2.1 Vectors

2.1.1 The Slippage construct, P-Slip

The PGK promoter fragment, including an ATG translation initiation site, was PCR amplified from a PGK-*puro* cassette (YTC 49, a gift from Youzhong Chen) using primers, PGK-5'-*Sal* I and PGK-3'-*Bam*HI, which add *Bam*HI and *Sal*I sites to the ends of the amplified frangment. The *PurobpA* fragment was prepared by PCR from the same PGK-*puro* cassette using PCR primers Puro-5'-*Spe*I and bPA-3'-*Not*I. The PCR fragment was gel purified and digested with *Spe*I and *Not*I to generate a *PurobpA* fragment with *Spe*I and *Sal*I ends. The $(CA)_{17}$ repeat sequence was constructed from oligos with *Bam*HI or *Spe*I ends, then ligated with the PGK promoter fragment and the *PurobpA* fragment into a pBluescript (pBS) plasmid (Stratagene) to create the P-Slip (Puro Slippage) cassette. Using the same strategy, the repeat sequence $(CA)_{16}$ was used to generate the P-Slip-ON plasmid.

Oligonucleotide for $(CA)_{17}$: *Spe*I-$(CA)_{17}$, 5'-CTA GTG TAT C$(TG)_{17}$ TTG and *Bam*HI-$(CA)_{17}$, 5'-GAT CCA A$(CA)_{17}$ GAT ACA.

Oligonucleotide for $(CA)_{16}$: *Spe*I-$(CA)_{16}$, 5'-CTA GTG TAT C$(TG)_{16}$ TTG and *Bam*HI-$(CA)_{16}$, 5'-GAT CCA A$(CA)_{16}$ GAT ACA.

PCR primers for PGK promoter fragment:

PGK-5'-*Sal*I,

5'-ACG CGT CGA CAG GTC GTC GAA ATT CTA CCG GGT AGG GGA GGC GCT TTT

PGK-3'-*Bam*HI,

5'-CGC GGA TCC GTA CTC GGT CCC CAT GGT GGC GTT GGC

PCR primers for *PurobpA* fragment:

Puro-5'-*Spe*I,

5'-GGA CTA GTA AGC CCA CGG TGC GCC TCG CCA CCC G

bPA-3'-*Not*I,

5'-ata aga atg cgg CCg cAG CTG GTT CTT TCC GCC TCA GAA gc

2.1.2 Gene-targeting vectors

Gdf-9-TV1: Gene-targeting construct for *Gdf-9* (growth and differentiation factor 9) locus. This contains a PGK-*Hprt* minigene as the drug selection marker for gene-targeting (Dong et al., 1996).

Gdf-9-TV2: Derived from *Gdf-9-TV1*, in which a *loxP*-flanked PGK-*neo* cassette was inserted into a *Cla*I site in front of the PGK-*Hprt* cassette.

ROSA26/Slip: Gene-targeting construct for the *ROSA26* locus containing the P-Slip cassette. To create the *ROSA26/Slip* gene-targeting vector, a 1.4 kb *Sal*I/*Xho*I fragment from pL313 (a gift from Dr. Pentao Liu) containing a PGK-*BSD* cassette was inserted into the *Sal*I site of the P-Slip vector. A 2.9 kb *Sal*I/*Not*I fragment containing the *PGK/BSD/pSlip* cassette was ligated into the multiple cloning site of a modified pBS vector, which flanks the inserted *PGK/BSD/Slip* cassette with *Nhe*I and *Bgl*II restriction sites. The *PGK/BSD/Slip* cassette was then released by digesting with *Bgl* II and *Nhe* I, and ligated into pROSA26-1 plasmid containing the genomic fragment of *ROSA26* gene and digested with *Nhe*I and *Bgl*II (Zambrowicz et al., 1997).

Dnmt1-V1: A replacement gene-targeting vector for the *Dnmt1* locus assembled using *E.coli* recombination. In *Dnmt1-V1,* a *loxP*- flanked *Neo/Kan* cassette replaces a 5.5 Kb genomic fragment, resulting in deletion of *Dnmt1* exons 2, 3 and 4. This construct was generated by a colleague, Wei Wang.

129Rb-puro: Targeting vector containing a 129Ola-derived retinoblastoma (*Rb*) genomic fragment.

B/cRb-puro: Targeting vector constructed with a BALB/c-derived *Rb* genomic fragment.

2.1.3 Retroviral vectors

pBabeEGFP: A pBabe derivative, containing a *SV40-EGFP* cassette between the pBabe LTRs (a gift from Dr. Xiaozhong Wang).

pBabeOligo: A pBabe derivative, with minimal cis-elements for viral packaging and a multiple cloning sites between pBabeLTRs (a gift from Dr. Xiaozhong Wang).

pLTRloxP: A 1.5 kb *Eco*RI / *Kpn*I fragment containing the pBabe 3'LTR and *SV40-EGFP* cassette was subcloned into pLitmus (Clontech). A *loxP* site flanked by *Xba*I and *Nhe*I sites was synthesized and cloned into the *Xba*I and *Nhe*I sites in the subcloned 3'LTR to create pLTRLoxP.

pBaER: pBabeOligo was linearized using *Eco*RI and then partially digested with *Kpn*I to obtain a 3.5 kb DNA fragment. pLTRloxP was double-digested with *Eco*RI and *Kpn*I to generate the 1.5 kb DNA fragment containing a *SV40-EGFP* cassette and the LTRloxP fragment, which was then ligated with the 3.5 kb DNA fragment from pBabeOligo to create pBaER.

pBaOR: pBabeOligo was linearized by *Bgl*II and then partially digested with *Kpn*I to generate a 3.7 kb DNA fragment. pLTRloxP was digested with *Bgl*II and *Kpn*I to generate a 350 bp LTRLoxP fragment, which was ligated to the 3.7 kb *Kpn*I / *Bgl*II fragment from pBabeOligo to create pBaOR. The SV40 origin was PCR amplified from the pBabepuro vector (Morgenstern and Land, 1990) and cloned into the *Not*I restriction site of pBaOR. PCR primers: *Not*I restriction site underlined
5'Primer_*Not*I: AGA ATG CGG CCG CTT TTT GCA AAA GCC TAG
3'Promer_*Not*I: AGA ATG CGG CCG CGA CCC TGT GGA ATG TGT G;
PCR cycling conditions: 94 °C 30 seconds, 58 °C 30 seconds, 72 °C 1 minute, for 30 cycles.

