Supplementary Materials and Methods

Investigational Agents

Self-activating lentiviral vectors encoding either GFP or WASp were utilized in these studies. A variety of promoters were tested in the mouse model as well as in cultured human cells using one or the other of the investigational agents. Vectors having insulator elements were also tested in these studies.

Vector assembly

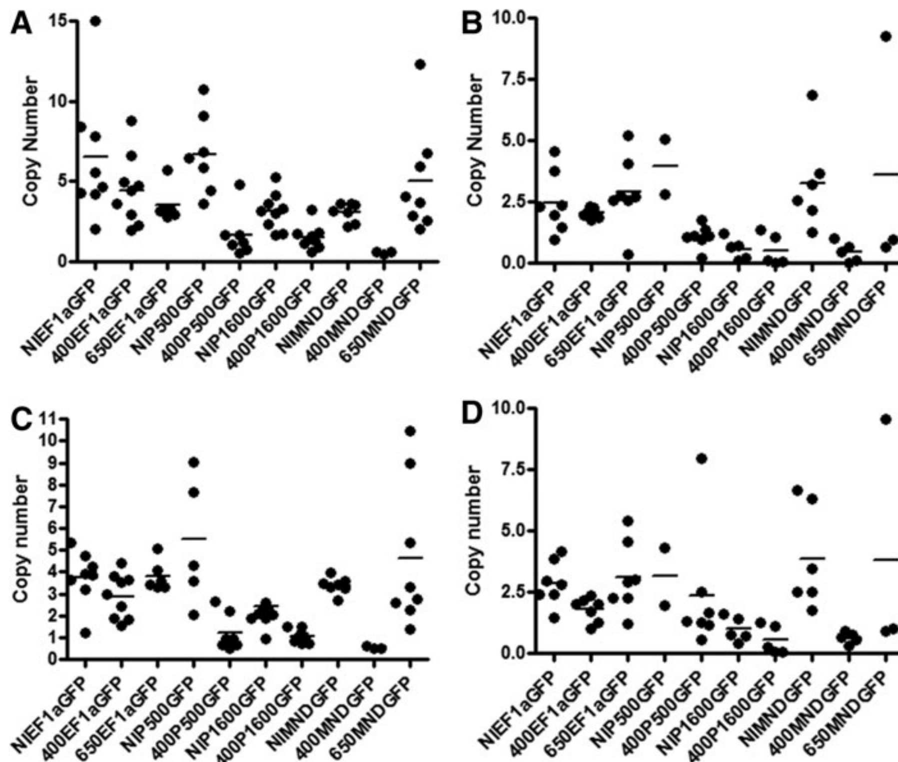
All vectors used in this study were derived from the pCL20cw series of lentiviral vectors (Hanawa *et al*., 2002, 2009) and were constructed using standard DNA subcloning techniques (sequences available on request). The MND promoter used in this study is a composite sequence of the pMBC-2T plasmid (GenBank: x777501, nt112–421), followed by the CMV TATAA box and TSS (sequence TCAGATCG). The EF1α promoter used in this study is identical to the first 272b of the *Homo sapiens* EF1α mRNA cds sequence (GenBank: EF362804.1). The TSS of the EF1α promoter is 24 b downstream of the EF1α TATAA box, beginning with CTTTTTCG. The MND promoter, originally described by Challita *et al*. (1995), was obtained from David Rawlings. We used the short variant of the EF1α promoter as it does not include intron sequences. Insulator elements were added to the LTRs as described in more detail in the section Summary of Data.

