Supporting Information

Miccio *et al.* 10.1073/pnas.0711666105

SI Materials and Methods

Protein Analysis. Globin polypeptide chain composition was determined by HPLC as described (1). Hemoglobins in peripheral blood collected from transplanted mice were resolved by cellulose acetate electrophoresis as previously described (2).

Antibody Staining for FACS Analysis and Sorting. RBCs were fixed in 0.05% glutaraldehyde for 10 min at room temperature, washed once with PBS, and permeabilized in 0.1% Triton X-100. Cells were incubated with anti-CD16 antibody (PharMingen) and 1% FBS (FBS) to block unspecific binding, stained with phycoerythrin (PE)-conjugated anti-human β-globin antibody (SC-21757, Santa Cruz Biotechnology), and analyzed by FACScan employing the Cell Quest System (Becton Dickinson). Splenic and BM cells were incubated with anti-CD16 antibody and 1% FBS, stained with a PE-conjugated anti-mouse Ter119 antibody (PharMingen), and analyzed by FACScan. To purify Ter119$^+$ cells, BM cells were stained with PE-conjugated anti-mouse Ter119 antibody, magnetically labeled with anti-PE Microbeads (Miltenyi Biotec), and separated by using the Mini MACS separator (Miltenyi Biotec).

1. Schroeder WA, Shelton JB, Shelton JR, Huynh V, Teplow DB (1985) High performance liquid chromatographic separation of the globin chains of non-human hemoglobins. *Hemoglobin* 9:461–482.
2. May C, *et al.* (2000) Therapeutic haemoglobin synthesis in beta-thalassaemic mice expressing lentivirus-encoded human beta-globin. *Nature* 406:82–86.

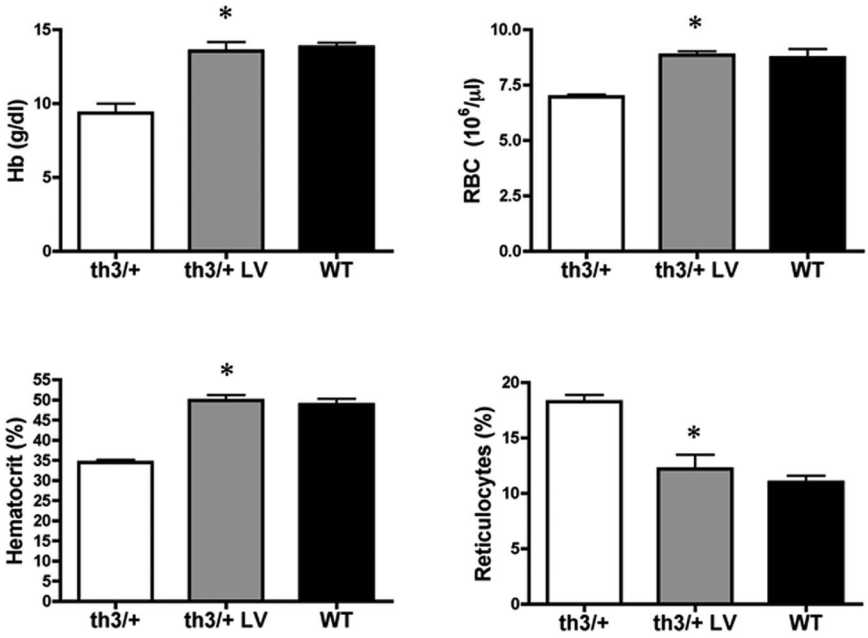

Fig. S1. Correction of hematological parameters in secondary recipients transplanted with transduced BM. Hemoglobin (Hb) concentration, RBC count, hematocrit, and percentages of reticulocytes in peripheral blood of secondary recipients transplanted with HSCs from primary recipients of mock-transduced *th3*/+ (white column; $n = 4$), normal (WT; black column; $n = 3$), and transduced *th3*/+ BM cells (*th3*/+ LV; gray column; $n = 5$), were measured 6 months after BM transplantation. Data represent the mean values and SEM for each group of animals. Hematological parameters are significantly different in recipients transplanted with transduced HSCs vs. control *th3*/+ mice (Hb: $P < 0.05$; RBC count: $P < 0.01$; hematocrit: $P < 0.01$; reticulocyte counts: $P < 0.05$). No statistically significant differences were found compared with mice transplanted with normal cells.

