Supplementary Materials and Methods

Chimerism analysis.

Transplanted mice were observed for up to 32 weeks. 100-150 µl of peripheral blood (PB) was taken by retro-orbital bleeding or from the tail vein. Blood cells were counted using an automatic analyzer (ABC Counter, Scil, Vierheim, Germany). Donor cells (CD45.2+) and EGFP+ cells were determined by flow cytometry after lysis of erythrocytes (FACS Calibur, Becton Dickinson). Dead cells were excluded by propidium iodide staining.

Vector insertion site (VIS) identification by LMPCR (DBISA).

To extract genomic DNA from peripheral blood, bone marrow or spleen QIAamp DNA blood kit (Qiagen, Hilden, Germany) was used. Gamma-retroviral and lentiviral integration sites analysis was performed by LMPCR, as described.[51] Briefly, to obtain γ-retroviral integration sites, LM-PCR was performed to amplify sequences flanking the 5'LTR. Genomic DNA was digested with restriction enzyme Tsp509I or MseI (New England BioLabs, Frankfurt a. Main, Germany) for 2 h at 65°C or 37°C correspondently. For primer extension biotinylated retroviral specific primer A1RV (5'-CTGGGGACCATCTGTTCTTGGCCTC-3') was used. The first PCR (94°C for 2 min; 94°C for 15 sec, 60°C for 30 sec, 68°C for 2 min for 30 cycles; 68°C for 10 min) was performed using Extensor Hi-Fidelity PCR Master Mix (ABgene, Hamburg, Germany), retroviral primer A2RV (5'-GCCCTTGATCTGAACTTCTC-3') and linker specific primer OCI.[51,52] The nested PCR (94°C for 2 min; 94°C for 15 sec, 60°C for 30 sec, 68°C for 2 min for 30 cycles; 68°C for 10 min) was performed using Extensor Hi-Fidelity PCR Master Mix (ABgene), retroviral primer A3RV (5'-CCATGCCTTGCAAAATGGC-3') and linker specific primer OCII. To obtain lentiviral integration sites, LM-PCR was performed to amplify sequences flanking the 3'LTR. Genomic DNA was digested with restriction enzyme Tsp509I. Reaction conditions, primers and configuration of PCR program described before.[51,52] For primer extension, first and nested PCR, biotinylated lentiviral specific primer LVFW1 (5'-GAACCCACTGCTTAAGCCTCA-3'), lentiviral primers LVFW2 (5'-AGCTTGCCTTGAGTGCTTCA-3') and LVFW3 (5'-AGTAGTGTGTGCCCGTCTGT-3') were used, respectively.[51] PCR products were isolated after gel electrophoresis using QIA quick Gel Extraction Kit (Qiagen) and sequenced directly or after subcloning into TA vector (Invitrogen, Karlsruhe, Germany) using primer RAseq (5'-CTTGCAAAATGGCGTTAC-3') or LVFWseq (5'-CTGTTGTGTGACTCTGGTAAC-3').

Dominant band insertion site analysis (DBISA).

Recovered insertion sites sequences were screened as described.[49,52] Briefly, fragments containing retroviral (gamma-retroviral or lentiviral) genomic junctions and polylinker were further analyzed using BLAST searches were performed at http://www.ncbi.nlm.nih.gov/BLAST or locally using blast 2.2.17, NCBI Mus musculus genome build 37. Integration sites were compared to previously identified hits in the mouse retrovirus tagged cancer gene databases (RTCGD) at http://rtcgd.ncifcrf.gov. Integrations retrieved from the studies presented here and from the IDDb[24] including only primary recipients from experiments with LTR-driven vectors expressing fluorescent or surface markers (161 sequences) were also aligned to mouse NCBIv37 genome, using BLAST 2.2.17.

Real-time RT-PCR.

Total RNA was extracted from peripheral blood or bone marrow cells of experimental and control mice using RNeasy Micro kit (Qiagen). 20-500 ng RNA was reverse transcribed into cDNA with Power Script Reverse Transcriptase (Takara, Clontech, Palo Alto, CA) or Quantiscript Reverse Transcriptase (QuantiTect Reverse Transcription Kit, Qiagen) using random hexamer primers and according to the manufacturer instructions. Real-time PCR analysis was performed on a 7300 Real Time PCR system (Applied Biosystems, Foster City, CA), using TaqMan Gene Expression Assays-on Demand (Applied Biosystems). The cDNA equivalent to 1-25 ng of input RNA was analyzed one to three times in duplicates or triplicates. Real-time PCR was performed with cDNA in 25 μl using 96-well plates. In addition, real-time RT-PCR reaction was verified in 2% agarose 0.5x TBE gel electrophoresis. As control materials CD45.2 healthy animals (bone marrow pooled from 20 mice) or CD45.1 mice transplanted with only competitor cells CD45.1 were taken. Relative quantification of a target gene transcript in comparison to a reference β-Actin transcript was calculated using the method described by Pfaffl.[53]

Locus specific qPCR amplification of 5'-LTR –flanking sequence.

5 to 60 ng of genomic DNA were amplified by using a forward 5'insertion-specific primer (see table below) with 5'LTR specific reverse primer by 40-cycle qPCR on an Applied Biosystems StepOnePlus System (Foster City, CA) using the Quantitect SYBR Green Kit (Qiagen, Hilden, Germany). Relative quantification of a target gene amplicon was estimated in comparison to amplicon at early time points.

The locus specific signal was normalized by the signal of a housekeeping gene Flk-1 (accession number NT_039306). To design the locus specific primers the Primer3Plus program (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi) was used. Real-time PCR efficiencies were determined for each set of designed primers for selected target genes by using the method described by Pfaffl.[53]

Sequence of primers used for locus specific qPCR.

Primer	Name	Sequence (5' – 3')
RV1	5'LTR uni	AGCTAGCTTGCCAAACCTACA
RV2	Evi1#12	CAAACCTACAGGTGGGGTCT
RV3	Flk1	GGTTTCAATGTCCCGTATCCTT
FW1	Mafg#11	GAATCCCCACTTAGCCCAAC
FW2	5330437I02Rik#11	TGGGGGTTGCACCTAGAGTA
FW3	Evi1#12	AGGGCTTGGGACGACTACTT
FW4	Ak2#12	CCCTTGCCAACTTCACCTAC
FW5	Brd7#12	TGTTTCCAGCCTTTGAATCC
FW6	Sox7#15	AGTCGGAAGAAGCACCAGAA
FW7	Olfr56/Ifi47#15	CCCTCACTAACCCCACTCCT
FW8	LOC668840#15	TGTTAGCAGATAACGGGGAAA
FW9	Evi1#32	CCCAGTTGATCCGTTTCAGT
FW10	Smad3#32	GAGAGAGAGAATCTTTGTCAGCA
FW11	Edem2#32	GGTGGTTTCCTTTCGTCCTT
FW12	Mrpl1#32	GTAGGCAGAGGCCCGTAAGT
FW13	Flk1	CTTTGCCCCAGTCCCAGTTA

Supplementary References

51. Schmidt M, Hoffmann G, Wissler M, et al. (2001) Detection and direct genomic sequencing of multiple rare unknown flanking DNA in highly complex samples. *Hum Gene Ther*.**12**:743-749.
52. Kustikova OS, Baum C, Fehse B. (2008) Retroviral integration site analysis in hematopoietic stem cells. *Methods Mol Biol* **430**:255-267.
53. Pfaffl MW. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* **29**:e45..