Tissue culture

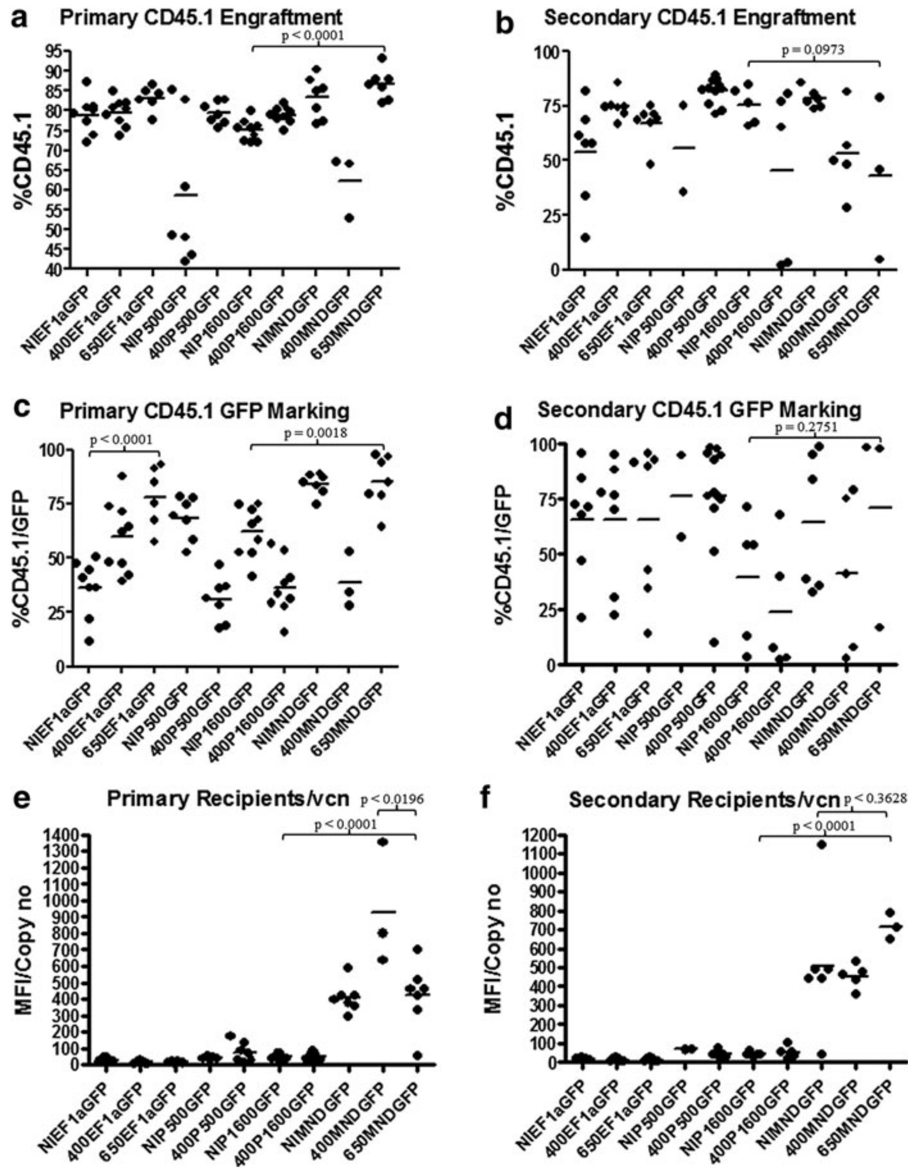
The 293T and Hela cells were maintained in Dulbecco's modified Eagle's medium (Cellgro, Manassas, VA), 10% fetal bovine serum (FBS; Thermo Fisher Scientific, Waltham, MA), 2 m*M* L-Glutamine (Gibco, Grand Island, NY), and 1× penicillin/streptomycin (Gibco) (referred to hereafter as D10). EVBWAS$^+$ and EBVWAS5-2 cells were kindly provided by Dr. Mary Ellen Conley and maintained in RPMI (Lonza, Basel, Switzerland), 15% Foundation FBS (Gemini Bioproducts, West Sacramento, CA), 2 m*M* L-Glutamine, 1× penicillin/streptomycin, 55 μM β-ME (Sigma-Aldrich, St. Louis, MO), and 20 μg/ml Ciprofloxacin (Sicor, Irvine, CA).

Vector production

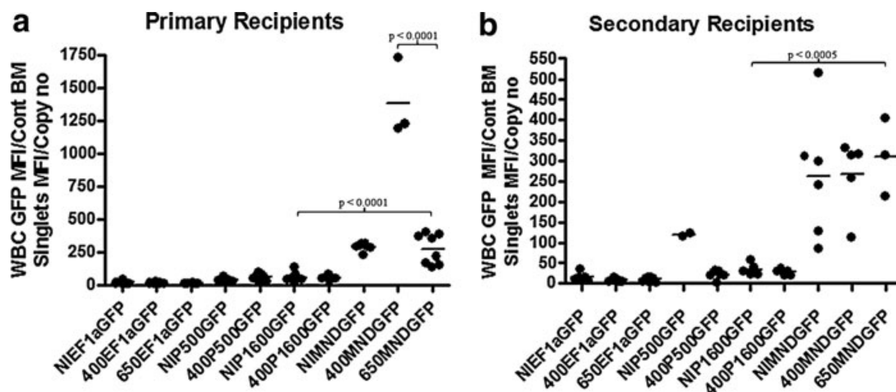
The 293T cells were plated at 3.5×10^6 cells per 10 cm dish in 10 ml of D10 24 hr before $CaCl_2$-mediated transfection with 6 μg pCAGK-GP1.1R, 2 μg pCAGA-RtR2, 2 μg pCAG-VSVg (Hanawa *et al*., 2002), and 10 μg of vector variant. About 24 hr later, the cells were rinsed with 1× PBS and 10 ml D10 was added. After a further 24 hr, the virus was collected, clarified via filtration through a 0.2 μm cellulose