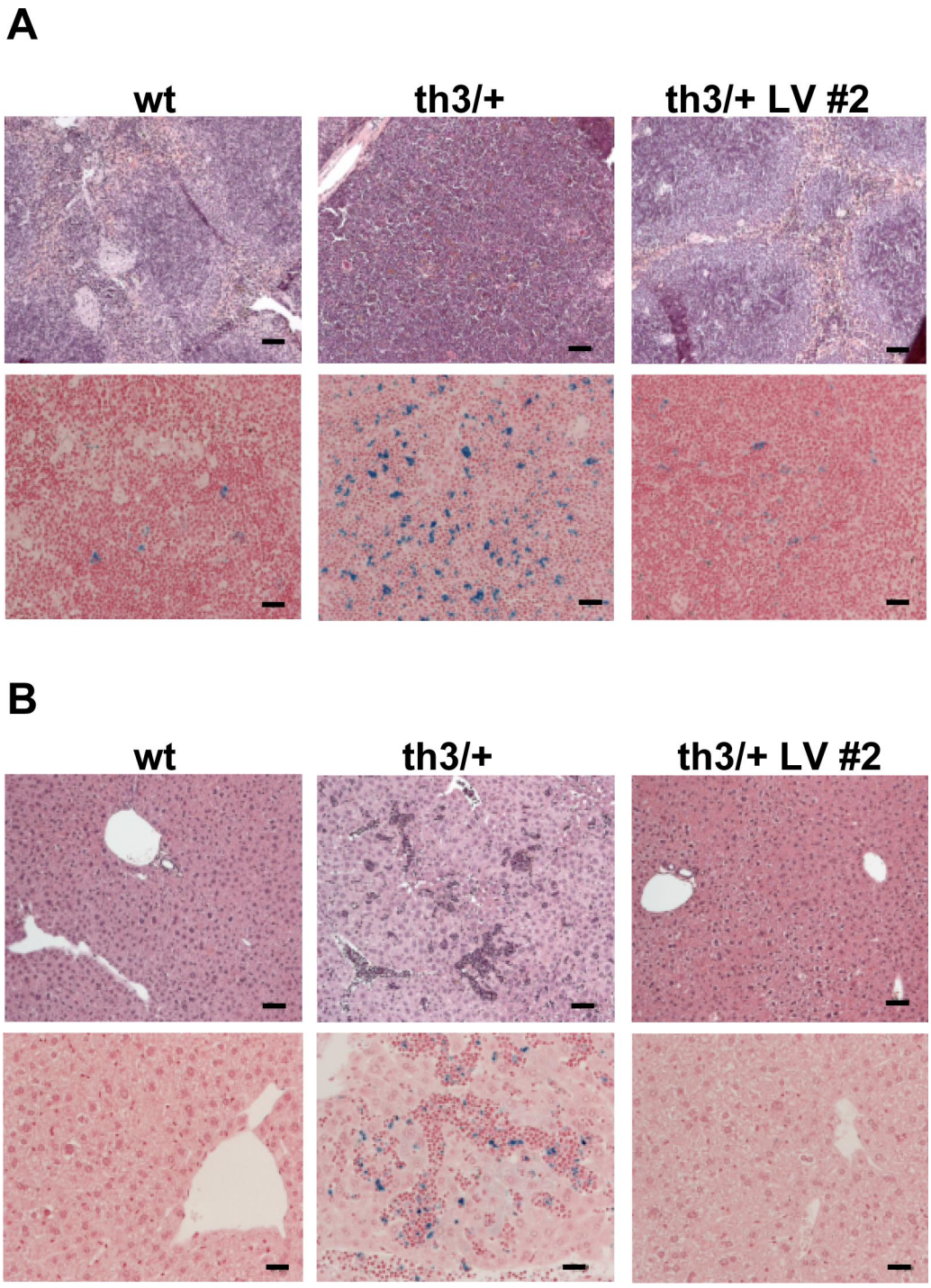

Fig. S2. Correction of spleen and liver pathology in recipients of transduced HSCs. (*A*) (*Upper*) H&E staining of splenic sections of mice transplanted with mock-transduced normal (WT), *th3*/+, and transduced *th3*/+ cells (*th3*/+ LV #2). Treated animals show a restoration of splenic architecture comparable to that of normal controls. (Original magnification ×10.) (*Lower*) Determination of the iron deposition extension (Perls staining) in splenic sections. *th3*/+ controls displayed numerous and large iron particles, whereas only trace amounts of iron, comparable to normal, were found in treated mice. (Original magnification ×20.) (*B*) (*Upper*) EMH determination (H&E staining) in liver sections. Many foci of EMH were distinguishable in *th3*/+ controls. Conversely, no evidence of EMH was found in recipients transplanted with transduced HSCs. (*Lower*) Absence of iron deposition in the liver of mice transplanted with transduced *th3*/+ cells. Large iron deposits can be readily detected in *th3*/+ control mice. (Original magnification for all ×20.)

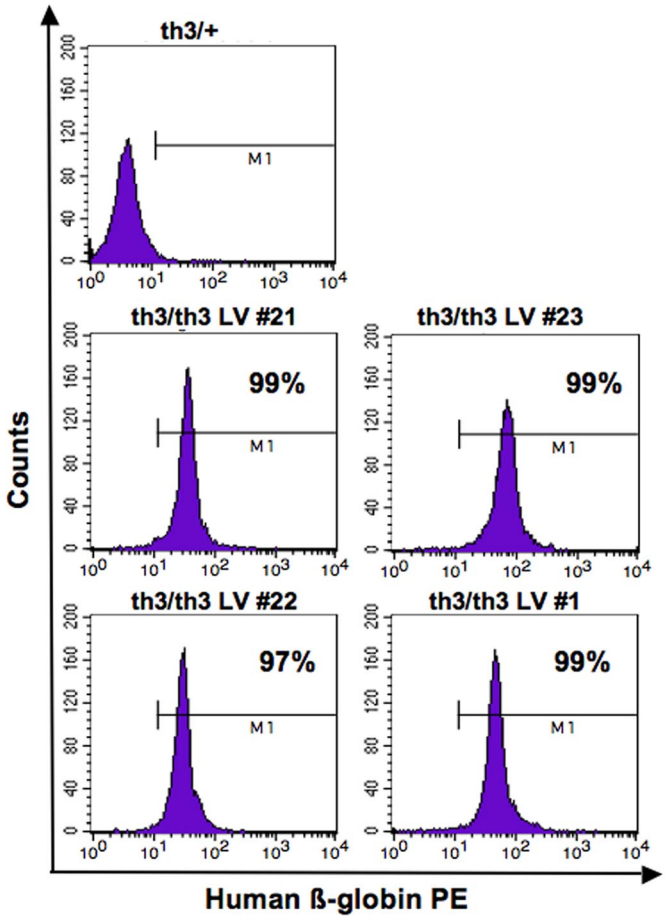

Fig. S3. GLOBE LV-derived β-globin expression in *th3/th3* hematopoietic chimeras. FACS analysis of peripheral blood samples collected 6 months after transplantation from mice transplanted with transduced *th3/th3* FLCs (*th3/th3* LV mice #21, #23, #22, and #1) and control mice transplanted with mock-transduced *th3*/+ cells. RBCs were stained by using an antibody that specifically recognizes the human β-globin protein and were analyzed by flow cytometry.