pCbOR: To generate a retroviral vector with a CMV enhancer and promoter, a 1.2 kb *Xba*I/*Cla*I fragment containing the 5'LTR from pBaOR was first subcloned into pBS. The 110 bp promoter region was replaced using an oligonuclotide containing *Xba*I and *Hind*III sites. This modified 5'LTR was digested with *Xba*I and *Cla*I, and cloned back into pBaOR to replace the original 5'LTR. A 1.6 kb fragment spanning the viral 5' LTR to the end of the 3' LTR was amplified by PCR. The PCR product was digested with *Hind*III and *Apa*I and cloned into *Hind*III and *Apa*I digested pcDNA3-EGFP vector (a gift from Dr. Xiaozhong Wang) to place the viral backbone following CMV promoter.
PCR primers: 5'primer: CGG TCC AGC CCT CAG CAG;
3'Primer_*Apa*I: CGG GGC CCT GAT ACA TGC TGC ATG TG
PCR condition: 94 °C 1.5 minutes, 57 °C 3 seconds and 68 °C 2 minutes for 25 cycles using the Roche Expand Long PCR System (Roche).

pBaERneo: A *Xho*I and *Bam*HI–flanked DNA fragment containing a PGK-*neo* cassette was cloned into pBaER digested with *Sal*I and *Bam*HI.

pBaORneo, pCBaORneo: A *Xho*I and *Bam*HI–flanked DNA fragment containing a PGK-*neo* cassette was cloned into pBaOR and pCBaOR (respectively) digested with *Xho*I and *Bam*HI.

pBeGTV, pCbGTV (RGTV-1): Gene trap retroviral vectors containing *SAβgeo* gene trap cassette. To introduce the *SAβgeo* cassette into the retroviral backbone so that the transcription of *βgeo* is reversed in relation to the viral transcription, a *Sna*BI restriction site was introduced into pSAβgeo (Friedrich and Soriano, 1991) between the *Sal*I and *Kpn*I restriction sites which follow the polyA signal sequence, and the *Xho*I site between *Sal*I and *Kpn*I was deleted. The *SAβgeo* cassette was then obtained by *Xho*I and *Sna*BI double digestion and cloned into *Eco*RV/*Sal*I digested pBaER or *Eco*RV/*Xho*I digested pCbOR to create pBeGTV and pCbGTV, respectively.

2.2 Cell culture

2.2.1 ES cell culture

ES cell culture has been described in detail (Ramirez-Solis et al., 1993). Briefly, ES cells were maintained on mitotically inactivated feeder cell layers (SNL76/7) in standard M15 medium (Knockout Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 15% foetal bovine serum (FBS), 2 mM L-Glutamine, 50 units/ml Penicillin, 40 μg/ml Streptomycin and 100 μM β-Mercaptoethanol (β-ME) (Invitrogen)). Cells were cultured at 37 °C with 5% CO_2. If not specified, ES cell medium was changed daily.

2.2.2 Chemicals used for selection in ES cells

Blasticidin, Blasticidin S HCl (Invitrogen), 1000x stock (2.5 mM) was made in Phosphate Buffered Saline (PBS).
FIAU, 1-(2'-deoxy-2'-fluoro-β-D-arabinofuranosyl)-5-iodouracil, 1000X stock (200 μM) was made in PBS and 5 M NaOH was added dropwise until it dissolved.
G418, Geneticin (Invitrogen), was bought as liquid containing 50 mg active ingredient per milliliter.
Puromycin, ($C_{22}H_{29}N_7O_5$.2HCL, Sigma) 1000X stock (3 mg/ml) was made in MiliQ water.
50X HAT supplement (Hypoxanthine-aminopterin-thymidine) (Invitrogen).
100X HT supplement (Hypozanthine-thymidine) (Invitrogen).
6TG: 2-amino-6-mercaptopurine (Sigma) 10 mM stock was made in PBS and 5 M NaOH was added dropwise until it dissolved.
.

2.2.3 Transfection of DNA into ES cells by electroporation

DNA to be used for electroporation was either prepared by the standard alkaline lysis method followed by the CsCl banding purificaion or was prepared using QIAgen Plasmid Purification Kits (QIAGEN). Before electroporation, DNA was purified by ethanol precipitation and air-dried in a tissue culture (TC) hood. If DNA linearization was required, for example for gene-targeting, plasmid DNA was first digested by an appropriate restriction enzyme before ethanol precipitation. The air-dried DNA was dissolved in TC hood in Tris-Cl (1 mM, pH 8.2) to a final concentration of 0.5-1 μg/ μl and 20 μl was used for each electroporation.

ES cell electroporation has been described in detail (Ramirez-Solis et al., 1993). Briefly, $1x10^7$ ES cells were eletroporated in a 0.4 cm gap cuvette with 10 to 20 μg DNA using Biorad Gene Pulser at 230 V, 500 μF. After electroporation, ES cells were plated onto a 90 mm feeder plate and followed by an appropriate drug selection as needed. If not specified, the drug selection procedure followed the description in (Ramirez-Solis et al., 1993).

2.2.4 Rb-targeting using Isogenic and non-isogenic gene targeting vectors

The Rb-targeting vectors, Rb129Rb-puro and B/cRb-puro were linearized with *Hind*III and electroporated into ES cells and selected with Puromycin (3 μM) for 8 days. 96 Puromycin resistant clones from each cell line were picked and genomic DNA was extracted and digested with *Eco*RI for Southern analysis using a *Rb* probe for gene-targeting events (expected sizes of detected bands were 9.7 kb for wild type and 4.7 kb for targeted). Targeting efficiency was determined as the number of targeted clones versus the number of samples exhibiting the 9.7 kb wild type band on the Southern-blot.