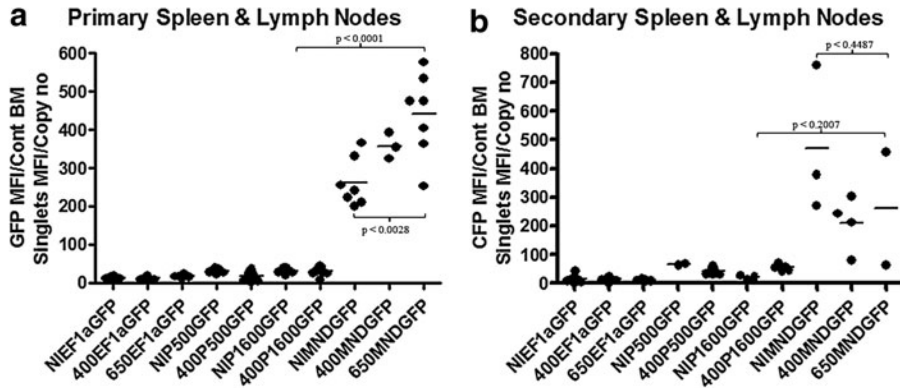
SUPPLEMENTARY FIG. S1. Copy number as determined by qRT-PCR. **(A)** Copy number in the BM of primary recipients. **(B)** Copy number in the BM of secondary recipients. **(C)** Copy number in the WBC of primary recipients. **(D)** Copy number in the WBC of secondary recipients. BM, bone marrow; qRT-PCR, quantitative real-time polymerase chain reaction; WBC, white blood cell.



SUPPLEMENTARY FIG. S2. Engraftment and GFP expression in the BM. **(a)** CD45.1 engraftment in primary recipient. **(b)** CD45.1 engraftment in secondary recipients. **(c)** CD45.1/GFP marking in primary recipients. **(d)** CD45.1/GFP marking in secondary recipients. **(e)** GFP expression per vector copy in primary recipients. **(f)** GFP expression per copy in secondary recipients. NI, non-insulated; 400, 400 bp insulated; 650, 650 bp insulated.

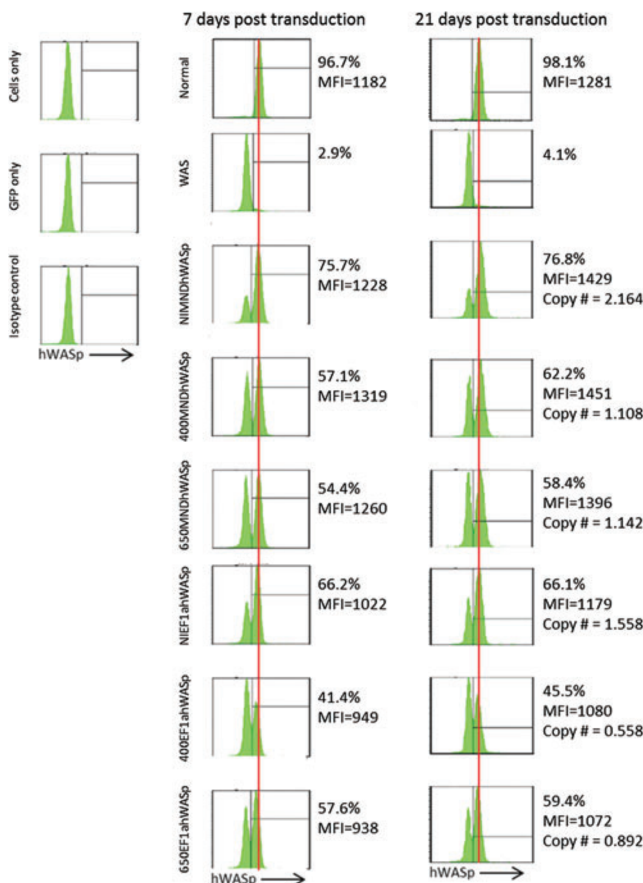


SUPPLEMENTARY FIG. S3. GFP expression in WBCs. **(a)** GFP/copy number in primary recipients. **(b)** GFP/copy in secondary recipients.



SUPPLEMENTARY FIG. S4. Expression of GFP in peripheral tissues. **(a)** Expression per vector copy in the spleen and lymph nodes of primary recipients. **(b)** Expression per vector copy in the spleen and lymph nodes of secondary recipients.

acetate filter, and, if required, concentrated via ultracentrifugation in an SW28 rotor at 28,000 rpm at 4°C for 90 min. The resulting pellet was resuspended in 1/60th the volume of Stemspan (Stemcell Technologies, Vancouver, Canada) at 4°C overnight. The virus was aliquoted and stored at −80°C



SUPPLEMENTARY FIG. S5. hWASp expression from NIMNDhWASp, 400MNDhWASp, 650MNDhWASp, N1EF1αhWASp, 400EF1αhWASp, and N1EF1αhWASp in patient peripheral blood mononuclear cells. Cells were prestimulated with T cell activation beads and transduced with virus at an MOI of 40. hWASp expression was determined at 7 and 21 days by intracellular staining for hWASp and flow cytometry.

and subsequently titered on Hela cells as previously described (Hanawa *et al.*, 2002).

