999. Mutations altering the Moloney murine leukemia virus p12 Gag protein affect virion production and early events of the virus life cycle. *EMBO J* 18:4700–4710. <https://doi.org/10.1093/emboj/18.17.4700>.

68. Wanaguru M, Barry DJ, Benton DJ, O'Reilly NJ, Bishop KN. 2018. Murine leukemia virus p12 tethers the capsid-containing pre-integration complex to chromatin by binding directly to host nucleosomes in mitosis. *PLoS Pathog* 14:e1007117. <https://doi.org/10.1371/journal.ppat.1007117>.
69. Elis E, Ehrlich M, Prizan-Ravid A, Laham-Karam N, Bacharach E. 2012. p12 tethers the murine leukemia virus pre-integration complex to mitotic chromosomes. *PLoS Pathog* 8:e1003103. <https://doi.org/10.1371/journal.ppat.1003103>.
70. Brzezinski JD, Felkner R, Modi A, Liu M, Roth MJ. 2016. Phosphorylation requirement of murine leukemia virus p12. *J Virol* 90:11208–11219. <https://doi.org/10.1128/JVI.01178-16>.
71. Yueh A, Goff SP. 2003. Phosphorylated serine residues and an arginine-rich domain of the Moloney murine leukemia virus p12 protein are required for early events of viral infection. *J Virol* 77:1820–1829. <https://doi.org/10.1128/jvi.77.3.1820-1829.2003>.
72. Schneider WM, Brzezinski JD, Aiyyer S, Malani N, Gyuricza M, Bushman FD, Roth MJ. 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. <https://doi.org/10.1073/pnas.1221736110>.
73. Voigt P, Reinberg D. 2011. BRD4 jump-starts transcription after mitotic silencing. *Genome Biol* 12:133. <https://doi.org/10.1186/gb-2011-12-11-133>.
74. Wagner BM. 1999. Book reviews: Pathology of the mouse, reference and atlas. R. R. Maronpot, G. A. Boorman and B. W. Gaul, Editors. Vienna, Illinois, Cache River Press, 1999, 699 pp. *Toxicol Pathol* 27:689. <https://doi.org/10.1177/019262339902700611>.
75. Terstappen LW, Loken MR. 1990. Myeloid cell differentiation in normal bone marrow and acute myeloid leukemia assessed by multi-dimensional flow cytometry. *Anal Cell Pathol* 2:229–240.
76. McGrath MS, Ng VL. 1989. Human retrovirus-associated malignancy. *Cancer Treat Res* 47:267–284. [https://doi.org/10.1007/978-1-4613-1599-5\\_11](https://doi.org/10.1007/978-1-4613-1599-5_11).
77. Huser CA, Gilroy KL, de Ridder J, Kilbey A, Borland G, Mackay N, Jenkins A, Bell M, Herzyk P, van der Weyden L, Adams DJ, Rust AG, Cameron E, Neil JC. 2014. Insertional mutagenesis and deep profiling reveals gene hierarchies and a Myc/p53-dependent bottleneck in lymphomagenesis. *PLoS Genet* 10:e1004167. <https://doi.org/10.1371/journal.pgen.1004167>.
78. Jia Y, Chng W-J, Zhou J. 2019. Super-enhancers: critical roles and therapeutic targets in hematologic malignancies. *J Hematol Oncol* 12:77. <https://doi.org/10.1186/s13045-019-0757-y>.
79. Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, Rahl PB, Lee TI, Young RA. 2013. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* 153:307–319. <https://doi.org/10.1016/j.cell.2013.03.035>.
80. Sabari BR, Dall'Agnese A, Bojja A, Klein IA, Coffey EL, Shrinivas K, Abraham BJ, Hannett NM, Zamudio AV, Manteiga JC, Li CH, Guo YE, Day DS, Schuijers J, Vasile E, Malik S, Hnisz D, Lee TI, Cisse II, Roeder RG, Sharp PA, Chakraborty AK, Young RA. 2018. Coactivator condensation at super-enhancers links phase separation and gene control. *Science* 361:eaar3958. <https://doi.org/10.1126/science.aar3958>.