2.2.5 Gene-targeting of *ROSA26/Slip-TV1* construct

10 μg of *ROSA26/Slip-TV1* DNA was linearized with *Kpn*I restriction enzyme and electroporated into ES cells. Blasticidin selection was initiated 48 hours post electroporation and continued for 4 days. At this stage feeder cells and most of the ES cells were dying and detaching from the bottom of the plate. Fresh M15 medium with $2x10^6$ mitoticaly inactivated SNL76/7 feeder cells were added to the plates. The surviving ES cells were allowed to grow in fresh M15 medium for 6 days to form ES cell colonies. 48 ES cell clones from each targeted cell line were picked into 96 well tissue culture plates and expanded for further analysis. For Southern analysis, genomic DNA was extracted and digested with *Eco*RI and hybridized with a *ROSA26* probe.

2.2.6 Cre-mediated recombination

20 μg of Cre-expressing plasmids pOG231 (CMV-Cre from Steve O'Gorman) or pCAAG-Cre (Araki et al., 1995) was electroporated into 0.5 -$1x10^7$ ES cells. After electroporation, cells were diluted in M15 and about 1,500 ES cells were plated onto a 90 mm feeder plate and cultured for 10 days to allow the formation of single ES cell colonies. 96 ES cell clones were picked into a 96 well tissue culture plate. To identify Cre-mediated recombination events, cells cultured in 96 well plates were duplicated into several 96-well tissue culture plates and sib-selection was performed to identify ES clones that lost the drug selection marker flanked by *loxP* sites. The revertant clones were expanded and loss of the drug selection marker was confirmed by Southern analysis.

2.2.7 Clonal survival assay

Cells were seeded at clonal density (200 to 250 cells per well) in one well of a 6-well tissue culture plate, with or without drug selection. 10 days later, the ES

clones were stained with 2% methylene blue in 70% ethanol for 5 minutes and the number was counted.

2.2.8 EMS mutagenesis in ES cells

EMS (Ethyl methanesulphonate) was purchased from Sigma (1.17 g/ml). A stock (20 mg/ml) was made in PBS and diluted to its final concentration immediately before treatment of ES cells. ES cells at approximately 50% confluence were fed with EMS supplemented M15 medium for 16 hours, then cells were washed three times with PBS, and re-fed with fresh M15. Three hours later, cells were trypsinized and counted. A small portion of cells were diluted to low density and plated onto one well of a 6 well tissue culture plate to determine the survival rate. The survival rate was determined by comparing the plating efficiency of EMS treated cells with non-EMS treated cells of the same genotype. The remaining cells were passaged onto fresh feeder plates.

2.2.9 Retroviral approaches

2.2.9.1 Producing retrovirus by transient transfection

Cells for transfection: The Phoenix ecotropic retroviral packaging cell line, a derivative of human embryonic kidney 293T line expressing retroviral gal, pol and env proteins, was obtained from the American Tissue Culture Collection (ATCC, Manassas, Virginia, USA). In general, Phoenix cells were cultured in M10 medium (DMEM supplemented with10% FBS, 2 mM L-Glutamine, 50 units/ml Penicillin, 40 μg/ml Streptomycin) at 37 °C with 5% CO_2. 18 hours prior to transfection, Phoenix cells were plated at a density of 2.1 million cells per 90 mm plate in M10. Three hours before transfection, cells were fed with 9 ml fresh M10 medium (at this time the cells were about 60% confluent).

DNA preparation: DNA for transfection was prepared by QIAgen Plasmid Purification Kit (QIAGEN). 20-25 μg DNA was used for one transfection of cells cultured on each 90 mm plate. DNA was precipitated with ethanol and air-dried in a TC hood then dissolved in 20 μl TE.

Transient transfection: 500 μl of $CaCl_2$ (0.25 M) was added to DNA and mixed. 500 μl HEBS (0.28 M NaCl, 0.05 M HEPES, 1.5 mM Na_2HPO_4, pH 7.10) was added to the DNA, mixed quickly by bubbling vigorously with an automatic pipettor for 5 seconds. The DNA mixture was kept at room temperature for 5 minutes, and then added to cells cultured in 9 ml of M10 medium. 24 hours later, cell medium was removed and 5 ml of 1% DMSO in PBS was added. After 2 minutes at room temperature, cells were washed twice with PBS and 10 ml of viral producer medium (M10 with heat-inactivated FBS) was layered on each plate. Viral supernatant was harvested 48, 60 and 72 hours after transfection.

2.2.9.2 Viral Infection

ES cells were plated at a density of $3X10^6$ per 90 mm feeder plate about 18 hours before infection. The viral supernatant was filtered through 0.45 μm filter, and polybrene (Hexadimethrine Bromide, Sigma, H9268) was added to the viral supernatant to a final concentration of 4 μg/ml, then added to ES cells.

2.2.9.3 Determination of the transfection efficiency

48 hours post transfection, the viral producer cells were trypsinized and collected in M10. Cells were pelleted and resuspended in PBS at a final concentration of $1x10^6$ cells per ml for flow cytometric analysis of EGFP expression.

2.2.9.4 Titration of the retrovirus

ES cells were plated in 6-well tissue culture plates at a density of $0.5x10^6$ per well in 3 ml M15 medium. 24 hours later, viral supernatant was applied. For the virus carrying a *Neo* cassette, G418 selection (180 μg/ml) was initiated 24 hours after viral infection and continued for 8 days. The drug-resistant ES colonies

were stained with 2% methylene blue in 70% ethanol and counted. The titer of the retrovirus is the number of drug resistant ES cell colonies per milliliter of viral supernatant used to infect the cells.

2.2.10 Gene trap mutagenesis and 6TG resistance screen

2.2.10.1 Gene trap mutants by RGTV-1 retrovirus on NGG5-3 cells.

RGTV-1 retrovirus was produced by transient transfection of Phoenix viral packaging cells. 400 ml of viral supernatant was harvested and filtered through 0.45 μm filter. NGG5-3 ES cells were plated on seventeen of 90 mm feeder plates at a density of $2.5x10^6$ cells per plate. 24 hours later, cells were infected with 5 ml of viral supernatant for at least 5 hours. Viral infection was repeated 5 times. G418 selection (180 μg/ml) was initiated 48 hours after first infection and continued for 8 days. One plate was stained by 2% methylene blue in 70% ethanol to determine the number of gene trap clones obtained. The G418 resistant ES cells from the other 16 infected plates were collected by trypsinization, and cells from two plates were combined and plated to 90 mm feeder plates, generating a total of 8 pools. These cells were cultured for 4 days and frozen down for the subsequent selection in 6TG.