Assays

Genomic DNA analysis for proviral copy number

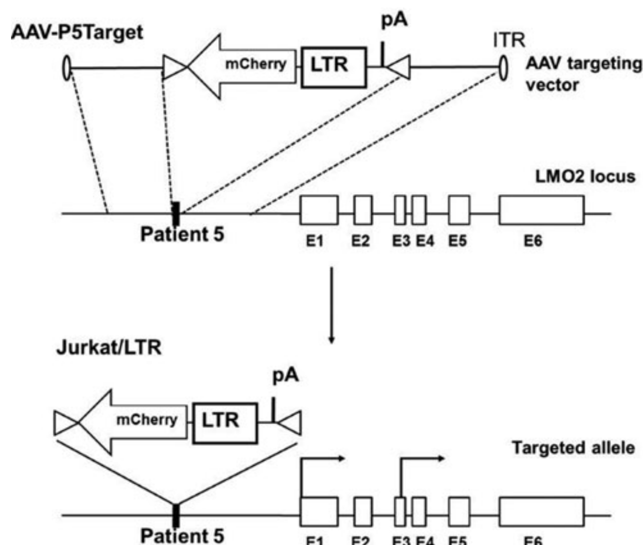
Proviral copy number was determined using quantitative real-time polymerase chain reaction (qRT-PCR) on gDNA samples prepared using Gentra gDNA Isolation Kit (Qiagen, Valencia, CA). Each sample was assayed using an assay detecting the vector sequence (forward 5′CCTCAGACCC TTTTAGT CAGTGT3′, reverse 5′CTTTCGCTTTCAAGTCCC TGTTC3′, probe 5′6FAMCCACTGCTAGAGATTTTMGB NFQ3′) and a control assay detecting mFrizzled (assay ID m*M* 02524776_s1; Applied Biosystems, Carlsbad, CA). A CFU-S colony that had previously been validated by Southern blot was utilized as the single-copy control. Samples were analyzed on a StepOnePlus (Applied Biosystems) using TaqMan Fast Universal PCR mix, no AmpErase UNG Part # 4352047 (Applied Biosystems). Copy number was calculated using the ΔΔCt method.

Western blotting

Cell lysates were prepared with M-Per lysis buffer (Themo Scientific, Waltham, WA) containing Protease Inhibitor Cocktail (Themo Scientific). About 40 μg samples of total lysate were analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis on 4–12% Bis-Tris gels (Invitrogen, Grand Island, NY) and transferred to polyvinylidene difluoride membranes (Invitrogen). Membranes were blocked overnight in 5% milk (Bio-Rad, Hercules, CA)– phosphate buffered saline with Tween 20 (PBST). The primary antibody, anti-WASP(B-9) (Santa Cruz Biotechnology, Santa Cruz, CA), was applied at a dilution of 1:200 for 1 hr at room temperature in 5% milk–PBST. After washing, horse radish peroxidase-conjugated antibodies goat anti-mouse (GE Healthcare, Piscataway Township, NJ), at a dilution of 1:2000, and goat anti-*β*-actin (Santa Cruz Biotechnology), at a dilution of 1:5000, were incubated with the membrane for 1 hr at room temperature in 5% milk–PBST. Detection was performed by ECL (GE Healthcare).

Proto-oncogene activation

Adeno-associated viral vectors were used to achieve targeted insertion of a *γ*-retroviral long terminal repeat driving



SUPPLEMENTARY FIG. S6. Method for generation of the G27 Clones. An AAV-vector for targeting the patient 5 (P5) site of the *LM02* gene was constructed using standard molecular techniques. Jurkat cells were transduced with the AAV vector at an MOI of 3×10^5. Seventeen days later, cells expressing mCherry were sorted into twenty 96-well plates, 1 cell/well, using fluorescence-activated cell sorting. Approximately 900 wells had active cell growth 2 weeks later. Genomic DNA from these clones was used as template for PCR to identify clones in which the LTR-mCherry expression cassette, which is flanked by the Loxp and Lox511 site, has been successfully targeted to the P5 site. Six clones were isolated and further confirmed using Southern blot to contain the targeted allele. All these clones showed significant activation of *LM02* expression, both at the mRNA and protein level. One of the six clones, designated as G27, was used for further experiments.

a GFP expression cassette with flanking loxp-sites in a human T-cell line at the locations of vector integrations in patients with X-SCID who develop leukemia (Ryu *et al*., 2008; Zhou *et al*., 2012). One targeted insertion was immediately upstream from the LM02 distal promoter (Zhou *et al*., 2012) (Supplementary Fig. S6), and a second was within the first intron (Ryu *et al*., 2008). *LM02* mRNA levels were estimated by qRT-PCR and LM02 protein levels were estimated by Western blotting (Ryu *et al*., 2008).

