Official journal of the American Society of Gene & Cell Therapy

MTOpen

Evaluating the Safety of Retroviral Vectors Based on Insertional Oncogene Activation and Blocked Differentiation in Cultured Thymocytes

Sheng Zhou[1], Soghra Fatima[1], Zhijun Ma[1], Yong-Dong Wang[2], Taihe Lu[1], Laura J Janke[3], Yang Du[4] and Brian P Sorrentino[1]

[1]*Division of Experimental Hematology, Department of Hematology, Memphis, Tennessee, USA;* [2]*Department of Computational Biology, Memphis, Tennessee, USA;* [3]*Department of Pathology, St. Jude Children's Research Hospital, Memphis, Tennessee, USA;* [4]*Department of Pediatrics, Uniformed Services University of the Health Sciences, Bethesda, Maryland, USA*

Insertional oncogenesis due to retroviral (RV) vector integration has caused recurrent leukemia in multiple gene therapy trials, predominantly due to vector integration effects at the *LMO2* locus. While currently available preclinical safety models have been used for evaluating vector safety, none have predicted or reproduced the recurrent *LMO2* integrations seen in previous X-linked severe combined immunodeficiency (X-SCID) and Wiskott–Aldrich clinical gene therapy trials. We now describe a new assay for assessing vector safety that recapitulates naturally occurring insertions into *Lmo2* and other T-cell proto-oncogenes leading to a preleukemic developmental arrest in primary murine thymocytes cultured *in vitro*. This assay was used to compare the relative oncogenic potential of a variety of gamma-RV and lentiviral vectors and to assess the risk conferred by various transcriptional elements contained in these genomes. Gamma-RV vectors that contained full viral long-terminal repeats were most prone to causing double negative 2 (DN2) arrest and led to repeated cases of Lmo2 pathway activation, while lentiviral vectors containing these same elements were significantly less prone to activate proto-oncogenes or cause DN2 arrest. This work provides a new preclinical assay that is especially relevant for assessing safety in SCID disorders and provides a new tool for designing safer RV vectors.

Received 6 January 2016; accepted 24 February 2016; advance online publication 12 April 2016. doi:10.1038/mt.2016.55

INTRODUCTION

Gene therapy trials for X-linked severe combined immunodeficiency (X-SCID) and Wiskott–Aldrich syndrome (WAS) have been complicated by acute lymphoid leukemia due to insertional activation of cellular proto-oncogenes, mostly notably *LMO2* (refs. [1–3]). These events were not predicted by any preclinical testing or animal models available at that time, despite mouse transplant assays done in multiple different labs. Several cellular assays and mouse transplant models have been subsequently developed to better assess vector safety. An *in vitro* assay based on the ability of vectors to immortalize murine myeloid cells can successfully detect the mutagenic activity of different types of retroviral (RV) vectors.[4,5] This assay has been very useful, but is relatively specific for insertions into the *MDS/EVI-1* locus and therefore may be most relevant for evaluating genotoxicity in myeloid rather than lymphocyte-deficiency cell disorders. Mouse transplant models have been developed that can be used to compare vector-associated genotoxicity, but are relatively expensive, take 6–12 months to complete, and can be prone to a high rate of background tumors.[6,7] None of these newer assays have reproduced the recurrent *LMO2* vector integration events seen in 12 patients enrolled in gene therapy trials for X-SCID and WAS. Clinically relevant *LMO2* vector insertions can be experimentally recreated using Cre-Lox cassette engineering in a human T-cell line. This system has been used to assess the relative ability of RV vectors to transactivate LMO2 expression,[8,9] but this approach is limited to examination of vector effects at predetermined insertion sites and is not based on natural vector integration events. Therefore, investigators and regulatory agencies agree that more effort is needed to further develop preclinical assays for assessing the relative safety of clinical RV vectors.[10]

While the precise mechanism explaining why recurrent LMO2 insertions specifically occur in certain SCID disorders is not well understood, one possibility is that the target cell for these diseases is distinct from that in other blood disorders. For SCID disorders, gene therapy likely targets an early T-lymphocyte progenitor that is generated in the bone marrow and capable of seeding and repopulating the thymus. This may be particularly true when cells are transplanted in the absence of any myeloablative conditioning. This early thymic progenitor likely differs from more primitive hematopoietic stem cells in gene expression profile and accessibility of certain genomic loci to vector integration events. We therefore hypothesized that transduction of primitive undifferentiated thymocytes could be used as the basis for a SCID-specific vector safety assay. Primitive double negative thymocytes ($CD4^-$, $CD8^-$) display several functional characteristics of hematopoietic stem cells including the ability to engraft in irradiated recipients[11,12] and

The first two authors contributed equally.

Correspondence: *Brian P. Sorrentino, St. Jude Children's Research Hospital, Memphis, Tennessee, 38105 USA. Email: brian.sorrentino@stjude.org*

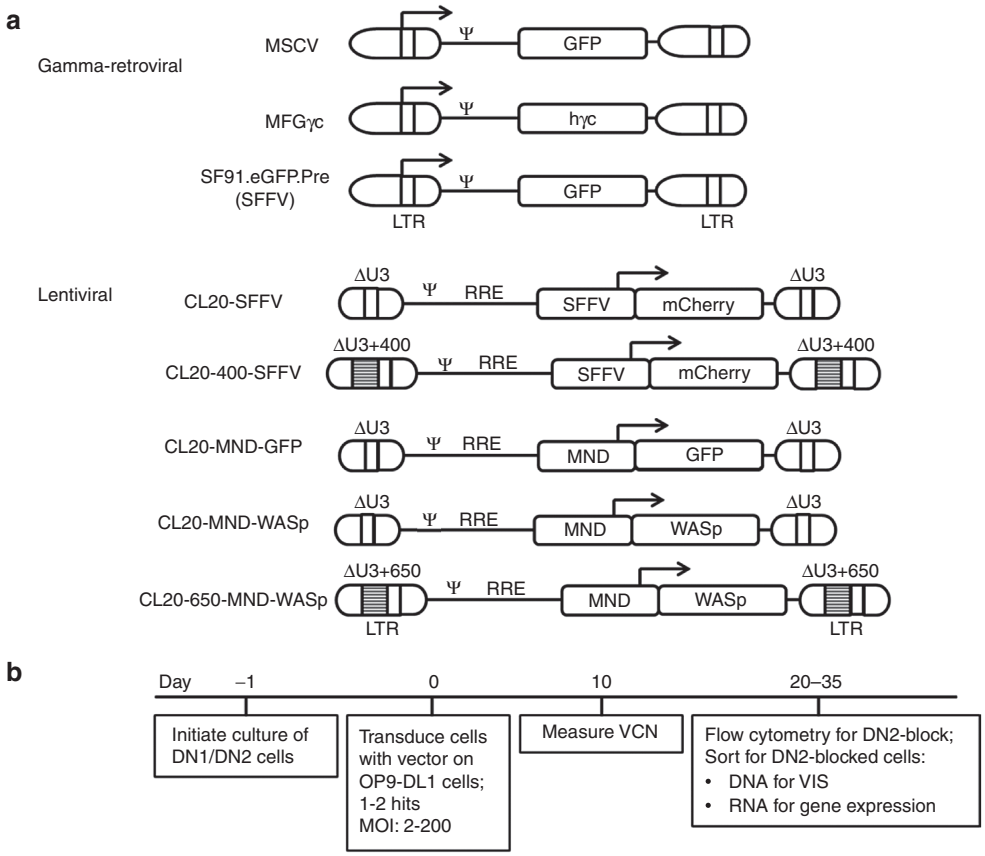


Figure 1 Retroviral vectors and thymocyte culture assay. Schematic representation of gamma-retroviral and lentiviral vectors (**a**) and outline of the DN2 differentiation block assay (**b**). ΔU3, deletion of lentiviral enhancer/promoter in the LTR; DN2, double negative 2; hγc, human IL2RG cDNA; LTR, long terminal repeat; MFG, a hybrid vector containing moloney murine leukemia virus (MoMLV) and myeloproliferative sarcoma virus (MPSV) LTR elements; MND, modified myeloproliferative sarcoma virus enhancer; MSCV, mouse stem cell virus; ψ, viral packaging signal; RRE: lentiviral rev response element; SFFV, spleen focus-forming virus; VCN, vector copy number; WASp, Wiskott–Aldrich Syndrome protein; +400 and +650, 400bp and 650bp chromatin insulator elements from the chicken β–globin locus.