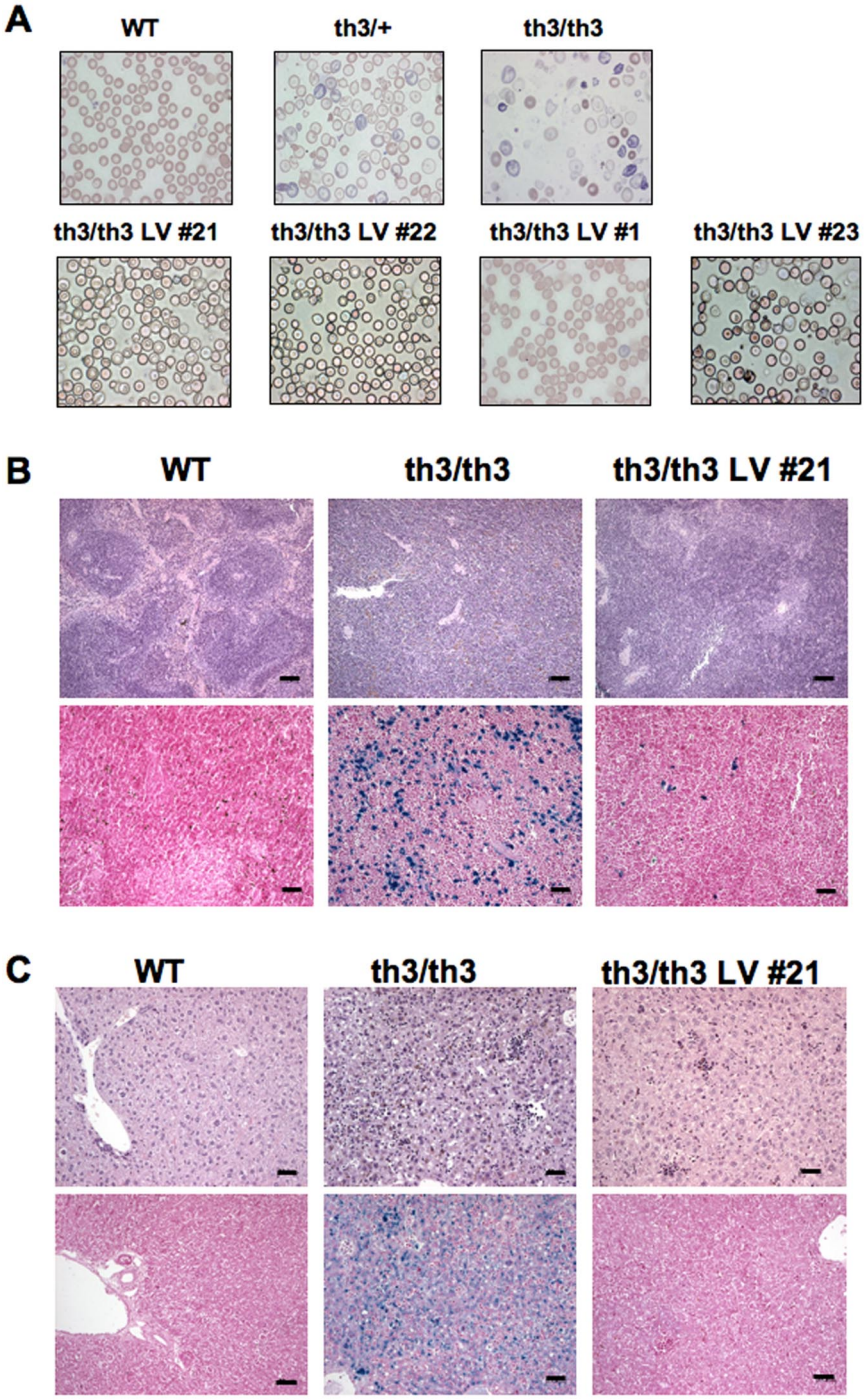

Fig. S4. Amelioration of RBC morphology and correction of organ pathology in long-term *th3/th3* chimeras. (*A*) May–Grunwald/Giemsa staining of blood smears from mice transplanted with mock-transduced normal (WT), *th3*/+, *th3/th3*, and transduced *th3/th3* (*th3/th3* GLOBE) cells was performed 7–12 months after transplantation except for mice transplanted with mock-transduced *th3/th3* FLCs, which were killed between 30 and 80 days after transplantation when profoundly anemic. Abnormal anisocytosis and poikilocytosis were strongly ameliorated in recipients of transduced FLCs. (Original magnification ×40.) (*B*) (*Upper*) H&E staining of splenic sections of mice transplanted with mock-transduced normal, *th3/th3*, and transduced *th3/th3* FLCs (*th3/th3* LV #21). Splenic red pulp was highly expanded in *th3/th3* controls, whereas treated animals showed restoration of splenic architecture. (Original magnification ×10.) (*Lower*) Determination of the iron deposition extension (Perls staining) in the spleen of mice transplanted with mock-transduced WT, *th3/th3*, and transduced *th3/th3* FLCs (*th3/th3* LV #21). Mice transplanted with *th3/th3* cells displayed high levels of iron deposition, while those receiving transduced HSCs showed reduced numbers of iron deposits. (Bars: 100 μm; original magnification ×20.) (*C*) (*Upper*) EMH determination (H&E staining) in liver sections of mice transplanted with mock-transduced normal, *th3/th3*, and transduced *th3/th3* cells (*th3/th3* LV #21). The liver of mice transplanted with *th3/th3* FLCs showed abundant foci of EMH. Conversely, a low level of EMH was found in recipients of transduced HSCs. (*Lower*) Reduction of iron deposition (Perls staining) in the liver of mice transplanted with transduced *th3/th3* cells (*th3/th3* LV #21). (Original magnification ×10.)