2.2.10.2 Screen for 6TG resistant gene trap mutants

Gene trap mutants which had been expanded for more than 14 population doublings were plated on 90 mm tissue culture plates at a density of 0.7-1 $x10^7$ cells per plate. For high stringency 6TG selection, 6TG selection (2 μM) was initiated 16 hours later and the 6TG-supplemented M15 medium was changed every day for 8 days. After culturing 4 days in fresh M15 medium, the 6TG resistant colonies were picked. For low stringency 6TG selection, cells were plated directly in 6TG (0.5 μM) -supplemented M15 medium which was changed

every other day for 10 days. After culturing in fresh M15 for 2 days, surviving ES cell clones were picked.

2.2.11 Fluctuation analysis of the MSI rate of P-Slip.

ES cell lines with the targeted P-Slip cassette were plated at single cell density on 90 mm tissue culture plates and cultured for 10 days to allow formation of single ES cell colonies. ES cell clones were picked to a 96 well tissue culture plate and independently expanded to the desired number of cells. The expanded clones were trypsinized and the number of cells in each clone were determined using a Beckman-Coulter blood cell counter, then selected in M15 medium containing Puromycin (3 μM) for 8 days. Puromycin-resistant ES cells were allowed to grow and visible colonies were stained with 2% methylene blue in 70% ethanol. The colony number was counted and the MSI rate of P-Slip was calculated by the Luria-Delbruck method of means with equation: $r = aN \ln(NCa)$, where "a" represents the mutation rate "r" represents the mean number of variation per culture (puromycin resistant clones); "C" represents the number of parallel cultures (luria, 1943).

2.3 DNA methods

2.3.1 Probes

2.3.1.1 General probes

***LacZ* probe**: A probe for gene trap viruses containing the *SAβgeo* gene trap cassette, consisting of a 1.4 kb *Cla*I fragment from pSAβgeo, a plasmid containing the *SAβgeo* cassette in pBS (from Dr. Philippe Soriano).
***Neo* probe**: A probe for gene trap viruses containing the SAβgeo gene trap cassette and consisting of a 700 bp *Pst*I /*Xba*I fragment from the PGK-*Neo* cassette.

***Rb* probe**: 450bp *Pst*I/*Pvu*II fragment from pPHA153 (a gift from Dr. Hein te Riele).

***Gdf-9* probe**: 650bp *Bam*HI/*Sal*I fragment from pGDF9-212D (a gift from Dr. Martin Matzuk).

γ*SAT* probe: Probe for paracentromeric gamma satellite repeats. 1.9kb *Not*I and *Sal*I fragment from pγSAT plasmid (a gift from Niall Dillon).

2.3.1.2 PCR amplified genomic DNA probe

Genomic DNA probes were PCR amplified from AB2.2 mouse genomic DNA and used for Southern-blot analysis.

Msh6 exon2 (F), 5'-GCAACAGTTCTTGTGACTTCTCACCA
(R), 5'-CCTCTTACCTGTATATGGCTTTAACAT, 180 bp

Dnmt1 (F), 5'-GCAGTTTGTTTAAATAGAAGTGTGCATAGT
(R), 5'-GTCCCCTAACACATACCTTCGTGTAT, 685 bp

Tgif (F), 5'-CGCCAGCGCGCTCCGACTTCTTAACT
(R), 5'-GAGCAGCGACGTCACCGCCGGTG, 1.1kb

Rbpsuh (F) 5'-GAATTCCCTTATCTCTAAAAGGAGCATAT
(R) 5'-GACTCCACATTAACACAGAGATGTTAAG, 721 bp

mMRG9 (F), 5'-CGACTGTGGGCCGAAGGTTCGAGGCTGT
(R), 5'-CCGCCTGTCCTTGTACATCGATTAATTAAAACGT, 900 bp

ROSA26 (F) 5'- CTGGATCCTCCCCAATCAAAAGTATAGG
(R) 5'- CTCCCTGTGGCGTATGCCCCAGTATCC, 660 bp

STA1.2 (F), 5'-GAGAGGTCACCATTATTTCTAGAATGGCCTA
(R), 5'-CTGAAGAAATACAGCCTGGATATCCACAGCT, 1.2 kb

AldpS (F), 5'-GCTTCCCAAGTGCTGGGATTAAAGGTATGTGT
(R), 5'-CAGGGTACTGCAGCAAAGGAGCCCAGGT

AldpL (F), 5'-CCATTCAGGACAACCACAGAGTACTGGATCA
(R), 5'-CTCATGTGAGTATATGGACATGTAAGTTGGGTAT

2.3.1.3 cDNA probe

cDNA probes were PCR amplified from AB2.2 cDNA and used for Northern-blot analysis or Southern-blot analysis.

Msh6 exon1 (F), 5'-CGTCAGCCTTCATCATGTCCCGACAA
exon4 (R), 5'-GGTGCCTAGGTCCTCACTATC, 876 bp
Msh2 exon1 (F), 5'-GCAGCCTAAGGAGACGCTGCAGTTG
exon18 (R), 5'-CCGTGAAATGATCTCGTTTACGAAGCTG, 2.8 kb
Mlh1 exon1 (F), 5'-GGCGTTTGTAGCAGGAGTTATTCGG
exon19 (R), 5'-AGACTTTGTATAGATCTGGCAGGTTGGC, 2.3 kb
Dnmt1 exon1 (F), 5'-GCTCCAGCCCGAGTGCCTGCGCTTG
exon6 (R), 5'-CTCTGTGTCTACAACTCTGCGTTTc, 543 bp

2.3.2 Southern blotting and hybridization

1). Southern blotting: Genomic DNA was digested with an appropriate restriction enzyme and the digested fragments were separated by electrophoresis on 0.8% agarose gel in 1XTAE buffer. The gel was soaked in Depurination Buffer (0.25 M HCl) for 10 minutes with gentle agitation, and then transferred into Denaturation Buffer (0.5 M NaOH, 1.5 M NaCl) for 30 minutes with gentle agitation. A capillary blot was set up according to standard methods, using Hybond-N+ membrane (Amersham) using Denaturation Buffer as the transfer buffer. Following overnight transfer, the blot was neutralized in Rinse Buffer (0.2 M Tris-Cl (pH7.4), 2X SSC) for 5 minutes, and baked at 80 °C for 1 hour.