the capacity to reconstitute both myeloid and lymphoid lineages.[13] Furthermore, enforced expression of LMO2 results in a block in thymocyte development at the double negative 2 (DN2) stage, identified by the $CD4^-$ $CD8^-$ $CD25^+$ $CD44^+$ phenotype,[14] ultimately leading to acute T cell leukemia (T-ALL). Cells in this expanding DN2 pool possess enhanced self-renewal and serve as the first clonal event leading to a fully transformed T-ALL phenotype.[11,12,15]

To test the hypothesis that a DN2 thymocyte block could serve as a surrogate endpoint for vector-induced leukemia, we established an assay based on culture of early murine thymic precursor cells on OP9-DL1 stromal cells (see **Supplementary Figure S1**), which allows progression of thymic differentiation through all double negative, double positive, and single positive stages.[16] This system was used to test whether transduction of thymic progenitor cells with murine gamma-RV or HIV-derived lentiviral vectors could induce a DN2 differentiation block and reproduce *LMO2* insertions as has been noted in certain human SCID gene therapy trials. This system was also used to quantitatively assess the relative oncogenic potential of lentiviral versus gamma-RV vectors, different vector-encoded promoters, and chromatin insulator fragments from the chicken β-globin locus.[17]

RESULTS

Vector design and the thymocyte differentiation assay

We generated and tested a variety of murine gamma-RV and human lentiviral vectors designed to encompass a wide spectrum of oncogenic potential (**Figure 1a**). The gamma-RV vectors contained complete long terminal repeats (LTRs) with viral promoters and enhancers and included the MFGγ_c vector that caused T-ALL in an early X-SCID clinical trial[18] and the spleen focus-forming virus (SFFV) vector backbone that caused myelodysplasia in a clinical gene therapy trial for chronic granulomatous disease.[19] We also tested lentiviral vectors based on the CL20 lentiviral backbone[20] containing either an internal SFFV LTR to express mCherry or the MND LTR[21] to drive expression of a therapeutic Wiskott–Aldrich Syndrome protein cDNA. Two of these lentiviral vectors also contained insulator fragments from the chicken β-globin locus[22] designed to shield surrounding cellular genes from vector-induced activation of cellular genes and to prevent transgene silencing. The CL20-650-MND-Wiskott–Aldrich Syndrome protein vector is being developed for use in a clinical gene therapy trial for WAS.[23]

High-titer vector preps were used to transduce murine DN1/DN2 thymic precursor cells that were cocultured on OP9-DL1 stromal cells for up to 35 days (**Figure 1b**). This assay uses a lower

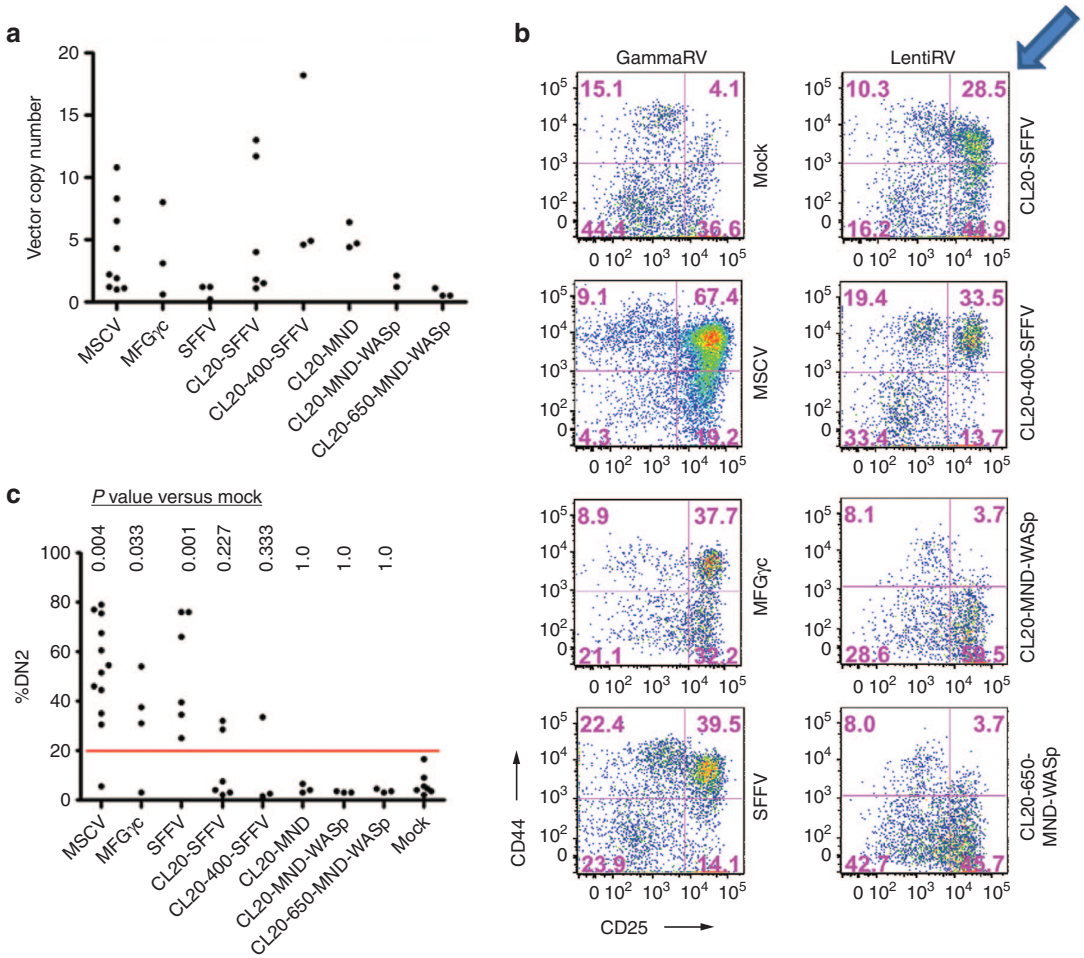


Figure 2 Gamma-retroviral vectors and lentiviral vectors induce DN2 differentiation block in cultured thymic precursor cells. (**a**) Vector copy number (VCN) in samples taken at 10 days after transduction. Each point represents an independent transduction sample. (**b**) Representative examples of flow cytometry analysis for each vector showing the DN1–DN4 subpopulations. The DN2 population is shown in the upper right hand quadrant of each panel and blocked DN2 populations are indicated by the arrows. Accumulations of greater than 20% DN2 cells represent blocked differentiation and were seen in some cases. (**c**) Percentage of DN2-blocked cells in the $CD4^{-}CD8^{-}$ subpopulation with each independent experiment shown as an individual point. The red line represents the normal cutoff defined by mock transduced samples. DN2, double negative 2; MSCV, murine stem cell virus; SFFV, spleen focus-forming virus; WASp, Wiskott–Aldrich Syndrome protein.

concentration of interleukin 7 to increase selective pressure for transformed clones. Transduction efficiency was measured 10 days later using quantitative real-time PCR to quantify the mean vector DNA copy number in the cultured thymocyte pool. The cellular differentiation status was monitored by flow cytometry using cell surface markers CD4, CD8, CD25 and CD44 starting 12–15 days post transduction. If a DN2-blocked subpopulation emerged as greater than 20% of all double negative cells, these cells were then sorted and analyzed for vector insertion site (VIS) using our quantitative shearing linear amplification PCR method.[24] The cut-off at 20% was defined by the mock-transduced populations, which developed normally and never showed greater than 20% DN2 cells.

Comparison of gamma-RV and lentiviral vectors for induction of DN2-blocked populations

Sorted double negative thymocytes were transduced with these vectors in a total of six independent experiments that each included untransduced negative control cells. The vector copy number (VCN) for each transduction pool ranged between 1 and 18 vector genomes (vg) per cell in the population and generally correlated with the multiplicity of infection and the number of vector exposures (**Figure 2a**). Cells were analyzed by flow cytometry for a DN2 differentiation block every 3 days starting 12–15 days after transduction. High-grade DN2 accumulations occurred in most samples that were transduced with any of the gamma-RV vectors (**Figure 2b,c**). While the likelihood of a DN2 block was correlated with higher average VCNs, DN2 blocks occurred at VCNs as low as 1.0 vg/cell with the murine stem cell virus (MSCV) vector, 0.24 vg/cell with the SFFV vector, and 3.1 vg/cell with the MFG-γ_c vector (**Table 1**).