Table S1. Correction of thalassemic phenotype in mice transplanted with transduced BM HSCs

Mouse no.	% huβ^+ RBCs	% huβ/mα	Hb, g/dl	RBCs $\times 10^6/\mu$l	Hct, %	% Reticulocytes	% BM Ter119$^+$ cells	% Splenic Ter119$^+$ cells	Splenic iron deposits	Liver EMH	Liver iron deposits	VCN
14	64	ND	9.69	7.72	37	18.9	55	26	3	1	2	0.23
19	99	19	12.48	8.40	42	12.3	46	8	2	0	1	0.91
21	95	19	14.01	10.99	53	7.7	46	8	2	0	0	0.74
4	99	23	12.30	10.16	48	11.8	42	11	1	0	0	0.55
6	86	14	9.40	7.51	36	20.8	48	21	2	0	2	0.27
5	99	37	13.86	9.99	51	11.8	37	5	3	0	0	0.83
2	95	30	13.20	9.30	48	4.1	26	3	1	0	0	2.20
55	96	ND	11.65	9.15	43	15.5	42	10	1	0	0	0.55
th3/+	NA	NA	8.03 ± 0.39	5.89 ± 0.17	29 ± 1	25.8 ± 2.0	60 ± 3	47 ± 8	2–3	0–3	0–3	NA
WT	NA	NA	13.01 ± 0.87	8.97 ± 0.29	46 ± 2	5.6 ± 0.7	37 ± 2	6 ± 2	0–2	0–1	0–1	NA

NA, not applicable; ND, not determined.

Table S2. Correction of organ pathology in *th3*/+ mice transplanted with transduced BM cells

Mice*	% BM Ter119$^+$ cells†	% Splenic Ter119$^+$ cells†	Splenic iron deposits‡	Liver EMH§	Liver iron deposits‡
WT ($n = 5$)	37.13±1.78	5.88±1.61	0–2	0–1	0–1
th3/+ ($n = 5$)	60.±2.9	47.13±8	2–3	0–3	0–3
th3/+ LV ($n = 8$)	43.5±2.92	11.50±2.81	1–3	0–1	0–2

*Irradiated mice were transplanted with mock-transduced normal (WT) or *th3*/+ cells, and with GLOBE-transduced (*th3*/+ LV) cells, respectively.
†percentage of Ter119$^+$ cells, as determined by FACS staining, after *in silico* gating on mononucleated cells ($P = 0.003$ *th3*/+ vs. *th3*/+ GLOBE; $P > 0.05$ WT vs. *th3*/+ LV)
‡The amount of iron deposition was scored semiquantitatively on a scale of 0 (no iron present) to 3 (maximum amount of iron identified in the organ).
§The extent of EMH was evaluated semiquantitatively on a scale of 0 (no EMH present) to 3 (maximum extent of EMH identified in the organ).

Table S3. List of proviral integration sites

Mouse no.	Chromosomal band	Target gene	Location	RefSeq no.	Biological function
Vector: GLOBE; Target: BM cells from the *th3*/+ transplanted mice					
4	chr8 qE1	Usp10	intron 1	NM_009462.1	ubiquitin specific peptidase 10
	chrX qA5	Hs6st2	intron 1	NM_015819.2	heparan sulfate 6-O-sulfotransferase 2
5	chr9 qB	Neo 1	intron 7	NM_008684.1	neogenin
	chr9 qE1		intergenic		
	chr16 qC3.1		intergenic		
	chrX qF4		intergenic		
6	chr5 qE1		intergenic		
	chr10 qB5.1	D10Ucla1	intron 2	NM_178606.2	DNA segment, Chr 10, University of California at Los Angeles 1
	chr12 qE1		intergenic		
	chr16 qA1	Fgd4	intron 1	NM_139232.1	FYVE RhoGEF and PH domain containing 4 isoform
	chrX qE3	Timm8a	downstream	NM_013898.1	translocase of inner mitochondrial membrane 8
19	chr1 qD		intergenic		
	chr12 qC1	Lrfn5	downstream	NM_178714.2	leucine rich repeat and fibronectin type III domain containing 5
	chr17 qB1	9130008F23Rik	exon 2	NM_027834.1	RIKEN cDNA 9130008F23 gene
	chr18 qB3		intergenic		
2	chr3 qG1	Usp53	intron 13	NM_133857.1	ubiquitin specific peptidase 53
	chr6 qA1	Foxp2	upstream	NM_053242.4	forkhead box P2
	chr9 qA1	Alkbh8	intron 2	NM_026303.1	alkB, alkylation repair homolog 8 (E. coli)
	chr9 qF4	Snrk	upstream	NM_133741.1	SNF related kinase
	chrX qD	P2ry10	intron 2	NM_172435.2	purinergic receptor P2Y, G-protein coupled 10
55	chr2 qC1.1		intergenic		
	chr2 qH3	Eya2	intron 1	NM_010165.2	eyes absent 2 homolog
	chr3 qG1	Ndst4	intron 4	NM_022565.1	N-deacetylase/N-sulfotransferase
	chr3 qE3	Golph4	downstream	NM_175193.4	golgi phosphoprotein 4
	chr5 qB3		intergenic		
	chr6 qA1		intergenic		
	chr6 qA1	Peg10	intron 1	NM_130877.2	paternally expressed 10 isoform RF1/RF2
	chr6 qA2		intergenic		
	chr7 qE3	Gvin1	intron 1	NM_001039160.2	very large inducible GTPase 1
	chr8 qB3.1	Tll1	intron 6	NM_009390.2	tolloid-like
	chr10 qD2		intergenic		
	chr11 qA1	Sfi1	intron 20	NM_030207.2	Sfi1 homolog spindle assembly associated
	chr18 qA2		intergenic		
21	chr1 qA2	Mybl1	intron 4	NM_008651.2	myeloblastosis oncogene-like 1
	chr1 qB		intergenic		
	chr3 qA3		intergenic		
	chr10 qB4	Pkib	intron 2	NM_008863.3	cAMP-dependent protein kinase inhibitor beta
	chr10 qD2	Cpsf6	intron 3	NM_001013391.1	cleavage and polyadenylation specific factor 6
	chr11 qA5	Mat2b	intron 1	NM_134017.1	methionine adenosyltransferase II beta
	chr12 qD3 (1)		intergenic		
	chr12 qD3 (2)		intergenic		
	chr15 qB3.2		intergenic		
	chr19 qA	Ehd1	intron 2	NM_010119.3	EH-domain containing 1
	chrX qF5	Glra2	intron 4	NM_183427.2	glycine receptor, alpha 2 subunit
14	chr2 qB	Dennd1a	intron 2	NM_146122.2	DENN/MADD domain containing 1A
	chr2 q C1.1		intergenic		
	chr2 qC1.3		intergenic		
	chr3 qA2		intergenic		
	chr4 qA1	Gem	3′UTR	NM_010276.1	GTP binding protein(gene overexpressed in skeletal muscle)
	chr6 qA1		intergenic		
	chr7 qD3	Rab38	intron 1	NM_028238	Rab38, member of RAS oncogene family
	chr8 qB3.1		intergenic		
	chr9 qA1	Dync2h1	exon>10	NM_029851.2	dynein cytoplasmic heavy chain 2
	chr10 qD1	Osbpl8	intron 1	NM_001003717.1	oxysterol-binding protein-like protein 8 isoform
	chr12 qC3	Ppp2r5e	intron 3	NM_012024.2	epsilon isoform of regulatory subunit B56
	chr15 qA1	Lmbrd2	intron 4	NM_177178.3	LMBR1 domain containing 2
	chr15 qF3	Nckap1l	intron 21	NM_153505.2	NCK associated protein 1 like
	chr16 qA1	2310008H04Rik	intron 1	NM_146068	hypothetical protein LOC224008
	chr17 qE1.1	Arhgap28	intron 1	NM_172964	Rho GTPase activating protein 28