Procedures

Mice

All mice were treated under protocols approved by the Institutional Animal Use and Care Committee.

Bone marrow purification and transduction

The C57BL6/CD45.1 (Jackson Laboratories, Bar Harbor, Maine) mice were sacrificed by CO_2 asphyxiation and cervical dislocation. Bone marrow (BM) was harvested from the tibia and femur via flushing with Stemspan. The cells were passed through a 100 μm cell strainer and purified using a lineage cell depletion kit (Miltenyi, Cologne, Germany). The cells were then cultured at 1×10^6 cells/ml in Stemspan, 10 ng/ml mSCF, 20 ng/ml mTPO, 20 ng/ml mIGF-2/10, ng/ml hFGF (all from Peprotech, Rocky Hill, NJ), 10 μg/ml Heparin (Sigma-Aldrich), 2 m*M* L-glutamine and 1× penicillin/streptomycin (referred to hereafter as pre-stim media) on nontissue culture treated plates for 30 hr. The cells were harvested using HBSS Cell Dissociation Solution (Sigma-Aldrich) and plated at 1×10^6 cells/ml on retronectin (Takara Bio Inc., Shiga, Japan)–coated plates in pre-stim media supplemented with 7.5 μg/ml polybrene and lentivirus at an MOI = 10. After 16 hr the transduction medium was replaced with a half volume of a fresh transduction medium also containing 7.5 μg/ml polybrene and lentivirus at an MOI = 10. After 6 hr, the cells were harvested and transplanted into recipient mice. In addition, cells were placed into myeloid culture and methylcellulose to determine the level of GFP marking.

BM transplant

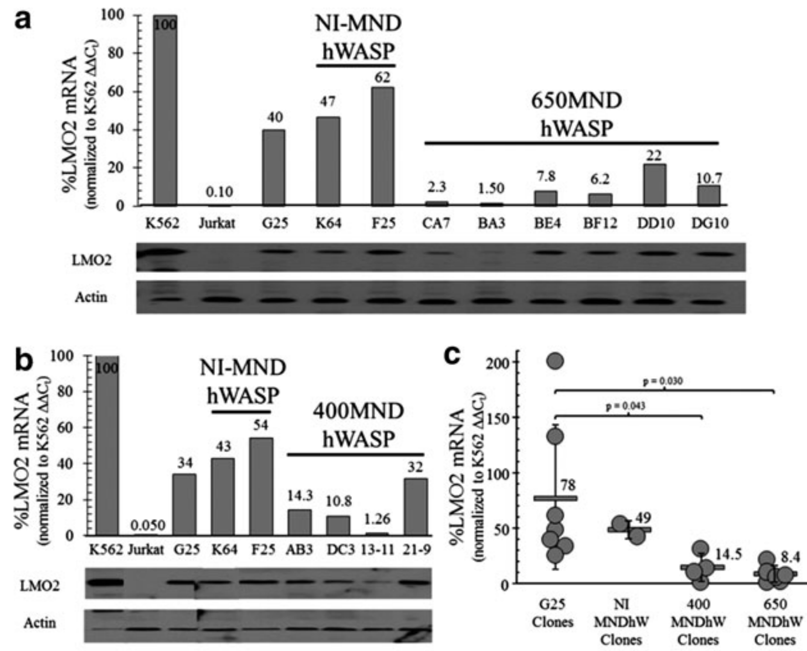
For primary transplants, 0.5×10^6 cells in 200 μl PBS, 2% iFBS, 2 m*M* L-glutamine, and 1× penicillin/streptomycin were administered per animal intravenously into C57BL6/CD45.2 mice (Jackson Laboratories or generated at SJRCH) that had been irradiated at 1125 rad. About 250 μg/ml Baytril (Bayer, Shawnee Mission, KS) was administered in the water bottle for 3 weeks post-transplant. For secondary transplants, 5×10^7 whole BM cells were administered per animal.

BM analysis

Cells were cultured under myeloid conditions at 0.5×10^6 cells/ml in Stemspan, 20 ng/ml mIL-3, 50 ng/ml mIL-6, 50 ng/ml mSCF (all from Peprotech), and 1× penicillin/streptomycin in nontissue culture-treated plates for 7 days. GFP expression was determined using flow cytometry and pro-viral copy number using real-time PCR. For primary CFU-C colonies, 250 cells were plated in 1 ml of Methocult M3434 (Stemcell Technologies) per 35 m*M* dish (two dishes per sample) and cultured at 37°C/5% CO_2 for 12 days. Colony counts and GFP expression were determined using fluorescent microscopy. For BM from primary and secondary transplant harvests, 5×10^4 and 2.5×10^4 cells were used and analyzed at 9 and 7 days, respectively.