Two out of six transductions with the CL20-SFFV vector and one of three transductions with the CL20-400-SFFV vector resulted in a DN2 block (**Figure 2b,c**) and were associated with relatively high VCNs (13.0, 11.7, and 4.9 vector copies per cell, respectively). The DN2-blocked subpopulations were highly enriched for vector-transduced cells relative to cells in the other gated populations (see **Supplementary Figure S2**). No DN2 blocks occurred with any of the three lentiviral vectors that utilized the MND LTR promoter. This included three experiments with the CL20 MND-GFP vector with

Table 1 Vector-specific DN2 blocks, copy number, and integration sites

Vector/ sample ID	Transduction efficiency (VCN) genomes/cell	Day of DN2 block	DN2 block (% DN2 cells)	Recurrent VIS (other VIS)
Mock				
Mock1			No (4.5%)	
Mock2			No (16.4%)	
Mock3			No (5.5%)	
Mock4			No (2.3%)	
MSCV-GFP				
MSCV1	ND	28	Yes (77%)	1x Lmo2
MSCV2	ND	22	Yes (79%)	1x Lmo2
MSCV3	ND	16	Yes (35%)	1x Mef2c
MSCV4	6.52	20	Yes (54.4%)	4x Mef2c, 1x Lmo2
MSCV5	8.29	20	Yes (51.5%)	2x Mef2c
MSCV6	1.89	20	Yes (60.8%)	1x Mef2c, 1x Lmo2
MSCV7	10.78	28	Yes (82.6%)	1x mef2c, 1x Lmo2
MSCV8	4.33	28	Yes (74.6)	1x Mef2c, 1x Lmo2
MSCV9	2.19	28	Yes (58.4%)	1x Mef2c
MSCV10	1.07		No (18.2%)	
MSCV11	1.03	25	Yes (30.6%)	ND
MSCV12	1.11		No (5.7)	
MSCV13	1.18	25	Yes (44.5%)	ND
SFFV-GFP				
SFFV1	ND	40	Yes (66%)	1x Lmo2
SFFV2	ND	28	Yes (76%)	(Acox1)
SFFV3	ND	22	Yes (76%)	1x Lmo2
SFFV4	1.22	24	Yes (27.8%)	(Prdm16)
SFFV5	1.22	24	Yes (40.8%)	2x Lmo2
SFFV6	0.24	24	Yes (36.1%)	(Mvb12b)
MFG-γc				
MFG-γc1	ND	22	Yes (54.1%)	1x Lmo2
MFG-γc2	3.12	28	Yes (40.4%)	2x Lmo2
MFG-γc3	0.59		No (14.3%)	
MFG-γc4	7.98	28	Yes (42.6%)	2x Lmo2
CL20-SFFV-mCherry				
CL20-SFFV1	13.04	32	Yes (29.7%)	3x Mef2c
CL20-SFFV2	11.71	24	Yes (30.9%)	1x Mef2c
CL20-SFFV3	3.97		No (1.9%)	
CL20-SFFV4	1.83		No (3.9%)	
CL20-SFFV5	1.49		No (3.3%)	
CL20-SFFV6	1.13		No (7.5%)	
CL20-400-SFFV-mCherry				
Cl20-400-SFFV1	4.88	24	Yes (35.4%)	1x Lmo2
Cl20-400-SFFV2	18.21		No (1.5%)	
Cl20-400-SFFV3	4.57		No (2.8%)	
CL20-MND-WASP				
WASP1	1.23		No (3%)	
WASP2	2.15		No (3.1%)	
WASP3	ND		No (3.8%)	
CL20-650-MND-WASP				
650WASP1	0.56		No (3.7%)	
650WASP2	0.54		No (3.3%)	
650WASp3	1.09		No (4.8%)	
CL20-MND-GFP				
CL20-MND1	4.45		No (4.2%)	
CL20-MND2	4.76		No (6.4%)	
CL20-MND3	6.38		No (3.3%)	

DN2, double negative 2; MSCV, murine stem cell virus; SFFV, spleen focus-forming virus; VCN, vector copy number; VIS, vector insertion site; WASp, Wiskott–Aldrich Syndrome protein.

VCNs ranging from 4.5 to 6.4 vg/cell (Table 1). These results demonstrate that the lentiviral SFFV platform may be less prone to causing a DN2 block compared to the SFFV gamma-RV vectors ($P = 0.06$), that the 400-bp insulator does not necessarily protect against the block when an internal SFFV LTR is present in the lentiviral vector, and that the three CL20 MND vectors did not cause any differentiation block.

Recurrent vector insertions in the *Lmo2* and *Mef2c* gene loci in DN2-blocked cells

To determine the genomic locations of the VIS associated with these DN2-blocked cultures, we sorted DN2 populations from 20 independent transductions, and mapped the VISs using our recently developed quantitative shearing linear amplification PCR assay.[24] We identified 741 unique VIS in the DN2-blocked samples and showed recurrent VIS in the *Lmo2* and *Mef2c* loci with both gamma-RV and lentiviral vectors in independent transduction pools (see **Supplementary Table S2**). Overall, *Lmo2* or *Mef2c* VIS occurred in 17 out of the 20 blocked samples from the gamma-RV group and in all 3 blocked samples from the lentiviral group (Figure 3). In a majority of these cases, the shear site counts for the *Lmo2* or *Mef2c* insertions were within the top 10 most frequent insertions identified by quantitative shearing linear amplification PCR (see **Supplementary Table S2**). For instance, *Lmo2* insertions with the SFFV gamma-RV vector were the #1 or 2 most frequent VIS in three out of six cases. One of the two CL20-SFFV-mCherry samples had a dominant *Mef2c* VIS that ranked second in overall VIS frequency and three other *Mef2c* VIS ranked within the top eight most frequent sites. These results show that most of these DN2 populations contained significant proportions of cells with VIS in these two proto-oncogenes.

Vector insertions into the *Lmo2* locus occurred with both gamma-RV and lentiviral vectors and were distributed in three clusters, in the first intron, in a 3-kb window 24-kb upstream of the first exon, and in a 3.5-kb window 60-kb upstream of the first exon (Figure 3a). Fifteen of these insertions occurred with the