Mouse no.	Chromosomal band	Target gene	Location	RefSeq no.	Biological function
Vector: GLOBE; Target: Ter119^{+} BM cells from *th3*/+ mice					
4	chr8 qE1	Usp10	intron 1	NM_009462.1	ubiquitin specific peptidase 10
	chrX qA5	Hs6st2	intron 1	NM_015819.2	heparan sulfate 6-O-sulfotransferase 2
5	chr16 qC3.1		intergenic		
	chrX qF4		intergenic		
6	chr5 qE1		intergenic		
	chr12 qE1		intergenic		
19	chr12 qC1	Lrfn5	downstream	NM_178714.2	leucine rich repeat and fibronectin type III domain containing 5
	chr17 qB1	9130008F23Rik	exon 2	NM_027834.1	RIKEN cDNA 9130008F23 gene
	chr18 qB3		intergenic		
2	chr3 qG1	Usp53	intron 13	NM_133857.1	ubiquitin specific peptidase 53
	chr6 qA1	Foxp2	upstream	NM_053242.4	forkhead box P2
	chr9 qA1	Alkbh8	intron 2	NM_026303.1	alkB, alkylation repair homolog 8 (E. coli)
	chr9 qF4	Snrk	upstream	NM_133741.1	SNF related kinase
	chrX qD	P2ry10	intron 2	NM_172435.2	purinergic receptor P2Y, G-protein coupled 10
55	chr2 qC1.1		intergenic		
	chr3 qG1	Ndst4	intron 4	NM_022565.1	N-deacetylase/N-sulfotransferase
	chr5 qB3		intergenic		
	chr6 qA1		intergenic		
	chr6 qA2		intergenic		
	chr10 qD2		intergenic		
	chr11 qA1	Sfi1	intron 20	NM_030207.2	Sfi1 homolog spindle assembly associated
	chr18 qA2		intergenic		
21	chr1 qB		intergenic		
	chr3 qA3		intergenic		
	chr11 qA5	Mat2b	intron 1	NM_134017.1	methionine adenosyltransferase II beta
	chr12 qD3 (1)		intergenic		
	chr12 qD3 (2)		intergenic		
	chr15 qB3.2		intergenic		
14	chr2 q C1.1		intergenic		
	chr6 qA1		intergenic		
	chr8 qB3.1		intergenic		
	chr9 qA1	Dync2h1	exon>10	NM_029851.2	dynein cytoplasmic heavy chain 2
	chr15 qA1	Lmbrd2	intron 4	NM_177178.3	LMBR1 domain containing 2
	chr15 qF3	Nckap1l	intron 21	NM_153505.2	NCK associated protein 1 like
	chr17 qE1.1	Arhgap28	intron 1	NM_172964	Rho GTPase activating protein 28
Vector: GLOBE; Target: Ter119^{-} BM cells from *th3*/+ mice					
4	chr8 qE1	Usp10	intron 1	NM_009462.1	ubiquitin specific peptidase 10
	chrX qA5	Hs6st2	intron 1	NM_015819.2	heparan sulfate 6-O-sulfotransferase 2
5	chr9 qB	Neo 1	intron 7	NM_008684.1	neogenin
	chr9 qE1		intergenic		
6	chr5 qE1		intergenic		
	chr10 qB5.1	D10Ucla1	intron 2	NM_178606.2	DNA segment, Chr 10, University of California at Los Angeles 1
	chr12 qE1		intergenic		
	chr16 qA1	Fgd4	intron 1	NM_139232.1	FYVE RhoGEF and PH domain containing 4 isoform
	chrX qE3	Timm8a	downstream	NM_013898.1	translocase of inner mitochondrial membrane 8
19	chr12 qC1	Lrfn5	downstream	NM_178714.2	leucine rich repeat and fibronectin type III domain containing 5
	chr17 qB1	9130008F23Rik	exon 2	NM_027834.1	RIKEN cDNA 9130008F23 gene
2	chr3 qG1	Usp53	intron 13	NM_133857.1	ubiquitin specific peptidase 53
	chr6 qA1	Foxp2	upstream	NM_053242.4	forkhead box P2
	chr9 qA1	Alkbh8	intron 2	NM_026303.1	alkB, alkylation repair homolog 8 (E. coli)
	chr9 qF4	Snrk	upstream	NM_133741.1	SNF related kinase
	chrX qD	P2ry10	intron 2	NM_172435.2	purinergic receptor P2Y, G-protein coupled 10
55	chr2 qC1.1		intergenic		
	chr2 qH3	Eya2	intron 1	NM_010165.2	eyes absent 2 homolog
	chr3 qG1	Ndst4	intron 4	NM_022565.1	N-deacetylase/N-sulfotransferase
	chr3 qE3	Golph4	downstream	NM_175193.4	Golgi phosphoprotein 4
	chr6 qA1		intergenic		
	chr6 qA1	Peg10	intron 1	NM_130877.2	paternally expressed 10 isoform RF1/RF2
	chr7 qE3	Gvin1	intron 1	NM_001039160.2	very large inducible GTPase 1
	chr8 qB3.1	Tll1	intron 6	NM_009390.2	tolloid-like

