Supplementary Table S3: Insertion sites recovered from gamma-retroviral LTHSC, STHSC and MPP EGFP progeny (DBISA)

Mouse ID 1° BMT	weeks after BMT	Source	locus: within gene or nearest gene	RTCGD (hits N°)	gene ID	chromo-some	(proposed) function	pos. to TSS (i, intron; e, exon)	orien-tation
LT#26									
#26	12, 19	PB	Dnm2	0	13430	9A3	dynamin 2: GTP binding / cytoskeleton	38118 bp (i2)	R
#26	26, 31	PB, BM, Spl	Samm50	0	68653	15E2	sorting and assembly machinery component 50 homolog (S. cerevisiae)	16492 bp (i12)	R
ST#31									
#31	26, 31	PB, BM, Spl, Liv	Nln	0	75805	13 D1	neurolysin (metallopeptidase M3 family)	64844 bp (i9)	F
#31	31	BM	D030011O10Rik	1	320560	6 G3	RIKEN cDNA D030011O10 gene: domains involved in Rab-, Ras-signaling	6438 bp (i1)	R
#31	26, 31	PB, BM	*Bach2*	14	12014	4 A5	BTB and CNC homology 2: neuronal cells differentiation	113297 bp (i2)	R
#31	31	liver	Txnip*	1	56338	3 F2	thioredoxin interacting protein: stress-responsive protein	-2676 bp	F
#31	31	liver	Eno1*	0	13806	4 E2	enolase 1, alpha non-neuron: glycolysis	-788 bp	F
#31	31	liver	Reep3	2	28193	10 B5.1	receptor accessory protein 3	-7123 bp	R
#31	26	PB	LOC100040414	0	100040414	1 D	similar to ribosomal protein L30	6897 bp	F
ST#32									
#32	31	PB, BM	Mrpl1	0	94061	5 E3	mitochondrial ribosomal protein L1	-811 bp	F
#32	12, 31	PB, BM	Edem2* / Procr*	0 / 0	108687 / 19124	2 H1	ER degradation enhancer, mannosidase alpha-like 2 / protein C receptor, endothelial	-8598 bp / -13368 bp	F / R
#32	12, 26, 31	PB, BM, Spl	Smad3	7	17127	9 C	MAD homolog 3 (Drosophila): TF	7426 bp (i1)	R
#32	6, 12, 26, 31	PB, BM, Spl	LOC100040799 / Ppp2r5c	0 / 3	100040799 / 26931	12 F1	hypothetical protein / protein phosphatase 2, regulatory subunit B, gamma isoform	14099 bp / -25348 bp	F / F
#32	12, 19, 31	PB, BM	Evi1	25	14013	3A3	ecotropic viral integration site 1	-69704 bp	R
#32	6, 31	PB, Spl	LOC665928* / Bst2*	0 / 0	665928 / 69550	8 B3.3	hypothetical protein / bone marrow stromal cell antigen 2	9608 bp / 10217 bp	R / F
#32	6	PB	9930012K11Rik* / Pdlim2*	0 / 0	268759 / 213019	14 D2	RIKEN cDNA 9930012K11 gene / PDZ and LIM domain 2	6053 bp (i1) / 17 kb	R / R
#32	6	PB	Efna1*	0	13636	3 F1	ephrin A1: signaling	-3658 bp	R
#32	6	PB	Prx*	0	19153	7 A3	periaxin: signaling	7090 bp (i2)	R
#32	6	PB	Chsy1	0	269941	7 C	carbohydrate (chondroitin) synthase 1	9322 bp (i1)	F
#32	6	PB	*Egfl7 / Notch1*	2 / 33	353156 / 18128	2 A3	EGF-like domain 7 / Notch gene homolog 1 (Drosophila)	-3052 bp / -74 kb	R / F
ST#33									
#33	31	Spl	EG668996*	0	668996	X A1.3	predicted gene	-303 bp	R
#33	31	Spl	Wdr26	2	226757	1 H4	WD repeat domain 26	1188 bp (i1)	R
#33	31	Spl	LOC665939 / Ifngr2	0 / 4	665939 / 15980	16 C3.3	similar to cell division cycle 20 homolog (S. cerevisiae) / interferon gamma receptor 2	8728 bp (i1) / -24544 bp	F / R
#33	31	Spl	Msi2	3	76626	11 B5-C	Musashi homolog 2: proliferation and maintenance of CNS stem cell populations	5946 bp (i1)	R
ST#34									
#34	31	BM	LOC100043676 / Ptk2b	0 / 1	100043676 / 19229	14 D1	hypothetical protein LOC100043676 / PTK2 protein tyrosine kinase 2 beta	12393 bp / 18 kb (i1)	R / R