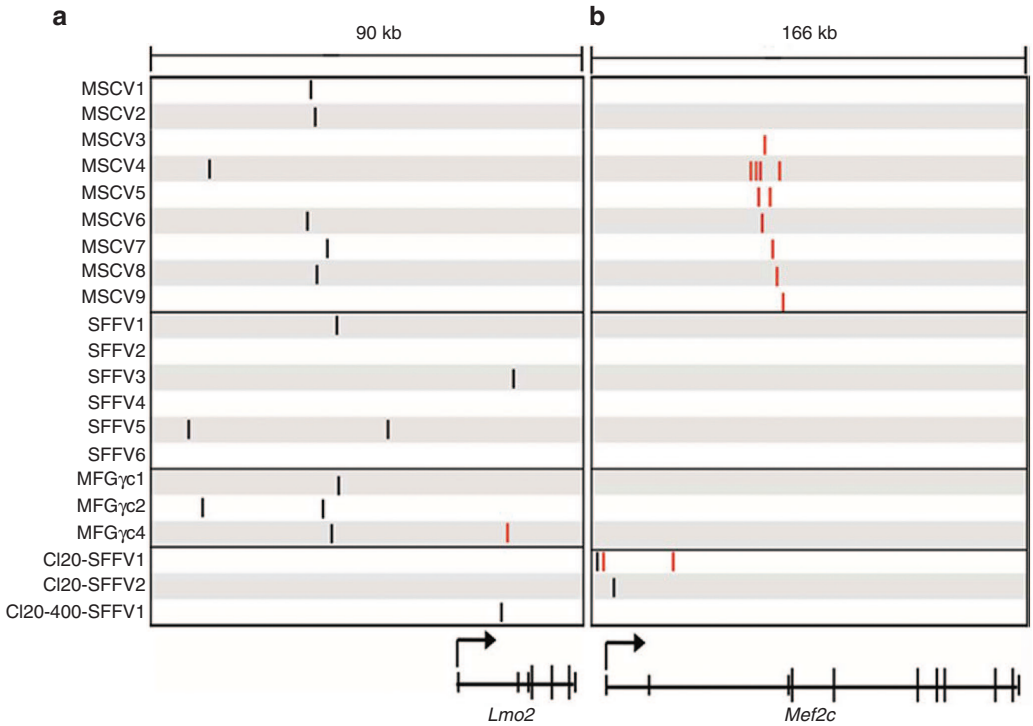


Figure 3 Recurrent vector insertion sites (VIS) identified in DN2-blocked samples. (**a**) VIS within a 90-kb window of the *Lmo2* locus. Each independent experiment is shown in one row with the vector type indicated on the left. Each vertical bar represents a unique VIS identified by sequencing. The black bars indicate the same orientation of vector and the gene, and the red bars indicate the opposite orientation of the vector and the gene. (**b**) VIS in the *Mef2c* loci. A 166-kb window is shown. Vector names are shown on the left. The relative position of *Lmo2* and *Mef2c* genomes are shown at the bottom. DN2, double negative 2; MSCV, murine stem cell virus; SFFV, spleen focus-forming virus.

gamma-RV vectors, the majority of which occurred upstream of the gene as has been seen in gamma-RV clinical trials for X-SCID and WAS. Five independent VIS in *Lmo2* were seen with the MFG-γ_c vector that was used in the French clinical trial.[25] A single *Lmo2* VIS was noted with the CL20-400-SFFV lentiviral vector insertion and occurred in the first intron of Lmo2.

Vector insertions into the *Mef2c* locus were the second most common VIS noted in these studies. *MEF2C* is a well-known T-cell oncogene[26,27] and has been identified as a SCL/TAL1 target[28] in myeloid and B cells. The *Mef2c* locus was recurrently targeted with the MSCV gamma-RV vector and with the lentiviral vectors containing the internal SFFV promoter (**Figure 3b**). All the *Mef2c* insertion sites with the MSCV gamma-RV vector were clustered in a 6-kb window within the second intron and were all in the opposite orientation for gene transcription. It has recently been demonstrated that gamma-RV vectors preferentially integrate into enhancers,[29] suggesting that this clustering may be explained by the potential presence of an enhancer within the *Mef2c* second intron. This possibility is consistent with an accumulation of enhancer-associated chromatin marks in this region previously seen in murine common lymphocyte progenitors (see **Supplementary Figure S3**). Integration sites were also identified with the CL20-SFFV lentiviral vector and were located in a 70-kb window around the first exon (**Figure 3b**). Other potential oncogenic targeting events were noted in other loci including a *Prdm16* intronic insertion with the SFFV gamma-RV vector, *Notch2* and *IL7r* insertions with the MSCV gamma-RV vector, and *Arid1b* with the SFFV lentiviral vector (see **Supplementary Table S2**).

Mef2c and *Lmo2* mRNA expression are upregulated by vector insertions

In order to determine if Mef2c was an upstream activator of Lmo2 in this assay, we measured expression of *Mef2c*, *Lmo2*, and *Hhex* (an Lmo2 target gene)[30,31] mRNAs from sorted DN2-blocked cells harboring *Lmo2* or *Mef2c* insertions. This analysis was performed in all eight samples in which adequate RNA preps were available. In the five samples with vector insertions in the *Lmo2* locus, *Lmo2* mRNA expression was upregulated significantly compared to the trace levels seen in normal DN2-stage thymocytes (**Figure 4**). In populations with *Lmo2* vector insertions, *Mef2c* expression was very low and approximated that seen in normal DN1–2 thymocytes. In three samples with *Mef2c* vector insertions, *Mef2c* expression was significantly upregulated and *Lmo2* mRNA expression was also increased relative to controls. Upregulation of *Hhex* mRNA expression correlated with the *Lmo2* expression levels and presumably reflects activation of the Lmo2 transcriptional program in both *Lmo2* VIS and *Mef2C* VIS-containing cells. These results show that vector insertions into these loci resulted in functionally significant increases in *Mef2c* expression and/or *Lmo2* expression and suggest that upregulation of *Mef2c* results in increased expression of the normal endogenous *Lmo2* gene, consistent with a prior report showing that MEF2C can transactivate LMO2 expression in Jurkat T cells via binding to the two *LMO2* promoters in human Loucy T cells.[26]

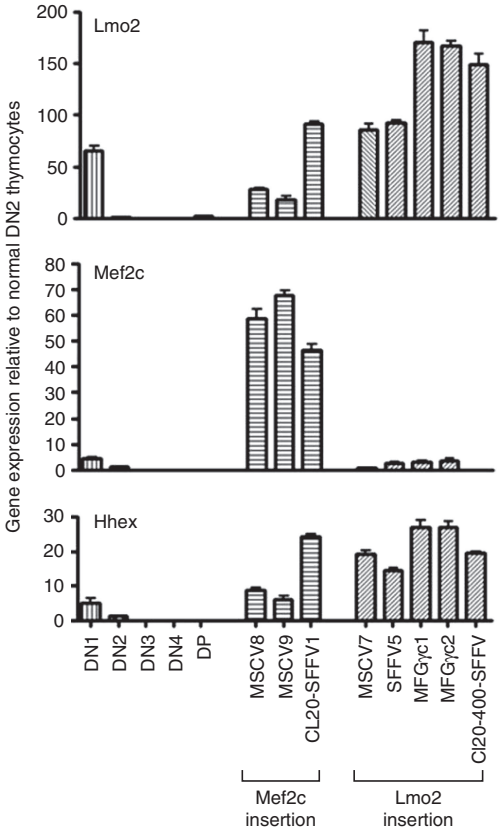


Figure 4 Activation of Lmo2 and Mef2c expression by vector integrations. Total RNA from sorted DN2-blocked cells from each sample was analyzed for expression of *Mef2c*, *Lmo2*, and *Hhex* mRNAs by quantitative real-time-PCR relative to the beta-actin internal control transcript. The sorted DN1, DN2, DN3, DN4, and DP cells from normal thymus were used as controls. The *x*-axis shows which vector populations were analyzed and are grouped by the presence of either Mef2c or Lmo2 insertions. The *y*-axis shows the fold change in gene expression relative to normal DN2 thymocytes. Error bars show the mean standard deviation from technical replicates done on the extracted mRNA samples. DN2, double negative 2; MSCV, murine stem cell virus; SFFV, spleen focus-forming virus.

Acute T-ALL develops in mice transplanted with SFFV gamma-RV-blocked thymocytes

To test whether the presence of a DN2 thymocyte block was a valid surrogate endpoint for vector-induced leukemia, we transplanted 5×10^6 thymocytes transduced with the SFFV gamma-RV vector into sublethally irradiated Rag2$^{-/-}$ $\gamma_c^{-/-}$ recipient mice. At 4 months post-transplant, 2 out of 10 mice in the SFFV vector-transduced group developed a CD3$^+$ T-ALL as manifested by highly elevated white blood cell numbers in the peripheral blood (119×10^3 and 215×10^3 cells/μl), enlarged spleens (0.62 and 0.60 g), circulating leukemic blasts, infiltration of CD3+ leukemic cells in the liver and kidney (**Figure 5a**), and high proportions of GFP+, CD3+ T cells in the bone marrow (**Figure 5b**), spleen, and peripheral blood. None of the 10 mice transplanted with the mock-transduced DN2 cells developed any malignancies. VIS analysis of genomic DNA from the leukemic cells identified a single identical VIS in the second intron of *Mef2c* gene (see **Supplementary Table S2**). These data show that *Mef2c* integration by the gamma-RV SFFV vector caused a classical T-ALL in transplanted mice and verify that the DN2 block serves as a surrogate endpoint for leukemia. The latency and incomplete penetrance seen in the transplant recipients presumably reflect the acquisition of other cooperating genetic mutations, such as loss of tumor suppressor function.