Mouse no.	Chromosomal band	Target gene	Location	RefSeq no.	Biological function
	chr10 qD2		intergenic		
	chr11 qA1	Sfi1	intron 20	NM_030207.2	Sfi1 homolog spindle assembly associated
	chr15 qA1	Zfr	intron 9	NM_011767.1	zinc finger RNA binding protein
21	chr1 qA2	Mybl1	intron 4	NM_008651.2	myeloblastosis oncogene-like 1
	chr10 qB4	Pkib	intron 2	NM_008863.3	cAMP-dependent protein kinase inhibitor beta
	chr10 qD2	Cpsf6	intron 3	NM_001013391.1	cleavage and polyadenylation specific factor 6
	chr12 qD3 (1)		intergenic		
	chr12 qD3 (2)		intergenic		
	chr19 qA	Ehd1	intron 2	NM_010119.3	EH-domain containing 1
	chrX qF5	Glra2	intron 4	NM_183427.2	glycine receptor, alpha 2 subunit
14	chr2 qB	Dennd1a	intron 2	NM_146122.2	DENN/MADD domain containing 1A
	chr2 qC1.3		intergenic		
	chr3 qA2		intergenic		
	chr4 qA1	Gem	3′UTR	NM_010276.1	GTP binding protein(gene overexpressed in skeletal muscle)
	chr6 qA1		intergenic		
	chr7 qD3	Rab38	intron 1	NM_028238	Rab38, member of RAS oncogene family
	chr8 qB3.1		intergenic		
	chr10 qD1	Osbpl8	intron 1	NM_001003717.1	oxysterol-binding protein-like protein 8 isoform
	chr12 qC3	Ppp2r5e	intron 3	NM_012024.2	epsilon isoform of regulatory subunit B56
	chr15 qA1	Lmbrd2	intron 4	NM_177178.3	LMBR1 domain containing 2
	chr16 qA1	2310008H04Rik	intron 1	NM_146068	hypothetical protein LOC224008
	chr17 qE1.1	Arhgap28	intron 1	NM_172964	Rho GTPase activating protein 28
Vector: GLOBE; Target: Lin$^-$ BM cells (*in vitro*)					
	chr1 qD		intergenic		
	chr1 qE2.1		intergenic		
	chr1 qH2.2	Kifap3	intron 18	NM_010629.2	kinesin-associated protein 3
	chr2 qC1.1	March7	upstream	NM_020575.2	membrane-associated ring finger (C3HC4) 7
	chr2 qE1	Ttc17	intron 1	NM_183106.2	tetratricopeptide repeat domain 17
	chr2 qE4		intergenic		
	chr2 qF3		intergenic		
	chr3 qE3		intergenic		
	chr4 qA5	Olfr157	intron 4	NM_019475.2	olfactory receptor 157
	chr4 qC1		intergenic		
	chr4 qC5		intergenic		
	chr4 qC6		intergenic		
	chr4 qC7	Btf3l4	intron 3	NM_027453.2	basic transcription factor 3-like 4
	chr5 qA1	Pftk1	intron 4	NM_011074.1	PFTAIRE protein kinase 1
	chr5 qB1	Bre	intron 3	NM_144541.1	brain and reproductive organ-expressed protein
	chr5 qC3.1	Slc30a9	downstream	NM_178651.3	solute carrier family 30 (zinc transporter)
	chr5 qG2	Slc29a4	downstream	NM_146257.2	solute carrier family 29 (nucleoside
	chr6 qA1	EG625347	upstream	NM_001039041.1	hypothetical protein LOC625347
	chr6 qA2	St7	upstream	NM_022332.2	low density lipoprotein-related protein 12
	chr6 qA3.3	Chchd3	intron 4	NM_025336.1	coiled-coil-helix-coiled-coil-helix domain
	chr6 qE3	Gucy1a3	upstream	NM_021896.4	guanylate cyclase 1 soluble alpha 3
	chr6 qE3	Rad18	intron 8	NM_021385.1	RAD18 homolog
	chr6 qF2	Clec4d	upstream	NM_010819.3	C-type lectin superfamily member 8
	chr7 qB4		intergenic		
	chr7 qD1		intergenic		
	chr8 qA1.1		intergenic		
	chr9 qE1	4930486G11Rik	intron 4	NM_175213.3	hypothetical protein LOC75033
	chr9 qF2	Traip	intron 10	NM_011634.3	TRAF-interacting protein
	chr10 qC2	Nedd1	intron 3	NM_008682.1	neural precursor cell expressed developmentally
	chr10 qD1		intergenic		
	chr10 qD3	Lrig3	intron 1	NM_177152.4	leucine-rich and immunoglobulin-like domains 3
	chr11 qB1.2	Olfr56	intron 1	NM_010999.2	olfactory receptor 56
	chr11 qB3	Map2k4	intron 2	NM_009157.3	mitogen activated protein kinase kinase 4
	chr11 qE2	Unk	intron 12	NM_172569.3	zinc finger CCCH type containing 5
	chr12 qA1.3	Taf1b	upstream	NM_020614.1	TATA box binding protein (Tbp)-associated
	chr13 qA1		intergenic		
	chr13 qA5	Aspn	intron 2	NM_025711.2	asporin
	chr13 qC3	Cryba4	downstream	NM_021351.1	crystallin beta A4
	chr13 qC3	Mef2c	intron 2	NM_025282.2	myocyte enhancer factor 2C
	chr13 qC3	Papd4	intron 10	NM_133905.2	PAP associated domain containing 4
	chr14 qB	Gcap14	intron 1	NM_028407.3	granule cell antiserum positive 14 isoform 2