2). Probe preparation: 5-20 ng of probe DNA was labeled using Rediprime^{TM} II Random Prime Labeling System (Amersham) according to the manufactor's protocol and purified with a G-50 column. The probe was denatured at 100 °C for 5 minutes, and chilled on ice for 5 minutes before use.

3). Hybridizaiton: The blot was pre-hybridized at 65 °C for 2 hours in Hybridization Buffer (1.5X SSCP, 1X Denhardts solution, 0.5% SDS, 10% Dextran Sulfate). The denatured probe was added into the hybridization tube and incubated at 65°C for at least 8 hours. The blot was first rinsed briefly in low stringency wash buffer (1X SSC, 0.1% SDS) at room temperature and then washed twice in high stringency wash buffer (0.5X SSC and 0.1% SDS) at 65 °C for 15 minutes. The blot was then exposed to X-ray film (Fuji).

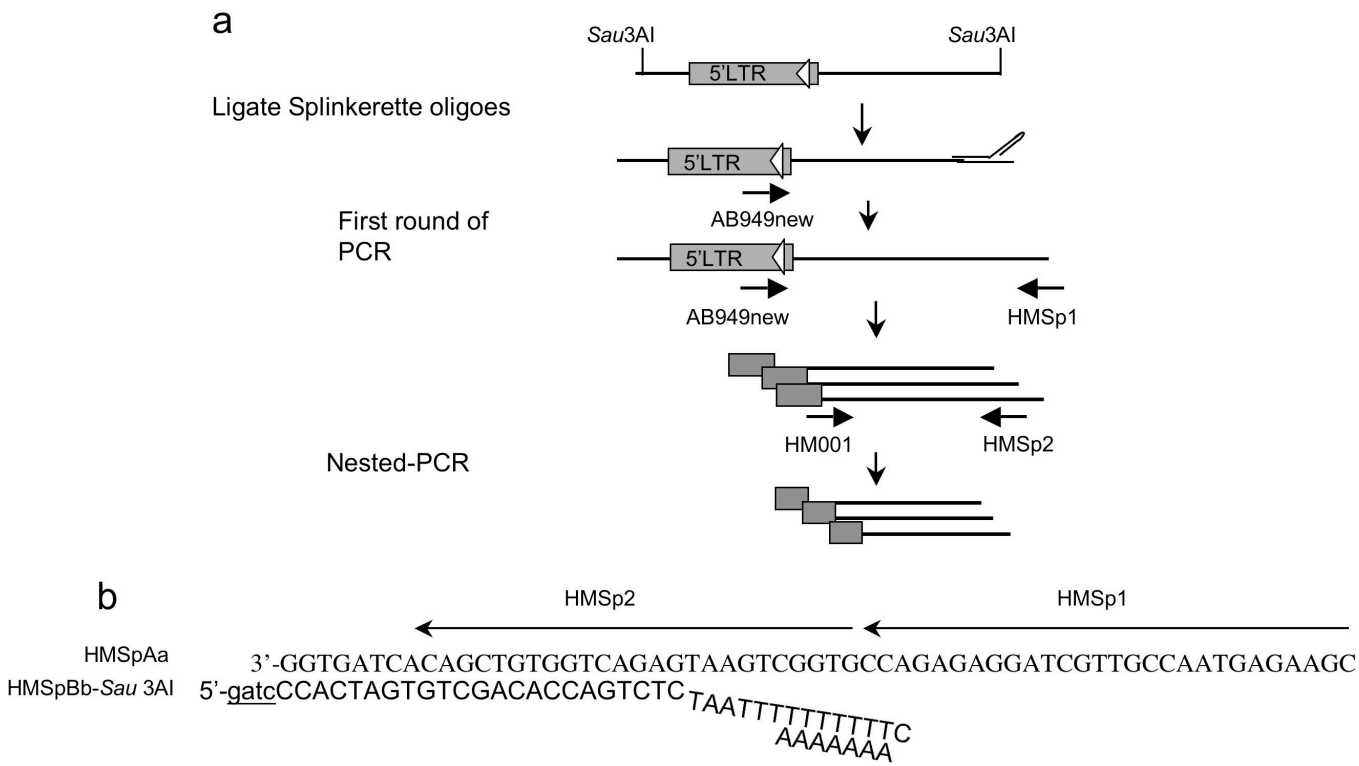

Figure 2-1. Splinkerette PCR. **a.** Schematic representation of Splinkerette PCR procedure**.** The *Sau*3AI digested DNA is ligated with an annealed Splinkerette oligoes (Splinkerette adaptor). The first round pf PCR is carried out with viral primer AB949new and Splinkerette primer HMSp1. The nested PCR is carried out with viral primer HMOO1 and Splikerette primer HMSp2. **b.** The structure of an annealed Splinkerette Oligos, showing the *Sau* 3AI site (under lined font), the loop structure and the position of the HMSp1 and HMSp2 primers.

2.3.3 Isolation of proviral/host junction by Splinkerette PCR

3 μg of genomic DNA was digested with *Sau*3AI in a 30 μl volume at 37 °C for 3 hours. The reaction was stopped by heating at 65 °C for 20 minutes. The Splinkerette Oligos were annealed in a reaction mix containing 150 pmol of HMSpAa-*Sau3* AI, 150 pmol of HMSpBb, 5 μl NEBuffer 2 (New England Biolabs) in a total of 100 μl. After a short incubation at 95 °C for 3 minutes, the mixture was cooled slowly to room temperature. 3 μl of the annealed Splinkerette oligos were used for each ligation reaction in 20 μl volume containing 5 μl digested genomic DNA, 2 μl 10X Ligation Buffer and 5 units T4 DNA Ligase (New England Biolabs). The ligation reaction was carried out at 16 °C overnight. The T4 DNA ligase was then heat-inactivated at 65 °C for 15 minutes. 20 μl of *Cla* I digestion mix containing 10 units *Cla*I (New England Biolabs), 4 μl 10X *Cla*I Buffer (New England Biolabs), 14 μl H_2O was added to the ligation reation and incubated at 37 °C for 2 hours. Unligated oligos were removed by gel filtration using Sephacryl^{TM}S-300 (Amersham) as described in section 2.4.4.