Quantifying the relative oncogenic capacity of different vectors

We defined the oncogenic insertion frequency as the odds that a given vector insertion would occur in a known proto-oncogene in a DN2-blocked population. This frequency was calculated by dividing the total number of unique VIS in cellular proto-oncogenes by the total number of vector insertions seen in all experiments using a given vector (**Figure 6**). The oncogenic insertions scored include mainly the *Lmo2* and *Mef2c* insertions but also other VIS such as the *Prdm16* insertions seen with the SF91 SFFV vector. The total number of VIS in a population was calculated by multiplying the average VCN with the total number of cells that were transduced.

The highest oncogenic insertion frequency was seen with the SFFV gamma-RV vector and was about 3×10^{-5} oncogenic insertions per VIS (**Figure 6a**). The MSCV and MFG gamma-RV vectors were associated with oncogenic insertion frequencies of about $1–2\times10^{-6}$ events per VIS, about fourfold less than that seen with the SFFV gamma-RV vector. In contrast, the oncogenic insertion frequency of the CL20-SFFV vector was approximately 3×10^{-7} events per VIS and may be slightly reduced by the inclusion of the 400-bp insulator (**Figure 6a**). This analysis shows that the SFFV, MSCV, MFGγc, and CL20-SFFV vectors are approximately 41, 11, 10 and 3-fold more likely to insert into T-cell oncogenes and cause a DN2 block relative to the CL20-400-SFFV vector, which had the lowest frequency of detectable events (**Figure 6b**). Statistical analysis using Fisher's exact test showed that the CL20-SFFV vector was significantly less oncogenic than the SFFV gamma-RV vector ($P = 0.002$), despite the presence of the SFFV LTR promoter in each vector, demonstrating that lentiviral vectors with internal RV LTRs are significantly less prone to cause functionally relevant oncogenic insertions. The inclusion of the 400-bp insulator element did not significantly attenuate the modest oncogenic insertion frequency of the CL20-SFFV vector ($P = 0.63$).

Lentiviral vectors containing the MND promoter driving the Wiskott–Aldrich Syndrome protein or GFP cDNA did not cause any detectable DN2 blocks (**Figure 6b**). However, the total number of vector insertions that were present in the MND LV vector pools ranged from only 1.1 to 7.8 million, based on the average VCN and the number of cells present in these pools. In comparison, multiple cases of DN2 blocks occurred with the SFFV gamma-RV vector in pools containing only 1.3 million insertion sites. These results suggest that the MND lentiviral vectors are relatively less prone to cause T-cell transformation than gamma-RV vectors and are consistent with human clinical trial data showing no evidence of vector-induced oncogenesis using a lentiviral vector that expressed the ABCD1 cDNA under control of the MND promoter in patients with adrenoleukodystrophy.[32] It is not clear if the 650-bp insulator provides additional shielding against oncogene activation from the internal MND promoter as has been suggested in other studies.[23,33] In summary, our overall results show that the CL20 vectors containing either the MND or the SFFV promoter are significantly less prone to cause DN2 blocks than

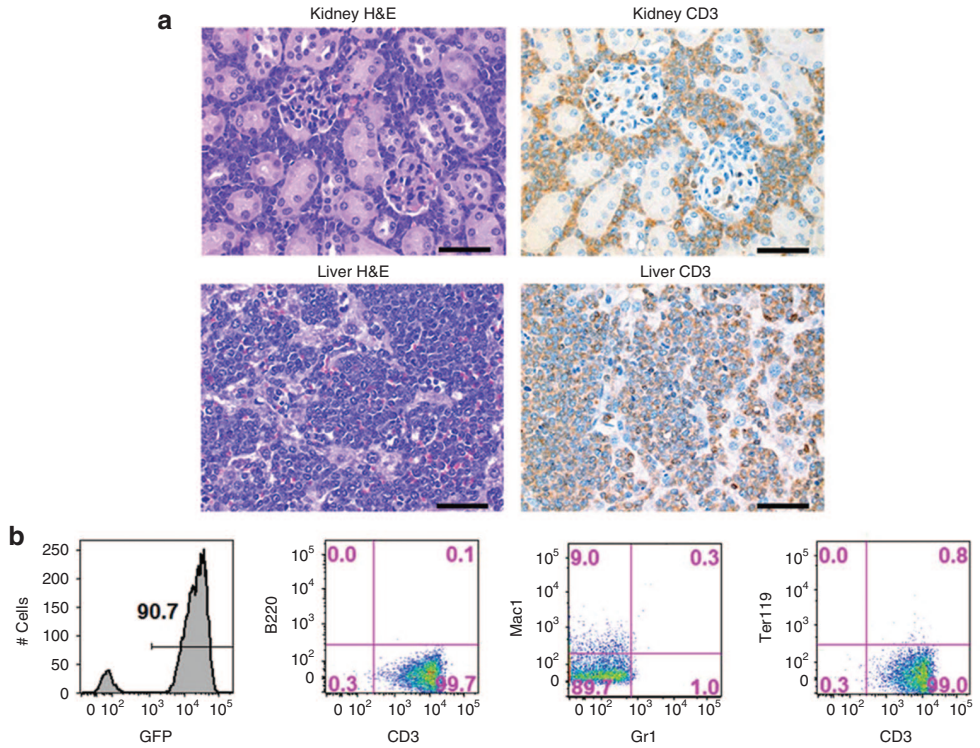
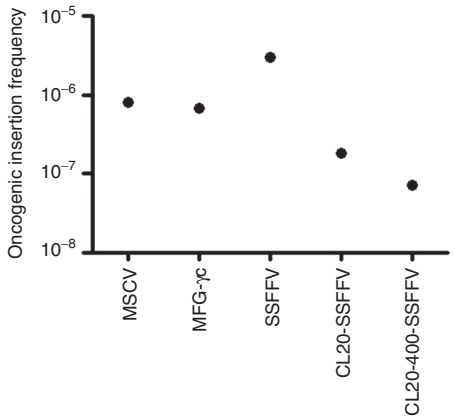


Figure 5 Insertional activation of Mef2c by the SFFV gamma-RV in thymocytes led to acute T-cell lymphocyte leukemia in transplanted mice. (**a**) Immunohistochemistry of kidney and liver sections of a leukemic mouse that arose from transplant of DN2 blocked, SFFV gamma-RV transduced thymocytes. Both H&E and CD3 staining are shown. Bar = 50 µm. (**b**) Flow cytometry analysis of bone marrow cells from one leukemic transplanted mouse showing GFP, CD3, and myeloid marker expression. DN2, double negative 2; SFFV, spleen focus-forming virus.

$$\text{Oncogenic insertion frequency} = \frac{\text{Number of unique proto-oncogene isertions in blocked DN2 cells}}{\text{Total number of vector insertions in all experiments}}$$

a



b

	Relative oncogenic insertion frequency	Total insertions analyzed ($\times 10^6$)
MSCV	11.0	17.5
SFFV	41.3	1.3
MFG-γc	9.5	5.8
Cl20-SFFV	2.9	14.3
Cl20-400-SFFV	1.0	13.8
Cl20-MND-GFP	Not detected	7.8
Cl20-MND-WASp	Not detected	1.7
Cl20-650-MND-WASp	Not detected	1.1

$P = 0.002$, $P = 0.625$, $P = 0.556$

Fisher's exact test used

Figure 6 Relative oncogenic insertion frequency of different vectors. The oncogenic insertion frequency is obtained by dividing the total number of recurrent VIS that occur within a known cellular proto-oncogene by the total number of VIS at day 10 post transduction as calculated by multiplying the vector copy number by the total cell number. (**a**) The average oncogenic insertion frequency of each vector is shown. The number on the *y*-axis indicated the number of integration events needed for one oncogenic VIS to occur that functionally lead to a DN2 block (**b**) This subtable shows the relative oncogenic insertion frequency and total number of insertions analyzed for each vector. The CL20-400-SFFV vector was assigned a frequency of 1 to allow relative comparisons to be made. DN2, double negative 2; MSCV, murine stem cell virus; SFFV, spleen focus-forming virus; VIS, vector insertion site; WASp, Wiskott–Aldrich Syndrome protein.

gamma-RV vectors due to the decreased occurrence of activating integrations in the *Lmo2* and/or *Mef2c* loci.