Transplant harvest

Mice were sacrificed via CO_2 asphyxiation. Blood was collected from the hepatic artery and analyzed for GFP expression in various lineages. About 5 μl was used to determine %GFP in platelets and red blood cells (RBCs) via flow cytometry. About 85 μl underwent RBC lysis with 155 m*M* NH_4Cl, 10 m*M* $KHCO_3$, and 1 m*M* EDTA and the white blood cells (WBCs) stained with CD45.1-APC-Cy7, CD45.2-PerCP-Cy5.5, Gr1-PE-CY7, NK1.1-PE, CD3-APC, B220-eFlour605, and CD11b-Alexa700/DAPI. The cells were analyzed via flow cytometry. The remaining blood underwent RBC lysis and the WBCs were retained for gDNA analysis. BM was collected from the tibia and femur, passed through a 100 μ*M* cell strainer, and used to establish methylcellulose plates and CFU-S assays. In addition, 5×10^6 cells were used for flow cytometry (stain 1, CD41-PE/CD45.2-PerCP-Cy5.5/CD11b-PE-Cy7/Gr1-APC/CD45.1-APC-Cy7/DAPI; stain 2, NK1.1-PE/CD45.2-PerCP-Cy5.5/B220-PE-Cy7/CD3-APC/CD45.1-APC-Cy7/DAPI), and 2×10^6 were



SUPPLEMENTARY FIG. S7. The cHS4 insulator's ability to diminish LMO2 activation from proviral MNDhWASP cell clones varies with the LMO2 insertion site. **(a)** Jurkat cell clones were screened by PCR to identify CRE-mediated, MNDhWASP exchange clones, in an anti-sense orientation, between Exons 1 and 2 of LMO2, identical to the site mapped for an X-SCID patient containing a γ_c retroviral insertion (Cavazzana-Calvo *et al*., 2000; Hacein-Bey-Abina *et al*., 2003). G25 was one of four AAV-targeted clones that contained a 592 bp retroviral LTR identical to the one used in the X-SCID trials driving GFP, anti-sense to LMO2, and was used for all CRE-mediated cassette exchange reactions (Ryu *et al*., 2008). Positive exchange clones were identified by PCR, expanded, and validated by Southern blot. *LMO2* mRNA (qRT-PCR) and protein (Western blot) were measured, relative to that in K562 cells (positive control) and Jurkat cells (negative control). NI-MNDhWASP exchange clones served as relevant, proviral-positive controls to assess the effect of the cHS4 650 bp insulator on LMO2 activation for 650MNDhWASP exchange clones. **(b)** *LMO2* mRNA and protein levels of CRE-mediated 400MNDhWASP exchange clones. **(c)** The relative amount of *LMO2* mRNA levels for all CRE-mediated, proviral MNDhWASP exchange clones, as assessed by qRT-PCR. In addition to the G25 parental clone, six subclones of G25 were also analyzed and served as positive controls to assess the ability of cHS4 insulator elements to diminish LMO2 activation in Jurkat clones containing a strong retroviral LTR promoter, between exons 1 and 2. A two-tailed, heteroscedastic variance *t*-test was used to assess *p*-values between the groups shown. cHS4, chicken hypersensitive site four; LMO2, LIM domain only two.

used for gDNA analysis. The cervical, auxiliary, mesenteric, and inguinal lymph nodes and spleen were pooled, homogenized, and passed through a 70 μM cell strainer. About 2×10^6 cells were taken for gDNA analysis and 2×10^6 were used to assess GFP expression via flow cytometry.

Peripheral blood mononuclear cell transduction

Peripheral blood was collected from WAS patients and healthy donors after informed consent under a protocol approved by the St. Jude Children's Research Hospital Institutional Review Board and studies were also conducted in

Supplementary Table S1. Vector-Positive Progenitor-Derived Colonies

	Percent GFP positive	
	Primary	Secondary
NIMND-GFP	59.1 ± 8.5 (*n* = 6)	65.5 ± 117 (*n* = 6)
400 MND-GFP	23.2 ± 7.1 (*n* = 2)	31.8 ± 21 (*n* = 5)
650 MND-GFP	73.0 ± 15.2 (*n* = 8)	94.2 ± 4.6 (*n* = 3)
N1EF1α-GFP	23.8 ± 7.9 (*n* = 8)	16.4 ± 16 (*n* = 7)
400EF1α-GFP	32.6 ± 8.4 (*n* = 9)	25.4 ± 15 (*n* = 13)
650EF1α-GFP	41.4 ± 11 (*n* = 6)	25.2 ± 11 (*n* = 7)
NIP1600-GFP	35.5 ± 8.2 (*n* = 9)	16.5 ± 17 (*n* = 3)
400P1600-GFP	40.1 ± 18 (*n* = 9)	8.3 ± 9.4 (*n* = 5)
NIP500-GFP	49.3 ± 11 (*n* = 7)	62.2 ± 18 (*n* = 2)
400P500-GFP	41.1 ± 7.9 (*n* = 7)	40 ± 18 (*n* = 12)