81. Wu X, Li Y, Crise B, Burgess SM. 2003. Transcription start regions in the human genome are favored targets for MLV integration. *Science* 300:1749–1751. <https://doi.org/10.1126/science.1083413>.
82. Li M, Jiang P, Cheng K, Zhang Z, Lan S, Li X, Zhao L, Wang Y, Wang X, Chen J, Ji T, Han B, Zhang J. 2021. Regulation of MYB by distal enhancer elements in human myeloid leukemia. *Cell Death Dis* 12:223. <https://doi.org/10.1038/s41419-021-03515-z>.
83. Zhang L, Liu S-Y, Yang X, Wang Y-Q, Cheng Y-X. 2020. TMEM206 is a potential prognostic marker of hepatocellular carcinoma. *Oncol Lett* 20:174. <https://doi.org/10.3892/ol.2020.12035>.
84. Korthuis PM, Berger G, Bakker B, Rozenveld-Geugien M, Jaques J, de Haan G, Schuringa JJ, Vellenga E, Schepers H. 2015. CITED2-mediated human hematopoietic stem cell maintenance is critical for acute myeloid leukemia. *Leukemia* 29:625–635. <https://doi.org/10.1038/leu.2014.259>.
85. Westin EH, Gallo RC, Arya SK, Eva A, Souza LM, Baluda MA, Aaronson SA, Wong-Staal F. 1982. Differential expression of the amv gene in human hematopoietic cells. *Proc Natl Acad Sci U S A* 79:2194–2198. <https://doi.org/10.1073/pnas.79.7.2194>.
86. Okada M, Tada M, Kanda N, Masuda M, Mizoguchi H, Kazuma M, Wada E, Kubota K, Nomura Y. 1990. c-myc gene analysis in T-cell malignancies with del(6q). *Cancer Genet Cytogenet* 48:229–236. [https://doi.org/10.1016/0165-4608\(90\)90125-t](https://doi.org/10.1016/0165-4608(90)90125-t).
87. Shin S-H, Lee GY, Lee M, Kang J, Shin H-W, Chun Y-S, Park J-W. 2018. Aberrant expression of CITED2 promotes prostate cancer metastasis by activating the nucleolin-AKT pathway. *Nat Commun* 9:4113. <https://doi.org/10.1038/s41467-018-06606-2>.
88. Weng AP, Millholland JM, Yashiro-Ohtani Y, Arcangeli ML, Lau A, Wai C, Del Bianco C, Rodriguez CG, Sai H, Tobias J, Li Y, Wolfe MS, Shachaf C, Felsher D, Blacklow SC, Pear WS, Aster JC. 2006. c-Myc is an important direct target of Notch1 in T-cell acute lymphoblastic leukemia/lymphoma. *Genes Dev* 20:2096–2109. <https://doi.org/10.1101/gad.1450406>.
89. Jundt F, Anagnostopoulos I, Forster R, Mathas S, Stein H, Dorken B. 2002. Activated Notch1 signaling promotes tumor cell proliferation and survival in Hodgkin and anaplastic large cell lymphoma. *Blood* 99:3398–3403. <https://doi.org/10.1182/blood.v99.9.3398>.
90. Kridel R, Meissner B, Rogic S, Boyle M, Telenius J, Li Y, Wolfe MS, Shachaf C, Jenkins C, Cochrane C, Ben-Neriah S, Tan K, Morin RD, Opat S, Sehn LH, Connors JM, Marra MA, Weng AP, Steidl C, Gascoyne RD. 2012. Whole transcriptome sequencing reveals recurrent NOTCH1 mutations in mantle cell lymphoma. *Blood* 119:1963–1971. <https://doi.org/10.1182/blood-2011-11-391474>.