DISCUSSION

We now describe a thymocyte-based culture assay that reproduces recurrent activating vector insertions in the *Lmo2* locus as well as other T-cell proto-oncogenes and functionally associates these vector insertion events with developmental arrest of thymocytes and induction of acute lymphocytic leukemia in transplanted mice. This assay uses primary thymocytes from wild-type C57BL/6J mice, takes about 35 days to complete, does not require mouse transplantation experiments, and relies on a relevant surrogate endpoint associated with preleukemic transformation to T-ALL.[11,12,15,34] This assay is particularly relevant to vectors used for SCID disorders such as WAS and X-SCID and therefore functionally complements other assays that may be most predictive for myeloid gene therapies such as chronic granulomatous disease[35] and hemoglobinopathies. Certain features are shared between our new assay and those previously described, such as the high degree of oncogenicity associated with the SFFV-based gamma-RV vectors and the decrease in transformation when using lentiviral vectors. What is unique about the thymocyte arrest assay is that it detects oncogenic activity of the MFG-γ_c vector that caused five cases of clinical leukemia in earlier X-SCID gene therapy trials[36] by reproducing the LMO2 activation seen in those cases.

Since overexpression of LMO2 by itself is sufficient to cause DN2 block, it is not surprising that *Lmo2* VIS occurred in many DN2-blocked samples. What was unexpected were the repeated VIS in the *Mef2c* loci in the absence of *Lmo2* VIS. *Mef2c* has been previously identified as a target for oncogenic insertions using replication competent retroviruses in mice.[37,38] This study showed that vector insertions into *Mef2c* caused overexpression of Mef2c, Lmo2, and Hhex suggesting that the leukemogenic effects of Mef2c may be mediated by activation of endogenous Lmo2 expression. Our assay also detected potential oncogenic insertions into other known proto-oncogenes associated with hematopoietic malignancy such as Prdm16 (refs. [35,39,40]) suggesting that the DN2 block assay may detect activation of other pathways.

Lentiviral vectors are being increasingly utilized in clinical trials based on the prediction that they will be safer and more effective for human gene therapy.[41] However, the design parameters that influence lentiviral vector safety are incompletely understood. For instance, it is not known whether inclusion of strong viral promoters in lentiviral vectors can safely be employed, particularly in risk-prone diseases such as X-SCID and WAS, and whether the use of insulators to block enhancer-promoter interactions would necessarily increase safety. Our data show that self-inactivating lentiviral vectors are much less prone to oncogene activation than gamma-RV vectors, even when the highly oncogenic SFFV promoter is contained internally in the vector, consistent with previous reports.[5,6,42] This is likely due to less frequent targeting of particular oncogenes using lentiviral versus gamma-RV vectors and because it is probable that less transactivation occurs when these loci are targeted with lentiviral versus gamma-RV vectors.

Inclusion of the 400-bp chicken beta-globin insulator to the CL20-SFFV vector decreased the mutagenic frequency by 2.9-fold, but this difference did not reach statistical significance. This could be due to the fact that the oncogenic activity of the CL20-SFFV vector was already relatively low and that not enough test replicates were performed to detect added activity with the 400-bp cHS4 insulator. It is also possible that this particular insulator does not provide sufficient enhancer blocking activity in T cells, despite its proven activity in myeloid cells.[17,22,43] It will be important to test alternative newly defined insulators to determine if more effective T cell insulators can be identified for use in SCID gene therapy.

Our vector safety assay is primarily based on the occurrence of a DN2 block in thymocyte development; a functional endpoint representing a known preleukemic change seen with activation of Lmo2 and a variety of other T-cell oncogenes. The other main endpoint is analysis of VISs within the DN2-blocked populations and allows identification of leukemia-initiating events, predominantly transactivation of Lmo2 and Mef2c in this study. We did not sequence VISs in thymocyte cultures that lacked the DN2-block phenotype and it is possible that vector insertions into proto-oncogenes could also be present in these populations, although by inference these insertion events would not be of functional significance. For instance, no DN2 blocks were seen with any of three different lentiviral vectors containing an internal MND promoter, suggesting that MDN-containing CL20 vectors are relatively safe for use in SCID disorders compared to gamma-RV vectors. However, our data cannot distinguish whether this effect was due to a lower frequency of insertions into proto-oncogenes, or whether this is due to less transactivation of proto-oncogenes that are in the vicinity of these insertions. It is clear from previous work that safety-modified lentiviral vectors can be much less prone to deregulate LMO2 than LTR gamma-RV vectors when directly targeted into the locus by cassette exchange,[9] suggesting that lack of transcriptional transactivation explains at least some of the safety advantage seen with the lentiviral vectors in our assay, particularly those lacking the powerful SFFV LTR. These safety considerations have led to the use of weak cellular promoters such as EF1α to drive transgene expression in lentiviral vectors,[9] although these cellular promoters are not always adequate to drive sufficient expression of the vector-encoded transgene. For example, cellular promoters may be suboptimal for obtaining fully physiologic levels of transgene expression in Wiskott–Aldrich syndrome.[23,44] In these cases, the use of a MND LTR promoter/enhancer may be ideal for driving therapeutic expression levels and MND lentiviral vectors have been safely used in a clinical trial for adrenoleukodystrophy.[32] Our thymocyte culture assay provides a useful new tool to further test these newly designed lentiviral vectors and to test newer enhancer-blocking insulator elements that are now available[45,46] and which may further enhance the safety of LV vectors with internal gamma-RV LTRs.

MATERIALS AND METHODS

Thymocyte separation. The female C57BL/6J mice were purchased from Jackson laboratory and used at 4–6 weeks. Female *IL2rg*$^{-/-}$*Rag2*$^{-/-}$ mice were purchased from Taconic Farms (Hudson, NY) and used as transplant recipients at 6–14 weeks. *p19Arf*$^{-/-}$ mice were obtained from Dr. Charles Sherr's lab.[47] All experimental procedures were reviewed and approved by the Institutional Animal Care and Use Committee of St. Jude Children's Research Hospital. Mice were killed and thymi were dissociated prior to staining with CD4-PE, CD8-PE and Ter119-PE antibodies. Followed by incubation with anti-PE microbeads, CD4$^+$, CD8$^+$, CD4$^+$CD8$^+$, and

Ter119$^+$ cells were depleted using MidiMACS columns and magnet. The CD4$^-$CD8$^-$ cells were then labeled with CD44-PE-Cy7 and CD25-APC antibodies and sorted for CD44$^+$CD25$^-$ DN1 and CD44$^+$CD25$^+$ DN2 cells using fluorescence-activated cell sorter.

RV vectors. The gamma-RV vectors MSCV-GFP[48] and SFFV-GFP[4] were obtained from their original sources. The MFG-γc has been generated based on available sequence data as previously described[7] and is identical to the vector used in the X-SCID clinical trial performed in Paris.[25] All vector were generated from stable cell lines based on GPE-86 ecotropic producer cells. The lentiviral vectors were based on the CL20 lentiviral backbone[20] and generated by standard cloning techniques. The CL20-650-MND-Wiskott–Aldrich Syndrome protein vectors have been previously described.[44,49] All lentiviral vectors were transiently produced in 293T cells by cotransfecting with plasmids expressing ecotropic envelope protein, Gag-Pol, and Rev-Tat and were titered on NIH3T3 cells.

Transduction of thymocytes. The stroma cell line expression cell surface notch ligand Delta-like 1 OP9-DL1 was kindly provided by J.C. Zuniga-Pflucker (University of Toronto, Toronto, Canada) and cultured in Alpha-Minimum Essential Medium containing 20% fetal bovine serum, 2 mmol/l L-glutamine and 2 mmol/l sodium pyruvate according to the published method.[16] 5×10^4 OP9-DL1 cells were seeded into each well of 12-well tissue culture plates. Two to four days later, when the OP9-DL1 cells had reached confluence, 5×10^5 freshly purified DN1/DN2 cells were inoculated into each well and prestimulated for overnight in the presence of rmFlt3 (5ng/ml) and rmIL7 (1ng/ml) in a total of 2 ml medium. The next day, 1 ml of supernatant was removed and vectors were added in a total volume of 1 ml, along with 6 µg/ml polybrene and fresh cytokines. The plates were centrifuged at 2,000 rpm for 1 hour at room temperature and then put back into the CO_2 incubator. Two days later, the medium was changed to fresh medium containing 5 ng/ml rmFlt3 and 0.2 ng/ml rmIL7. Every 3–6 days, up to 5×10^5 thymocytes were passaged to new 12-well plates that were preseeded with OP9-DL1 cells. At the second passage, the rmIL7 concentration was restored to 1 ng/ml.