Mouse no.	Chromosomal band	Target gene	Location	RefSeq no.	Biological function
	chr14 qE2.3	Ndfip2	intron 1	NM_029561.2	Nedd4 family interacting protein 2
	chr14 qE4	Slitrk5	downstream	NM_029273.2	SLIT and NTRK-like family member 5 isoform 1
	chr15 qA1	1110020G09Rik	intron 1	NM_001040395.2	hypothetical protein LOC68646
	chr15 qD3	Tssk5	upstream	NM_183099.2	testis-specific serine kinase 5
	chr16 qA1	Gspt1	intron 9	NM_146066.1	G1 to S phase transition 1
	chr16 qA3	Hira	intron 14	NM_010435.2	histone cell cycle regulation defective homolog
	chr17 qA1		intergenic		
	chr17 qA1	Park2	intron 1	NM_016694.3	parkin
	chr17 qB3	Supt3h	intron 2	NM_178652.2	suppressor of Ty 3 homolog
	chr17 qB5	Crk	intron 2	NM_133656.2	v-crk sarcoma virus CT10 oncogene homolog
	chr18 qA1	Rock1	intron 11	NM_009071.2	Rho-associated coiled-coil forming kinase 1
	chr18 qD3		intergenic		
	chrX qA1.2	Utx	intron 2	NM_009483.1	ubiquitously transcribed tetratricopeptide
	chrX qA1.2	Utx	intron 9	NM_009483.1	ubiquitously transcribed tetratricopeptide
	chrX qF2	Lrch2	intron 8	NM_001081173.1	leucine-rich repeats and calponin homology (CH)
	chrX qF3		intergenic		
	chrX qF5	Gpm6b	intron 1	NM_023122.2	glycoprotein m6b
Vector: LV-GFP; Target: BM cells from *th3*/+ mice					
1	chr2 qB	Ptges	intron 2	NM_022415.2	prostaglandin E synthase
	chr5 qA1		intergenic		
	chr16 qC1.2	Cldnd1	3′ UTR	NM_171826.2	claudin domain containing 1
	chr18 qA1		intergenic		
2	chr1 qA2		intergenic		
	chr1 qH5	Gpatc2	intron 5	NM_026367.3	G patch domain containing 2
	chr8 qA4	Tusc3	intron 5	NM_030254.3	tumor suppressor candidate 3
	chr10 qB3	Man1a	intron 4	NM_008548.2	mannosidase alpha class 1A member 1
	chr13 qC3	Xrcc4	intron 6	NM_028012.1	X-ray repair complementing defective repair in Chinese hamster cells 4
	chr15 qB3.1	Pgcp	intron 3	NM_018755.2	plasma glutamate carboxypeptidase
	chr18 qB3	Arhgap26	intron 1	NM_175164.4	Rho GTPase activating protein 26
	chr19 qA	1200004M23Rik	downstream(27.7)	NM_026169.3	FKSG44 homolog
	chrX qA1.1		intergenic		
	chrX qB	EG636104	downstream(2.9)	NM_001099307.1	hypothetical protein LOC636104
3	chr3 qB	Sclt1	intron 9	NM_001081411.1	sodium channel associated protein 1
	chr9 qE2	Phip	intron 3	NM_001081216.1	pleckstrin homology domain interacting protein
	chr12 qD3	4930534B04Rik	intron 18	NM_181815.2	hypothetical protein LOC75216
	chr19 qC3	Tmem20	downstream (5.1)	NM_175507.3	transmembrane protein 20
4	chr2 qB	Arhgap15	intron 8	NM_153820.3	Rho GTPase activating protein 15 isoform 1
	chr2 qH1	Rbm39	intron 4	NM_133242.2	RNA binding motif protein 39
5	chr13 qA3.1		intergenic		
	chr13 qA3.3	Gmds	intron 1	NM_146041.2	GDP-mannose 4,6-dehydratase
	chr13 qD1	Rnf180	upstream	NM_027934.2	ring finger protein 180
6	chr2 qC1.1	March7	intron 1	NM_020575.2	membrane-associated ring finger (C3HC4) 7
	chr5 qA1	Cacna2d1	intron 3	NM_009784.1	calcium channel voltage-dependent alpha2/delta
	chr5 qE2	Rchy1	intron 4	NM_026557.3	androgen receptor N-terminal-interacting
	chr9 qA4		intergenic		
	chr14 qE4		intergenic		

Columns indicate, from left to right: mouse number, chromosomal location of each integrated provirus, target gene symbol, RefSeq identifier no., and biological function of target genes. Integrations were distributed as inside, outside, or at a distance of <30 kb upstream or downstream from known genes (University of California Santa Cruz annotation).

Table S4. List of genomic control sequences randomly cloned from a library of murine BM cells DNA