First round of PCR: The 5'LTR proviral junction was amplified with the LTR specific primer, AB949new, and the Splinkerette primer, HMSp1 in 50 μl PCR system containing 20 μl purified genomic DNA, 2 μl AB949new (10 μM), 2 μl HMSp1 (10 μM), 5 μl 10x PCR buffer, 1.5 μl $MgCl_2$ (50 mM), 1 μl dNTPs (25 mM), 0.5 μl PlatinumTaq (5 units/μl, Invitrogen), ddH_2O 18 μl. The PCR cycling was performed at 94 °C 2 minutes, cycles of 94 °C 1 minutes, 68 °C 30 seconds, 72 °C 1.5 minutes and followed by 30 cycles of 94 °C 30 seconds, 65 °C 30 seconds, 72 °C 2 minutes and finished by 72 °C incubation for 10 minutes. The first round of PCR product was diluted at a 1 to 200 ratio and 5 μl of the diluted product was used as the template for the nested PCR with primers, HM001 and HMSp2, in 50 μl PCR system containing 1 μl HM001 (10 μM), 1 μl HMSp2 (10 μM), 5 μl 10x PCR buffer, 1.5 μl $MgCl_2$ (50 mM), 1 μl dNTPs (25 mM), 0.5 μl PlatinumTaq (5 units/μl, Invitrogen), ddH_2O 35 μl. PCR cycling was performed at

94 °C 1.5 minutes, 30 cycles of 94 °C 30 seconds, 60 °C 30 seconds, 72 °C 1.5 minutes and finished by 72 °C incubation for 7 minutes. The nested PCR products were separated on a 1% agarose gel. The specific PCR fragments were gel purified using QIAquick Gel Extraction Kit (QIAGEN) according to the manufacture's instructions. 5-20 ng of the purified DNA was used for sequencing using two sequencing primers.

To obtain longer provial/host juction fragments, genomic DNA was digested with restriction enzyme *Eco*RI*, Hind*III, *Xba*I*, Spe*I*, Nhe*I*, Bam*HI *or Nco*I. The Splinkerette adapters were made by annealing the related HMSpAa and HMSpBb oligos and ligated to the digested genomic DNA (Fig. 2-1b). The same Splinkerette PCR and LTR primers used for amplifying *Sau*3AI digested genomic DNA were used to amplify the proviral/host junction following the above protocol, with 3 minutes of PCR elongation time.

Splinkerette Oligos:
HMSpAa: 5'-CGA AGA GTA ACC GTT GCT AGG AGA GAC CGT GGC TGA ATG AGA CTG GTG TCG ACA CTA GTG G
HMSpBb-*Sau*3AI
5'-gat cCC ACT AGT GTC GAC ACC AGT CTC TAA $(T)_{10}C(A)_7$
HMSpBb-*Hind*III
5'-agc tCC ACT AGT GTC GAC ACC AGT CTC TAA $(T)_{10}C(A)_7$
HMSpBb -*Nco*I
5'-cat gCC ACT AGT GTC GAC ACC AGT CTC TAA $(T)_{10}C(A)_7$
HMSpBb –*Xba*I
5'-cta gCC ACT AGT GTC GAC ACC AGT CTC TAA $(T)_{10}C(A)_7$
HMSpBb -*Eco*RI
5'-aat tCC ACT AGT GTC GAC ACC AGT CTC TAA $(T)_{10}C(A)_7$
HMSpBb -*Bam*HI
5'-gat cCC ACT AGT GTC GAC ACC AGT CTC TAA $(T)_{10}C(A)_7$

PCR primers:
AB949new: 5'-GCT AGC TTG CCA AAC CTA CAG GTG G
HM001: 5'- GCC AAA CCT ACA GGT GGG GTC TTT
HMSp1: 5'-CGA AGA GTA ACC GTT GCT AGG AGA GAC C
HMSp2: 5'-GTG GCT GAA TGA GAC TGG TGT CGA C
Primers for sequencing:
HM002: 5'-ACA GGT GGG GTC TTT CA; HMSp3: 5'-GGT GTC GAC ACT AGT GG

2.3.4 Cre-*loxP* mediated reversal assay of integrated retrovirus

Cre-mediated recombination was performed as described in Section 2.2.6. To identify clones which had undergone Cre-mediated excision, PCR was performed on 200 ng of genomic DNA to amplify the SAβ*geo* gene trap cassette using the primers lacZ(F) and LacZ(R), which amplify a 335 basepair *LacZ* fragment. A pair of PCR primers, Ctbp2 (F) and Ctbp2 (R), were included in the PCR reaction as a control, which amplified a 490 basepair fragment from *CtBP2* (C-terminal binding protein 2) (Fig. 5-11a). Clones which did not show amplification of the *LacZ* fragment were expanded and the clonal survival were checked in 6TG (See Section 2.2.6). The primer sequences:
LacZ (F), 5'- CGA ATA CGC CCA CGC GAT GGG TAA CA
LacZ (R), 5'- CGC TAT GAC GGA ACA GGT ATT CGC TGG T
Ctbp2 (F), 5'-CTC GCC AGC AGC CTT GAT GTC CAC GTT GT
Ctbp2 (R), 5'-CCT GGT GGC ACT GCT GGA TGG CAG AGA CT.
PCR cycling was performed at 94 °C for 2 minutes; 30 cycles of 94 °C 30 seconds, 58 °C 30 seconds, 72 °C 1 minute and finished by incubation at 72 °C for 7 minutes.