#34	31	BM	Acot7*	2	70025	4 E2	acyl-CoA thioesterase 7: metabolic	12855 bp (i1)	F
#34	31	PB	EG381483	0	381483	3 H3	predicted gene	155033 bp	F
#34	31	Spl	Tnni2	1	21953	7 F5	troponin I, skeletal, fast 2	559 bp (i1)	F

#34	31	Spl	Sox4	75	20677	13 A3-A5	SRY-box containing gene 4: TF	2810 bp (e1)	F
#34	31	Spl	Pde4d	2	238871	13 D2.1-D2.2	phosphodiesterase 4D, cAMP specific: signaling	991750 bp (i44)	R
#34	31	Spl	Phf17	1	269424	3 B	PHD finger protein 17	-6093 bp	F
#34	6	PB	Dusp5	1	240672	19 D2	dual specificity phosphatase 5: MAPK signaling pathway	28600 bp	F
#34	6	PB	Tagln2*	2	21346	1 H3	transgelin 2: actin-associated protein / muscle development	1224 bp (i1)	F
#34	6	PB	Jun	1	16476	4 C5-C7	Jun oncogene: TF	-105551 bp	R
#34	6	PB	Oscar*	1	232790	7 A1	osteoclast associated receptor	6512 bp	R
#34	12	PB	EG381483	0	381483	3 H3	predicted gene	155033 bp	F
#34	12	PB	LOC666619	0	666619	9 A4	similar to ribosomal protein L21	13255 bp	F
#34	12	PB	Ubl7 / Sema7a	3 / 3	69459 / 20361	9 C	ubiquitin-like 7 / semaphorin7A	19 kb (i9) / -11587 bp	R / R
#34	12	PB	Rarg*	3	19411	15 E-F3	retinoic acid receptor, gamma	1 kb	F
#34	12	PB	Cenpj	0	219103	14 C3	centromere protein J	3913 bp (i1)	F
#34	19	PB	LOC666619	0	666619	9 A4	similar to ribosomal protein L21	13252 bp	R
#34	19	PB	Arrb1	0	109689	7 E2	arrestin, beta 1: MAPK signaling pathway	10 kb (i1)	R
#34	19	PB	BC030046*	0	277154	14 C3	cDNA sequence BC030046: pol-like protein	-485 bp	R
#34	19	PB	Ptk2b	1	19229	14 D1	PTK2 protein tyrosine kinase 2 beta	19 kb (i1)	R
#34	19	PB	Recql5	1	170472	11 E2	RecQ protein-like 5: DNA repair	21 kb (i7)	F
ST#35									
#35	31	BM, Spl	Galnt2	2	108148	8 E2	UDP-N-acetyl-alpha-D-galactosamine	95336 bp (i6)	R
#35	31	PB	Arid1a	5	93760	4 D3	AT rich interactive domain 1A (Swi1 like)	-62134 bp	R
#35	31	PB	Tcfec	0	21426	6 E4-G	transcription factor EC	2418 bp (i1)	F
#35	31	Spl	Zfp217	6	228913	2 H3	zinc finger protein 217	-29735 bp	F
#35	31	Spl	B930007M17Rik	1	381476	3 H1	RIKEN cDNA B930007M17 gene	41752 bp (i6)	R
MPP#36									
# 36	26	PB	Ablim1	2	226251	19 D2	actin-binding LIM protein 1	96551 bp (i6)	R
# 36	26	PB	EG628871	0	628871	X A1.3	predicted gene, EG628871	71615 bp	F
# 36	26	PB	Trim16*	0	94092	11 B2	tripartite motif-containing 16: zink finger	-719 bp	R
# 36	24	PB	P2ry6*	0	233571	7 E3	pyrimidinergic receptor P2Y, G-protein coupled, 6	14349 bp (i1)	R
MPP#37									
#37	31	PB, BM, Spl	Mgll	1	23945	6 D1	monoglyceride lipase	86 kb (i5)	R
#37	26, 31	PB	1810011O10Rik	1	69068	8 A2	RIKEN cDNA 1810011O10 gene: apoptosis	66883 bp	R
#37	26	PB	Aff4*	1	93736	11 B1.3	AF4/FMR2 family, member 4: TF	25902 bp (i4)	R
#37	26	PB	LOC666495	0	666495	2 A1	similar to Presenilin associated, rhomboid-like	-21813 bp	R
#37	26	PB	Serpinb1a*	0	66222	13 A4	serine (or cysteine) peptidase inhibitor, clade B, member 1a	18224 bp	F
#37	26	PB	Ctsg*	0	13035	14 C3	cathepsin G: proteolysis	625 bp (i2)	R
#37	26	PB	LOC667119 / Fos	0 / 6	667119 / 14281	12 D2	hypothetical LOC667119 / FBJ osteosarcoma oncogene	-15359 bp / -57393 bp	R / R

MPP#38									
#38	26, 31	PB, Liv	Snx31	0	66696	15 C	sorting nexin 31: signaling	-8114 bp	R