Supplementary Table S2. Titers of Lentiviral pCL20cw…Δ46 Transient Preparations

	Unconcentrated, all promoters, both transgenes		
	NI	400 INS	650 INS
Mean	2.34E + 07	5.15E + 07	1.60E + 07
SD	1.83E + 07	2.31E + 07	1.56E + 07
	Concentrated, all promoters, both transgenes		
	NI	400 INS	650 INS
Mean	4.80E + 08	1.83E + 09	8.45E + 07
SD	5.08E + 08	9.60E + 08	8.62E + 07

SD, standard deviation.

Supplementary Table S3. Platelet Analysis in Transplant Recipients

	Primary		Secondary	
	% GFP (N)	*MFI*/Copy*	*% GFP* (N)	*MFI/Copy*
A. NIMND-GFP	74.7 ± 8.4 (7)	231 ± 120 (*n* = 7)	90.2 ± 6.8 (5)	1143 ± 906 (*n* = 8)
400 MND-GFP	32.3 ± 1.9 (3)	902 ± 203 (*n* = 3)	29.7 ± 23.8 (5)	5542 ± 7081 (*n* = 4)
650 MND-GFP	77.7 ± 22.0 (8)	153 ± 183 (*n* = 8)	97.9 ± 1.2 (3)	5145 ± 4256 (*n* = 3)
B. N1EF1α-GFP	28.3 ± 13.4 (8)	5.0 ± 3.3 (*n* = 8)	2.9 ± 1.8 (7)	11.6 ± 6.9 (*n* = 7)
400EF1α-GFP	58.5 ± 4.8 (4)	29.8 ± 7.1 (*n* = 4)	1.5 ± 0.9 (7)	71.6 ± 108 (*n* = 13)
650EF1α-GFP	5.2 ± 4.9 (5)	42.2 ± 14 (*n* = 6)	1.3 ± 0.9 (7)	188 ± 191 (*n* = 7)
C. NIP1600-GFP	35.7 ± 19.0 (8)	10.7 ± 8 (*n* = 8)	0.8 ± 0.5 (5)	16.2 ± 13.8 (*n* = 5)
400P1600-GFP	29.6 ± 12.4 (9)	53.4 ± 40 (*n* = 9)	5.6 ± 10.8 (5)	119 ± 85 (*n* = 4)
D. NIP500-GFP	52.3 ± 29.6 (8)	2.8 ± 3.3 (*n* = 7)	64.2 ± 7.1 (3)	371 ± 68.2 (*n* = 2)
400P500-GFP	13.7 ± 7.7 (7)	22.8 ± 18 (*n* = 7)	14.6 ± 21.1 (7)	52.4 ± 68.6 (*n* = 12)

*MFI = mean fluorescent intensity.

accordance with the Declaration of Helsinki. PBMCs were isolated using Histopaque (Sigma-Aldrich) density medium and centrifugation. The purified cells were stimulated with T cell activation beads (Miltenyi) for 72 hr in XVivo15 (Lonza), 15% normal human serum (Innovative Research, Novi, MI), 1× penicillin/streptomycin, 2 m*M* L-Glutamine, 55 *μM* *β*-ME, and 10 U/ml hIL-2 (Peprotech) before 2 × 24 hr transductions on retronectin-coated plates with virus at an MOI = 20 (cumulative MOI = 40), with polybrene (7.5 mg/ml) in the culture medium. The cells were re-stimulated with T cell activation beads every 14 days. WASp expression was determined using flow cytometry. The cells were fixed and permeabilized using BD Phosflow Fix buffer 1 and BD Phosflow Perm/Wash Buffer 1 (both from BD Biosciences, Franklin Lakes, NJ) according to the manufacturer's directions and stained with PE-anti-WASP(B-9) (Santa Cruz Biotechnology).

EBVWAS5-2 transduction

EBVWS5-2 cells were plated at 1×10^6 cells per well in a retronectin-coated 12-well plate and transduced with virus as specified for 2 × 24 hr transductions. Cells were expanded and analyzed 21 days post-transduction via Western blotting and qRT-PCR.