Vector analysis in transduced cells. Ten days post transduction, the cultured thymocytes were sorted for CD45$^+$ expression by flow cytometry to eliminate contaminating OP9-DL1 stroma cells. DNA from sorted cells was extracted and the VCNs in transduced cells were measured by quantitative real-time PCR. For the VIS analysis, 100–400 ng of genomic DNA from sorted DN2-blocked cells was analyzed by the quantitative shearing linear amplification PCR method[24] with the Illumina MiSeq instrument. The processed reads were mapped to Genome Reference Consortium Mouse Build 38 (mm10).

Primer sequences for linear PCR were CCAATCAGTTCGCTTCTC (MSCV and MFG vector), CTGCTTCTCGCTTCTGTTC (SFFV vector), and AGTAGTGTGTGCCCGTCTGT (CL20 lentiviral vectors).

Primer sequences for the final nested-PCR were AATGATACGGCG ACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCG ATCTGCTGTTTGCATCCGAATC (MSCV vector), AATGATACGG CGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCC GATCTGTGGTCTCGCTGTTCCTT (MFG and SFFV vector), and A ATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC GCTCTTCCGATCTGATCCCTCAGACCCTTTTAGT (CL20 vectors).

Total RNAs were extracted from sorted DN2-blocked cells, fresh DN1, DN2, DN3, DN4, and DP subpopulations from wild-type thymus. The Lmo2, Mef2c, and Hhex mRNA levels were measured using the qRT-PCR method using beta-actin mRNA as an internal control for each sample.

Transplant of transduced DN2 cells in recipient mice. DN1/DN2 thymocytes from wild-type mice were transduced with the SFFV-mCherry gamma-RV vector and cultured on OP9-DL1 stroma cells for 21 days. 5×10^6 cultured thymocytes were injected into each female *IL2rg*$^{-/-}$*Rag2*$^{-/-}$ recipient mouse via the lateral tail vein in a total volume of 0.5 ml PBS containing 2% fetal bovine serum. The recipient mice were irradiated with 600 rad in a Cesium137 irradiator 2 hours before injection.

Statistics. Fisher's exact test was used to calculate the significance of the differences seen in DN2 block frequencies seen with the different vectors. *P* values were obtained for a two-tailed and one-tailed tests.

Vendor sources. The sources for key specific reagents are shown in the supplementary material (see **Supplementary Table S1**).

SUPPLEMENTARY MATERIAL

Figure S1. Thymocyte culture assay for insertional mutagenesis of retroviral vectors.
Figure S2. DN2-blocked cells are highly enriched for vector-transduced cells.
Figure S3. Bioinformatic analysis of enhancer chromatin marks in the Mef2c vector integration hotspot region.
Table S1. List of key reagents.
Table S2. Vector insertion sites in blocked DN2 thymocytes.

ACKNOWLEDGMENTS

We thank Dana Roeber and Scott Olsen for genome sequencing, flow cytometry core for cell sorting and flowcytometry analysis, Juan Carlos Zúñiga-Pflücker for providing the OP9-DL1 cells, and Christopher Baum for providing the SF91.eGFP.PRE vector. This work was supported by the National Heart, Lung, and Blood Institute (grant P01HL 53749), Cancer Center (support grant P30 CA 21765), the Assisi Foundation of Memphis, and the ALSAC.

REFERENCES

1. Hacein-Bey-Abina, S, Von Kalle, C, Schmidt, M, McCormack, MP, Wulffraat, N, Leboulch, P *et al.* (2003). LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science* **302**: 415–419.
2. Howe, SJ, Mansour, MR, Schwarzwaelder, K, Bartholomae, C, Hubank, M, Kempski, H *et al.* (2008). Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. *J Clin Invest* **118**: 3143–3150.
3. Braun, CJ, Boztug, K, Paruzynski, A, Witzel, M, Schwarzer, A, Rothe, M *et al.* (2014). Gene therapy for Wiskott–Aldrich syndrome—long-term efficacy and genotoxicity. *Sci Transl Med* **6**: 227ra33.
4. Modlich, U, Bohne, J, Schmidt, M, von Kalle, C, Knöss, S, Schambach, A *et al.* (2006). Cell-culture assays reveal the importance of retroviral vector design for insertional genotoxicity. *Blood* **108**: 2545–2553.
5. Modlich, U, Navarro, S, Zychlinski, D, Maetzig, T, Knoess, S, Brugman, MH *et al.* (2009). Insertional transformation of hematopoietic cells by self-inactivating lentiviral and gammaretroviral vectors. *Mol Ther* **17**: 1919–1928.
6. Montini, E, Cesana, D, Schmidt, M, Sanvito, F, Ponzoni, M, Bartholomae, C *et al.* (2006). Hematopoietic stem cell gene transfer in a tumor-prone mouse model uncovers low genotoxicity of lentiviral vector integration. *Nat Biotechnol* **24**: 687–696.
7. Zhou, S, Ma, Z, Lu, T, Janke, L, Gray, JT and Sorrentino, BP (2013). Mouse transplant models for evaluating the oncogenic risk of a self-inactivating XSCID lentiviral vector. *PLoS One* **8**: e62333.
8. Ryu, BY, Evans-Galea, MV, Gray, JT, Bodine, DM, Persons, DA and Nienhuis, AW (2008). An experimental system for the evaluation of retroviral vector design to diminish the risk for proto-oncogene activation. *Blood* **111**: 1866–1875.
9. Zhou, S, Mody, D, DeRavin, SS, Hauer, J, Lu, T, Ma, Z *et al.* (2010). A self-inactivating lentiviral vector for SCID-X1 gene therapy that does not activate LMO2 expression in human T cells. *Blood* **116**: 900–908.
10. Kiem, HP, Baum, C, Bushman, FD, Byrne, BJ, Carter, BJ, Cavagnaro, J *et al.* (2014). Charting a clear path: the ASGCT Standardized Pathways Conference. *Mol Ther* **22**: 1235–1238.
11. Treanor, LM, Volanakis, EJ, Zhou, S, Lu, T, Sherr, CJ and Sorrentino, BP (2011). Functional interactions between Lmo2, the Arf tumor suppressor, and Notch1 in murine T-cell malignancies. *Blood* **117**: 5453–5462.
12. McCormack, MP, Young, LF, Vasudevan, S, de Graaf, CA, Codrington, R, Rabbitts, TH *et al.* (2010). The Lmo2 oncogene initiates leukemia in mice by inducing thymocyte self-renewal. *Science* **327**: 879–883.
13. Bell, JJ and Bhandoola, A (2008). The earliest thymic progenitors for T cells possess myeloid lineage potential. *Nature* **452**: 764–767.
14. Larson, RC, Osada, H, Larson, TA, Lavenir, I and Rabbitts, TH (1995). The oncogenic LIM protein Rbtn2 causes thymic developmental aberrations that precede malignancy in transgenic mice. *Oncogene* **11**: 853–862.
15. Treanor, LM, Zhou, S, Janke, L, Churchman, ML, Ma, Z, Lu, T *et al.* (2014). Interleukin-7 receptor mutants initiate early T cell precursor leukemia in murine thymocyte progenitors with multipotent potential. *J Exp Med* **211**: 701–713.
16. Schmitt, TM and Zúñiga-Pflücker, JC (2002). Induction of T cell development from hematopoietic progenitor cells by delta-like-1 *in vitro*. *Immunity* **17**: 749–756.
17. Aker, M, Tubb, J, Groth, AC, Bukovsky, AA, Bell, AC, Felsenfeld, G *et al.* (2007). Extended core sequences from the cHS4 insulator are necessary for protecting retroviral vectors from silencing position effects. *Hum Gene Ther* **18**: 333–343.