#38	26	PB	*Bach2*	14	12014	4 A5	BTB and CNC homology 2: neuronal cells differentiation	132926 bp (i2)	F
#38	26, 31	PB, Spl, Liv	D16Ertd472e	3	67102	16 C3.1	DNA segment, Chr 16, ERATO Doi 472, expressed	73231 bp	R
#38	26	PB	Riok2	0	67045	17 A3.1	RIO kinase 2 (yeast): signal transduction	4266 bp (i1)	F
#38	26	PB	LOC100041377*	0	100041377	18 B1	hypothetical protein LOC100041377	3271 bp	F
#38	31	BM	Mtch2*	0	56428	2 E1	mitochondrial carrier homolog 2 (C. elegans)	-955 bp	F
#38	31	BM	Cradd	0	12905	10 C2	CASP2 and RIPK1 domain containing adaptor with death domain	187 kb	F
#38	31	BM	Slc22a16	0	70840	10 B1	solute carrier family 22 (organic cation transporter), member 16	-56050 bp	F
#38	31	PB, BM, Spl, Liv	*Egfl7 / Notch1*	2 / 33	353156 / 18128	2 A3	EGF-like domain 7 / Notch gene homolog 1 (Drosophila)	- 10411 bp / - 66 kb	F / R
#38	31	Spl	2310033K02Rik	0	69553	2 H3	RIKEN cDNA 2310033K02 gene	-450 bp	F
MPP#39									
#39	26	PB	Irf1	1	16362	11 B1.3	interferon regulatory factor 1: TF	-7541 bp	R
#39	26	PB	Cenpc1*	0	12617	5 E2-E5	centromere protein C1	-491 bp	F
#39	26	PB	Il4ra	3	16190	7 F3	interleukin 4 receptor, alpha: hematopoietic cell lineage	30973 bp	F
#39	26	PB	6720432D03Rik	0	77740	9 A5.1	RIKEN cDNA 6720432D03 gene	-21626 bp	F
#39	26	PB	Cd47	3	16423	16 B5	CD47 antigen (Rh-related antigen, integrin-associated signal transducer)	-17521 bp	R
#39	26	PB	Raf1	1	110157	6 E3	v-raf-leukemia viral oncogene 1	-19821 bp	F
#39	31	BM	LOC100039985	0	100039985	3 A1	hypothetical protein LOC100039985	30845 bp	R
#39	31	BM	Ptpn1	1	19246	2 H3	protein tyrosine phosphatase, non-receptor type 1: insulin receptor signalling pathway	2995 bp (i1)	R
#39	31	BM	H3f3a / BC031781	0 / 4	15078 / 208768	1 D2.3	H3 histone, family 3A / cDNA sequence BC031781	-883 bp / -35 kb	F / R
#39	31	PB	Mfhas1	0	52065	8 A4	malignant fibrous histiocytoma amplified sequence 1: Ras_like_GTPase	62 kb (i1)	F
#39	31	PB	Wdr42a*	0	98193	1 H3	WD repeat domain 42A	1030 bp (i1)	R
#39	31	Spl	Mdfic	0	16543	6 A1	MyoD family inhibitor domain containing: transcription	74202 bp (i3)	F
#39	31	Spl	Med27	0	68975	2 A3	mediator complex subunit 27: transcription	108 kb (i3)	R
MPP#40									
#40	26	PB	Centa1	0	231821	5 G2	centaurin, alpha 1: pleckstrin homology (PH) domain	8487 bp (i3)	R
#40	31	Spl	LOC546780	0	546780	2 H3	similar to ribosomal protein L9	-4936 bp	R
#40	31	Spl	Cbfa2t2	0	12396	2 C3.2	core-binding factor, runt domain, alpha subunit 2, translocated to, 2 (human): TF	-9182 bp	R
#40	31	Spl	LOC100042432* / Ptpn 7*	0 / 2	100042432 / 320139	1 E4	hypothetical protein / protein tyrosine phosphatase, non-receptor type 7	1 kb / -2kb	F / R

Insertional coordinates are given based on the NCBI database, January-July 2008; RTCGD membership is given based on the database's information accessible between January-July, 2008; PB, total peripheral blood; BM, total bone marrow; Spl, spleen; Liv, liver; TSS, transcriptional start site (NCBI definition); *, located in gene dense region: more than 3 genes within 50 kb from each RVIS side; in italics, common locus for retroviral integration sites within this set of data; n.a., not applicable.