91. Mendrysa SM, Akagi K, Roayaei J, Lien W-H, Copeland NG, Jenkins NA, Eisenman RN. 2010. An integrated genetic-genomic approach for the identification of novel cancer loci in mice sensitized to c-Myc-induced apoptosis. *Genes Cancer* 1:465–479. <https://doi.org/10.1177/1947601910374875>.
92. South AP, Cho RJ, Aster JC. 2012. The double-edged sword of Notch signaling in cancer. *Semin Cell Dev Biol* 23:458–464. <https://doi.org/10.1016/j.semcdb.2012.01.017>.
93. Tsuji H, Ishii-Ohba H, Ukai H, Katsube T, Ogiu T. 2003. Radiation-induced deletions in the 5' end region of Notch1 lead to the formation of truncated proteins and are involved in the development of mouse thymic lymphomas. *Carcinogenesis* 24:1257–1268. <https://doi.org/10.1093/carcin/bgg071>.
94. Roth MJ. 1991. Mutational analysis of the carboxyl terminus of the Moloney murine leukemia virus integration protein. *J Virol* 65:2141–2145. <https://doi.org/10.1128/JVI.65.4.2141-2145.1991>.
95. Blyth K, Terry A, Mackay N, Vaillant F, Bell M, Cameron ER, Neil JC, Stewart M. 2001. Runx2: a novel oncogenic effector revealed by in vivo complementation and retroviral tagging. *Oncogene* 20:295–302. <https://doi.org/10.1038/sj.onc.1204090>.
96. Kozak CA. 2014. Origins of the endogenous and infectious laboratory mouse gammaretroviruses. *Viruses* 7:1–26. <https://doi.org/10.3390/v7010001>.
97. Chiu J, March PE, Lee R, Tillett D. 2004. Site-directed, ligase-independent mutagenesis (SLIM): a single-tube methodology approaching 100% efficiency in 4 h. *Nucleic Acids Res* 32:e174. <https://doi.org/10.1093/nar/gnh172>.
98. Linney E, Davis B, Overhauser J, Chao E, Fan H. 1984. Non-function of a Moloney murine leukaemia virus regulatory sequence in F9 embryonal carcinoma cells. *Nature* 308:470–472. <https://doi.org/10.1038/308470a0>.
99. Overhauser J, Fan H. 1985. Generation of glucocorticoid-responsive Moloney murine leukemia virus by insertion of regulatory sequences from murine mammary tumor virus into the long terminal repeat. *J Virol* 54:133–144. <https://doi.org/10.1128/JVI.54.1.133-144.1985>.
100. Bachelier L, Fan H. 1981. Isolation of recombinant DNA clones carrying complete integrated proviruses of Moloney murine leukemia virus. *J Virol* 37:181–190. <https://doi.org/10.1128/JVI.37.1.181-190.1981>.
101. Pizzato M, Erlwein O, Bonsall D, Kaye S, Muir D, McClure MO. 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. <https://doi.org/10.1016/j.jviromet.2008.10.012>.
102. Stavrou S, Crawford D, Blouch K, Browne EP, Kohli RM, Ross SR. 2014. Different modes of retrovirus restriction by human APOBEC3A and APOBEC3G in vivo. *PLoS Pathog* 10:e1004145. <https://doi.org/10.1371/journal.ppat.1004145>.
103. Sherman E, Nobles C, Berry CC, Six E, Wu Y, Dryga A, Malani N, Male F, Reddy S, Bailey A, Bittinger K, Everett JK, Caccavelli L, Drake MJ, Bates P, Hacey-Bey-Abina S, Cavazzana M, Bushman FD. 2017. INSPIRED: a pipeline for quantitative analysis of sites of new DNA integration in cellular genomes. *Mol Ther Methods Clin Dev* 4:39–49. <https://doi.org/10.1016/j.omtm.2016.11.002>.
104. Gao T, Qian J. 2020. EnhancerAtlas 2.0: an updated resource with enhancer annotation in 586 tissue/cell types across nine species. *Nucleic Acids Res* 48:D58–D64. <https://doi.org/10.1093/nar/gkz980>.