18. Hacein-Bey-Abina, S, Le Deist, F, Carlier, F, Bouneaud, C, Hue, C, De Villartay, JP *et al.* (2002). Sustained correction of X-linked severe combined immunodeficiency by ex vivo gene therapy. *N Engl J Med* **346**: 1185–1193.
19. Stein, S, Ott, MG, Schultze-Strasser, S, Jauch, A, Burwinkel, B, Kinner, A *et al.* (2010). Genomic instability and myelodysplasia with monosomy 7 consequent to EVI1 activation after gene therapy for chronic granulomatous disease. *Nat Med* **16**: 198–204.
20. Hanawa, H, Kelly, PF, Nathwani, AC, Persons, DA, Vandergriff, JA, Hargrove, P *et al.* (2002). Comparison of various envelope proteins for their ability to pseudotype lentiviral vectors and transduce primitive hematopoietic cells from human blood. *Mol Ther* **5**: 242–251.
21. Challita, PM, Skelton, D, el-Khoueiry, A, Yu, XJ, Weinberg, K and Kohn, DB (1995). Multiple modifications in cis elements of the long terminal repeat of retroviral vectors lead to increased expression and decreased DNA methylation in embryonic carcinoma cells. *J Virol* **69**: 748–755.
22. Arumugam, PI, Urbinati, F, Velu, CS, Higashimoto, T, Grimes, HL and Malik, P (2009). The 3′ region of the chicken hypersensitive site-4 insulator has properties similar to its core and is required for full insulator activity. *PLoS One* **4**: e6995.
23. Koldej, RM, Carney, G, Wielgosz, MM, Zhou, S, Zhan, J, Sorrentino, BP *et al.* (2013). Comparison of insulators and promoters for expression of the Wiskott–Aldrich syndrome protein using lentiviral vectors. *Hum Gene Ther Clin Dev* **24**: 77–85.
24. Zhou, S, Bonner, MA, Wang, YD, Rapp, S, De Ravin, SS, Malech, HL *et al.* (2015). Quantitative shearing linear amplification polymerase chain reaction: an improved method for quantifying lentiviral vector insertion sites in transplanted hematopoietic cell systems. *Hum Gene Ther Methods* **26**: 4–12.
25. Cavazzana-Calvo, M, Hacein-Bey, S, de Saint Basile, G, Gross, F, Yvon, E, Nusbaum, P *et al.* (2000). Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science* **288**: 669–672.
26. Homminga, I, Pieters, R, Langerak, AW, de Rooi, JJ, Stubbs, A, Verstegen, M *et al.* (2011). Integrated transcript and genome analyses reveal NKX2-1 and MEF2C as potential oncogenes in T cell acute lymphoblastic leukemia. *Cancer Cell* **19**: 484–497.
27. Zuurbier, L, Gutierrez, A, Mullighan, CG, Canté-Barrett, K, Gevaert, AO, de Rooi, J *et al.* (2014). Immature MEF2C-dysregulated T-cell leukemia patients have an early T-cell precursor acute lymphoblastic leukemia gene signature and typically have non-rearranged T-cell receptors. *Haematologica* **99**: 94–102.
28. Gekas, C, Rhodes, KE, Gereige, LM, Helgadottir, H, Ferrari, R, Kurdistani, SK *et al.* (2009). Mef2C is a lineage-restricted target of Scl/Tal1 and regulates megakaryopoiesis and B-cell homeostasis. *Blood* **113**: 3461–3471.
29. De Ravin, SS, Su, L, Theobald, N, Choi, U, Macpherson, JL, Poidinger, M *et al.* (2014). Enhancers are major targets for murine leukemia virus vector integration. *J Virol* **88**: 4504–4513.
30. Smith, S, Tripathi, R, Goodings, C, Cleveland, S, Mathias, E, Hardaway, JA *et al.* (2014). LIM domain only-2 (LMO2) induces T-cell leukemia by two distinct pathways. *PLoS One* **9**: e85883.
31. Goodings, C, Smith, E, Mathias, E, Elliott, N, Cleveland, SM, Tripathi, RM *et al.* (2015). Hhex is required at multiple stages of adult hematopoietic stem and progenitor cell differentiation. *Stem Cells* **33**: 2628–2641.
32. Cartier, N, Hacein-Bey-Abina, S, Bartholomae, CC, Veres, G, Schmidt, M, Kutschera, I *et al.* (2009). Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy. *Science* **326**: 818–823.
33. Evans-Galea, MV, Wielgosz, MM, Hanawa, H, Srivastava, DK and Nienhuis, AW (2007). Suppression of clonal dominance in cultured human lymphoid cells by addition of the cHS4 insulator to a lentiviral vector. *Mol Ther* **15**: 801–809.
34. Curtis, DJ and McCormack, MP (2010). The molecular basis of Lmo2-induced T-cell acute lymphoblastic leukemia. *Clin Cancer Res* **16**: 5618–5623.
35. Ott, MG, Schmidt, M, Schwarzwaelder, K, Stein, S, Siler, U, Koehl, U *et al.* (2006). Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1. *Nat Med* **12**: 401–409.
36. McCormack, MP and Rabbitts, TH (2004). Activation of the T-cell oncogene LMO2 after gene therapy for X-linked severe combined immunodeficiency. *N Engl J Med* **350**: 913–922.
37. Du, Y, Jenkins, NA and Copeland, NG (2005). Insertional mutagenesis identifies genes that promote the immortalization of primary bone marrow progenitor cells. *Blood* **106**: 3932–3939.
38. Du, Y, Spence, SE, Jenkins, NA and Copeland, NG (2005). Cooperating cancer-gene identification through oncogenic-retrovirus-induced insertional mutagenesis. *Blood* **106**: 2498–2505.
39. Modlich, U, Schambach, A, Brugman, MH, Wicke, DC, Knoess, S, Li, Z *et al.* (2008). Leukemia induction after a single retroviral vector insertion in Evi1 or Prdm16. *Leukemia* **22**: 1519–1528.
40. Yu, H, Neale, G, Zhang, H, Lee, HM, Ma, Z, Zhou, S *et al.* (2014). Downregulation of Prdm16 mRNA is a specific antileukemic mechanism during HOXB4-mediated HSC expansion in vivo. *Blood* **124**: 1737–1747.
41. Naldini, L (2015). Gene therapy returns to centre stage. *Nature* **526**: 351–360.
42. Montini, E, Cesana, D, Schmidt, M, Sanvito, F, Bartholomae, CC, Ranzani, M *et al.* (2009). The genotoxic potential of retroviral vectors is strongly modulated by vector design and integration site selection in a mouse model of HSC gene therapy. *J Clin Invest* **119**: 964–975.
43. Arumugam, PI, Higashimoto, T, Urbinati, F, Modlich, U, Nestheide, S, Xia, P *et al.* (2009). Genotoxic potential of lineage-specific lentivirus vectors carrying the beta-globin locus control region. *Mol Ther* **17**: 1929–1937.
44. Astrakhan, A, Sather, BD, Ryu, BY, Khim, S, Singh, S, Humblet-Baron, S *et al.* (2012). Ubiquitous high-level gene expression in hematopoietic lineages provides effective lentiviral gene therapy of murine Wiskott–Aldrich syndrome. *Blood* **119**: 4395–4407.
45. Liu, M, Maurano, MT, Wang, H, Qi, H, Song, CZ, Navas, PA *et al.* (2015). Genomic discovery of potent chromatin insulators for human gene therapy. *Nat Biotechnol* **33**: 198–203.
46. Groth, AC, Liu, M, Wang, H, Lovelett, E and Emery, DW (2013). Identification and characterization of enhancer-blocking insulators to reduce retroviral vector genotoxicity. *PLoS One* **8**: e76528.
47. Kamijo, T, Bodner, S, van de Kamp, E, Randle, DH and Sherr, CJ (1999). Tumor spectrum in ARF-deficient mice. *Cancer Res* **59**: 2217–2222.
48. Persons, DA, Allay, JA, Allay, ER, Ashmun, RA, Orlic, D, Jane, SM *et al.* (1999). Enforced expression of the GATA-2 transcription factor blocks normal hematopoiesis. *Blood* **93**: 488–499.
49. Wielgosz, MM, Kim, YS, Carney, GG, Zhan, J, Reddivari, M, Coop, T *et al.* (2015). Generation of a lentiviral vector producer cell clone for human Wiskott–Aldrich syndrome gene therapy. *Mol Ther Methods Clin Dev* **2**: 14063.