2.3.5 Quantitative Southern analysis (QTSouthern)

2 to 4 μg of genomic DNA was digested using 20 units each of *Eco*RV and *Hind*III (New England Biolabs) in 40 μl, and separated on a 0.8% agarose gel. Southern blotting and hybridization were performed as described in section 2.3.2. 20 ng of each probe, *AldpS, AldpL* and *LacZ* was used for each hybridization. *AldpS* and *AldpL* probes were PCR amplified from AB2.2 genomic DNA for X-linked gene Adrenoleukodystrophy Protein Homolog using PCR primers AldpL(F), AldpL(R), AldpS(F), AldpS(R). *AldpS* recognizes a 2.5 kb *Hind*III fragment. *AldpL* recognizes a 3.7 kb *Eco*RV/*Hind*III fragment. The *lacZ* probe recognizes a 3.0 kb *Eco*RV/*Hind*III fragment from the gene trap cassette.

2.3.6 Sequencing analysis of the recovered P-Slip cassette

Individual Puromycin resistant ES colony was picked into a 96 well tissue culture plate, expanded, and genomic DNA was extracted. 50 μl 0.1xTE Buffer (0.1 mM EDTA, 1 mM Tris-Cl (pH 8.0)) was added to each well and incubated at 55 °C overnight to dissolve the genomic DNA. 5 μl dissolved genomic DNA was used for PCR reaction to amplify the 1.1 kb DNA fragment containing the $(CA)_{17}$ repeat sequences. The PCR reaction was performed with primers, PGK (F) and Rosa3'arm, in a total of 50 μl using the Expand Long Template PCR System (Roche) following the manufacturer's instructions. 15 μl of PCR product was treated with 2 units of Exonuclease I (*Exo*I, NEB) and 2 units of Shrimp Alkaline Phosphatase (SAP, Amersham) at 37 °C for one hour to degrade the single-stranded PCR primers and destroy the dNTPs. The enzymes were inactivated at 95 °C for 15 minutes, and 5 μl was used for sequencing with the PCR primers. Underlined is the *Eco* RI restriction site.

PGK (F) 5'-<u>cgg aat tc</u> G GGC AGC GGC CAA TAG CAG CTT TGC T

Rosa3'arm: 5'-<u>cgg aat tc</u>G ATA GAA CTT GAT GTG TAG ACC AGG CTG G

2.3.7 Determination of the CpG methylation pattern

2 μg genomic DNA was digested with 20 unit of *Msp*I or *Hpa*II (New England Biolabs) at 37 °C overnight. The digested DNA was separated on a 1% agarose gel and blotted onto Hybond-N+ membrane (2.3.2). Southern hybridization was carried out with the gamma satellite probe γSAT (2.3.1.1) in Rapid-Hyb Buffer (Amersham Bioscience) according to the manufacturer's instructions.

2.4 RNA methods

2.4.1 RT-PCR

Total RNA was prepared using an RNAqueous™ Kit (Ambion). 5 μg of total RNA was treated with 1 μl amplification grade DNase I (1 unit/μl, Invitrogen) in 10 μl volume for 10 minutes at room temperature (RT) to degrade the residual DNA. 1 μ1 of EDTA (25 mM) was added to the reaction, which was incubated at 65 °C for 15 minutes to inactivate the DNase1. 1 μl of Oligo-dT primer (10 μM) was added to the reaction, followed by incubation at 65°C for 10 minutes. The tube was placed on ice. First strand reaction mix (0.5 μl dNTPs (25 mM), 5 μl 5x first-strand buffer, 2.5 μl DTT (0.1 M), 1 μl Superscript™ II (5 units/μl)) was made and aliquoted into each sample, followed by an incubatation at 50 °C for 60 minutes. 1 μl of Ribonuclease H (2 U/μl) was added to each reaction and incubated at 37 °C for 30 minutes to remove RNA. The resultant cDNA was diluted at a ratio of 1:5 with ddH_2O and 2 – 5 μl was used for each PCR reaction. The PCR cycling conditions were: 94 °C for 2 minutes; 30 cycles of 94 °C for 30 seconds, 55 °C for 30 seconds, 72 °C for 1 minute; 72 °C for 7 minutes. dNTPs, Superscript™ II, Ribonuclease H and Platium Taq were purchased from Invitrogen.
Oligo-dT primer: GGC CAC GCG TCG ACT AGT AC $(T)_{17}$

2.4.2 RT-PCR primers

Dnmt1 **exon1 (F), 5'-GCT CCA GCC CGA GTG CCT GCG CTT G**
exon6 (R), 5'-CTC TGT GTC TAC AAC TCT GCG TTT C

Tgif

Tgif-α, exon1 (F) 5'-GCC ACT CCA CGG CTG CTG GCT CCT

Tgif-β, exon1 (F), 5'-GAG CTG AGG GAT GGA GAT GGT GCT CT

Tgif-γ, exon1 (F), 5'-CTG CCT CGA AAA GAT TTA TGC GAG CAG A
exon3 (R), 5'-TCT CAG CAT GTC AGG AAG GAG CCT G

Parp-2, exon1 (F), 5'-GCA GAG ATC AGG CTC TGG AAG GCG A
exon5 (R), 5'-GTG CTG GCA GCA TAG TCC ATC TGT A

Rbpsuh exon1 (F), 5'-CTC AGT CTC CAC GTA CGT CCC CGA G
exon4 (R), 5'-CAG AAC ATC CAT CTC GTT CCA TTT GCT CT

mMRG9 exon1 (F), 5'-GCG TCT GAC GCT GAG TTG GGT
exon6 (R), 5'-CCT CTC ATC TTG CCC TCT GCA

LacZ-Gsp2 (R): 5'- atg tgc tgc aag gcg att aag

2.4.3 Northern blotting and hybridation

RNA gel electrophoresis: Total RNA was prepared using the RNAqueous™ Kit (Ambion). 10-20 μg total RNA was mixed with 10 μl RNA loading dye, denatured by heating at 70 °C for 15 minutes and chilled on ice for 5 minutes. RNA was then separated by electrophoresis on a formaldehyde-agarose gel (1% agarose, 3% formadehyde, 1X MOPS solution in DEPC-treated H_2O) in RNA gel running buffer (1X MOPS in DEPC-treated H_2O).