Chromosomal band	Target gene	Location	RefSeq no.	Biological function
chr1 qA5	B3gat2	intron 1	NM_172124.2	galactosylgalactosylxylosylprotein
chr1 qC1.1	Gulp1	intron 1	NM_027506.2	PTB domain adaptor protein CED-6 isoform 1
chr1 qC3	Spag16	intron 1	NM_029160.2	sperm associated antigen 16 isoform 1
chr1 qD		intergenic		
chr1 qE1.2		intergenic		
chr1 qE2.1	Rnf152	intron 1	NM_178779.2	ring finger protein 152
chr1 qH4	Akt3	intron 1	NM_011785.2	thymoma viral proto-oncogene 3
chr2 qC1.1	Gpd2	downstream	NM_010274.2	glycerol phosphate dehydrogenase 2
chr2 qC1.1		intergenic		
chr2 qE1		intergenic		
chr2 qE1		intergenic		
chr2 qE3		intergenic		
chr2 qE3		intergenic		
chr2 qF3		intergenic		
chr2 qH4	Ctsz	intron 3	NM_022325.3	cathepsin Z preproprotein
chr2 qH4	Cdh4	intron 1	NM_009867.1	cadherin 4
chr3 qB		intergenic		
chr3 qC	Foxo1	intron 1	NM_019739.2	forkhead box O1a
chr3 qE3		intergenic		
chr3 qF1	D930015E06Rik	upstream	NM_172681.3	hypothetical protein LOC229473
chr4 qA3		intergenic		
chr4 qA3		intergenic		
chr4 qA5		intergenic		
chr4 qB1		intergenic		
chr4 qB2		intergenic		
chr4 qC3		intergenic		
chr4 qC6		intergenic		
chr4 qD2.1	BC057371	intron 1	NM_177572.3	hypothetical protein LOC194237
chr5 qA3	Gbx1	intron 1	NM_015739.2	gastrulation brain homeobox 1
chr5 qE1		intergenic		
chr5 qG3	Kl	intron 3	NM_013823.1	klotho
chr6 qD1		intergenic		
chr7 qE1	Gab2	intron 1	NM_010248.1	growth factor receptor bound protein
chr7 qF1	Acsm3	intron 11	NM_212441.2	SA hypertension-associated homolog
chr8 qB3.1		intergenic		
chr9 qA1		intergenic		
chr9 qB	Thsd4	intron 7	NM_001040426.1	thrombospondin type I domain containing 4
chr9 qD	Hcrtr2	upstream	NM_198962.1	hypocretin (orexin) receptor 2
chr9 qF1	Ky	intron 4	NM_024291.3	kyphoscoliosis
chr10 qA1		intergenic		
chr10 qA4	Lama2	intron 1	NM_008481.2	laminin alpha 2
chr10 qB2		intergenic		
chr10 qB3	Mcmdc1	intron 10	NM_027830.2	minichromosome maintenance deficient domain
chr10 qC2		intergenic		
chr11 qA5		intergenic		
chr11 qE1	Wdr68	intron 1	NM_027946.3	WD-repeat protein
chr12 qA1		intergenic		
chr12 qC1	Npas3	intron 1	NM_013780.1	neuronal PAS domain protein 3
chr12 qE1	1700029P11Rik	downstream	NM_025503.1	hypothetical protein LOC66346
chr13 qD2.1		intergenic		
chr14 qA2		intergenic		
chr14 qD3	Pcdh17	intron 1	NM_001013753.1	protocadherin 17
chr14 qE2.1	4921530L21Rik	downstream	NM_025733.1	hypothetical protein LOC66732
chr16 qC3.1	Nrip1	downstream	NM_173440.1	nuclear receptor interacting protein 1
chr16 qC3.1	Robo1	intron 2	NM_019413.2	roundabout homolog 1
chr17 qA1	Slc22a3	intron 1	NM_011395.1	solute carrier family 22 (organic cation
chr17 qB1	Olfr128	upstream	NM_206816.1	olfactory receptor 128
chr17 qB3	Dscr1l1	intron 2	NM_207649.1	Down syndrome critical region gene 1-like 1
chr18 qE3	Zbtb7c	intron 1	NM_145356.3	zinc finger and BTB domain containing 36
chrX qA2		intergenic		
chrX qE3		intergenic		

Columns indicate, from left to right: chromosomal location of each random genomic control sequence, target gene symbol, RefSeq identifier, and biological function of target genes. Control sequences were obtained from a randomly cloned library of MseI/NarI-restricted, LM-PCR-amplified murine BM DNA. Random genomic sequence were distributed as inside, outside, or at a distance of <30 kb upstream or downstream from known genes (University of California Santa Cruz annotation).

Table S5. Hematological parameters and erythroid expansion in *th3/th3* long-term hematopoietic chimeras

Mouse no.	Donor cells	Hb, g/dl	Hct, %	RBCs $\times 10^6/\mu$l	% BM Ter119$^+$ cells*	% Splenic Ter119$^+$ cells*	VCN
	th3/th3†	3.52 ± 0.50	12.64 ± 2.22	2.02 ± 0.22	80.10 ± 2.75	90.40 ± 1.75	NA
	th3/+‡	8.12 ± 0.20	33.33 ± 0.65	6.90 ± 0.25	51.3 ± 10.55	39.50 ± 10.25	NA
1	*th3/th3* LV§	8.6	35.01	5.9	66	59	3.2
23	*th3/th3* LV	8.1	33.54	5.3	65	65	4.2
21	*th3/th3* LV	11.9	45.36	7.2	48	74	5.3
22	*th3/th3* LV	12.2	44.04	7.2	67	56	3.4
	WT¶	13.42 ± 0.5	47.80 ± 3.22	9.22 ± 0.40	28.5 ± 5.75	3.25 ± 2.50	NA

NA, not applicable.

*percentage of Ter119$^+$ cells, determined by FACS staining after *in silico* gating on mononucleated cells.

†Recipients transplanted with mock-transduced *th3/th3* FLCs ($n = 9$). Analysis for this group was performed between 30 and 80 days after-transplant when mice were killed because they were moribund and profoundly anemic.

‡Recipients transplanted with mock-transduced *th3/+* FLCs ($n = 8$).

§Recipients transplanted with GLOBE-transduced *th3/th3* FLCs.

¶Recipients transplanted with mock-transduced normal (WT) FLCs ($n = 8$).

Table S6. Correction of organ pathology in mice transplanted with GLOBE-transduced FLCs

Mouse no.	Donor cells	Splenic iron deposits*	Liver EMH†	Liver iron deposits*
	th3/th3‡	2/4	2/3	2/4
	th3/+	2/3	0/3	0/3
1	*th3/th3* LV	2	1	2
23	*th3/th3* LV	2	3	2
21	*th3/th3* LV	1	2	0/1
22	*th3/th3* LV	2/3	1	0/1
	WT	0/2	0/1	0/1

*The amount of iron deposition was scored semiquantitatively on a scale from 0 (no iron present) to 4 (maximum amount of iron identified in the organ).

†The extent of EMH was scored semiquantitatively on a scale from 0 (no EMH present) to 3 (maximum extent of EMH identified in the organ).

‡Analyses for this group ($n = 9$) were performed between 30 and 80 days after transplant when mice were killed because they were moribund and profoundly anemic.