Quality assurance

Animal experiments were approved by the Animal Care and Use Committee of the St. Jude Children's Research Hospital and conducted in accordance with the Institute for Laboratory Animal Research Guide for the Care and Use of Animals. The animal care and use program of the hospital is accredited by the Association for the Assessment and Accreditation of Laboratory Animal Care International. Animal care and technical procedures were conducted using standard operating procedures that were reviewed and approved by the animal facility veterinary staff. Observations for morbidity and mortality were performed daily throughout the study. Hemizygous WAS-deficient mice on a C57Bl/6 background were bred from founders and housed in static polysulfonate microisolation cages in a pathogen-free barrier of the animal facility. Mice used for breeding were maintained on a 12-hr-light/dark cycle and provided with Purina rodent chow 5013 and reverse-osmosis-purified water chlorinated to 8 ppm *ad libitum* for the duration of the study. Mice were monitored throughout the study and maintained free of the following murine pathogens: EDIM virus, K virus, lymphocytic choriomeningitis virus, minute virus of mice, mouse adenovirus 1 and 2, mouse hepatitis virus, mouse norovirus, mouse parvovirus, mouse poxvirus, *Mycoplasma pulmonis*, pneumonia virus of mice, polyoma virus, Reovirus, rotavirus, Sendai virus, Theiler's encephalomyelitis virus, *Helicobacter* species, and internal and external parasites.

Results

Promoter activity and expression levels during hematopoiesis

Engraftment in primary recipients was generally equal across all groups, while it was somewhat more variable in secondary recipients as reflected by vector copy number (Supplementary Fig. S1) and GFP expression (Supplementary Fig. S2). GFP marking in engrafted cells was consistent within groups but variable between groups in primary recipients (Supplementary Fig. S2c). The higher variability of engraftment in secondary recipients is reflected in the marking levels in engrafted cells (Supplementary Fig. S1d). Vector copy number in the blood and BM of each group was also highly variable (Supplementary Fig. S3), and therefore the GFP mean fluorescent intensities (MFIs) were normalized by dividing by the GFP MFI of the negative control BM (to control for different flow runs) and by the copy number (Supplementary Fig. S2e and f). Expression from the EF1*α* promoter was very low in the BM. The P500 and P1600 groups, while slightly higher than EF1*α*, were also low in both primary (Supplementary Fig. S2e) and secondary (Supplementary Fig. S2f) recipients. In contrast, expression from the MND promoter was 5–10-fold higher than that from the EF1*α* promoter. While the 400MNDGFP vector showed higher expression in primary transplants than either NIMNDGFP or 650MNDGFP, in the secondary transplant 400MNDGFP expression dropped such that it was equal to NIMNDGFP and lower than 650MNDGFP expression. The BM was also stained for lineage markers to determine if there were any lineage-specific expression enhancement among the promoters, but there was no difference in either primary BM (Fig. 2) or WBCs of secondary recipients (Supplementary Fig. S3).

The frequency of GFP-positive CFU-C from primary recipients (Supplementary Table S1) closely mirrored the BM copy number (Supplementary Fig. S2). In secondary recipients, the EF1α, P500, and P1600 promoters maintained this pattern, with the presence of an insulator having no consistent effect on GFP marking (Supplementary Table S1).

Expression levels in WBCs from primary recipients were similar to those seen in the BM (Supplementary Fig. S3). The lowest expression was seen from EF1α, with the endogenous promoter showing slightly higher expression and the MND promoter showing significantly higher expression. The higher level of GFP expression from 400MNDGFP vector compared with NIMNDGFP and 650MNDGFP was greater in the blood than the BM. However, after secondary transplantation, expression from 400MNDGFP was reduced to NIMNDGFP and 650MNDGFP levels (Supplementary Fig. S3). Lineage analysis of WBCs from primary and secondary recipients did not reveal any further differences between the promoters or insulators.

To examine expression in peripheral lymphoid tissues, the auxiliary, inguinal, cervical, and mesenteric lymph nodes and spleen from each animal were pooled, homogenized, and examined by flow cytometry. As in other tissues examined, expression from MND promoter was much higher than that seen with EF1α, P500, and P1600 promoters (Supplementary Fig. S4). In contrast to what was seen in the blood and BM, expression in 400MNDGFP-treated animals was not increased in peripheral tissues compared with either NIMNDGFP or 650MNDGFP in either primary or secondary transplants. Supplementary Figure S5 compares expression from the various vectors in primary human T cells.

Proto-oncogene activation

We have used targeted integration to recreate two of the insertion sites observed in the *LM02* locus in children with immunodeficiency treated otherwise successfully by gene therapy who subsequently developed leukemia. The initial insertion was within the first intron of the *LM02* gene (Ryu *et al.*, 2008). Now, a new insertion upstream from the distal promoter of the *LM02* gene has been developed using AAV-mediated homologous recombination (Supplementary Fig. S6). We have studied a variety of vectors in this locus by cre-mediated cassette exchange. The results obtained at this integration site are provided in Figure 6. Shown in Supplementary Figure S7 are the data obtained from the integrations within the intron. Activation of the *LM02* gene was observed with the MND promoter, but both insulator elements reduced activation significantly.