Northern blotting: The gel was soaked in DEPC-treated H_2O for 30 minutes and then in 10X SSC for 30 minutes. RNA was blotted to Hybond-N+ membrane (Amersham) using standard capillary transfer method in 10X SSC for 12 hours. The blot was rinsed in 6X SSC for 5 minutes and baked at 90 °C for 1 hour. 20 ng of probe DNA was labeled using Rediprime™ II Random Prime Labeling System (Amersham) according to the manufacturer's protocol and purified using

a G-50 column. The labeled probe was denatured by heating to 100 °C for 5 minutes and then chilled on ice for 5 minutes before use. Hybridization was carried out in Rapid-hyb Buffer (Amersham) according to the manufacturer's instructions for 4 hours. The blot was washed in the low stringency wash buffer (1X SSC and 0.1% SDS) at room temperature for 5 minutes, and then washed once in high stringency wash buffer (0.5X SSC, 0.1% SDS) at 65 °C for 10 minutes. The blot was then exposed to X-ray film (Fuji).

2.4.4 Isolation of the 5' end of gene trap transcripts by 5'RACE

1). First strand cDNA synthesis: First strand cDNA synthesis was performed as described in Section 2.4.1 using a *lacZ* primer, LacZ-GSP1. cDNA was purified using a QIAquick PCR Purification Kit (QIAGEN), and eluted in 30 μl Tris-Cl (10 mM, pH8.2). (Note: If the cDNA synthesis was performed on samples in a 96-well plate, the resultant cDNA was purified using Sephacryl^{TM}S-300 following the protocol below).

2). Sample purification and size selection: Sephacryl^{TM}S-300 Media (Amersham Bioscience) was mixed at a 1:1 ratio with MilliQ water. 200 μl of this mixture was added to each well of a 0.2 μm PVDF filtration plate (Corning Inc.). The filtration plate was placed onto a collection plate and spun for 2 minutes at 600 g, and the flow-through was discarded. This step was repeated once. 200 μl of H_2O was added to each well of the filtration plate, spun for 2 minutes at 600 g, and then the flow-through was discarded. This step was repeated once. cDNA samples were loaded on to the filtration plate. This was placed onto a fresh collection plate and the samples were recovered by spinning at 800 g for 2 minutes.

3). PolyC tailing: On ice, 30 μl purified-cDNA was mixed with 10 μl of the tailing-mixture, containing 8 μl 5xTdT buffer, 2μl dCTP (4 mM), and 1 μl TdT enzyme The samples were kept at 37 °C 10 minutes.

4). First round of PCR: On ice, 10 μl of the tailed-cDNA was mixed with the PCR reaction mixture containing 5 μl 10X PCR buffer, 1.5 μl of $MgCl_2$ (50 mM),

0.5 μl dNTP (25 mM), 1μl LacZ-GSP2 (10 μM), 1 μl AAP (10 μM), 0.5 μl of PlatinumTaq (5 units/μl, Invitrogen) 30.5 μl MilliQ H_20. The PCR reaction was performed using the following conditions: 94 °C for 1.5 minutes; 35 cycles of 94 °C 30 seconds, 55 °C 30 seconds, 72 °C 1.5 minutes; 72 °C for 10 minutes.

5**). Nested-PCR:** First-round PCR products were diluted at a ratio of 1:100. 5 μl of the diluted PCR product was added to a PCR reaction mixture containing 5 μl 10X PCR buffer, 1.5 μl $MgCl_2$ (50 mM), 0.5 μl dNTP (25 mM), 1 μl LacZ-GSP3 (10 μM), 1 μl AUAP (10 μM), 0.5 μl Platinum Taq (5 units/μl, Invitrogen), 35.5 μl MiliQ H_20. PCR was performed using the following conditions:
94 °C for 1.5 minutes; 35 cycles of 94 °C 30 seconds, 55 °C 30 seconds, 72 °C 1.5 minutes; 72 °C for 10 minutes. 10 μl of the nested PCR products were loaded on a 1.0 % agarose gel. Samples with multiple PCR bands were gel purified and each band was cloned using TOPO TA Cloning kits (Invitrogen).

6). Sequencing the nested-PCR product: If the nested-PCR was performed on a small scale, the nested-PCR product was purified using a QIAquick PCR Purification Kit following the manufacturer's instructions and eluted in 30 μl of Tris-Cl (10 mM, pH 8.0). 5 μl was used for each sequencing reaction in a 10 μl volume, containing 4 μl of ABI PRISM^{TM} Big DyeTerminator Cycle Sequencing Ready Reaction Kit (PE Applied Biosystems), 1 μl of SA-seq (5 μM)).
Sequencing reaction was performed using the following conditions: 94 °C for 1.5 minutes; 40 cycles of 94 °C 30 seconds, 55 °C 30 seconds, 60 °C 4 minutes. If the nested-PCR was performed in a 96 well plate, 10 μl of the nested-PCR product was treated with 1U each of Exonuclease I (Exo I, NEB) and Shrimp Alkaline Phosphatase (SAP, Amersham) as described in Section 2.3.6. 5 μl of the treated PCR products was used for sequencing reaction using SA-seq primers with the same sequencing reaction conditions as above.

7): Purifying the sequencing product: 10 μl of MilliQ water was added to each sample to bring up the volume to 20 μl. To each sample, 50 μl of Precipitation Mix (100 ml 96% ethanol, 2 ml Na_2OAC (3 M, pH 5.2), 4 ml EDTA (0.1 mM, pH 8.0)) was added. The sequencing product in the 96 well PCR reaction plate was

collected by centrifugation at 4000 rpm at 4 °C for 25 minutes. The plate was placed upside-down on tissues to drain the liquid. The samples was washed twice using 100 µl of chilled 70% ethanol followed by a spin at 4000 rpm at 4 °C for 5 minutes. The residual liquid was drained by a final quick spin (200 rpm for 1 minute) with the sequencing plate upside-down on a tissue. The samples were dried at 65°C for 2 minutes. The sequencing reactions were run on a ABI PRISM^{TM} 377 DNA sequencer (Perkin Elmer).

5' RACE-PCR primers: LacZ-GSP1, 5'-GGG CCT CTT CGC TAT TAC GC; LacZ-GSP2, 5'-ATG TGC TGC AAG GCG ATT AAG; SA-GSP3, 5'-GTT GTA AAA CGA CGG GAT CCG CCAT; SA-seq, 5'-TGTCAC AGA TCA TCA AGC TTA TC, AAP and AUAP were purchased from Invitrogen.