SUPPLEMENTARY DATA

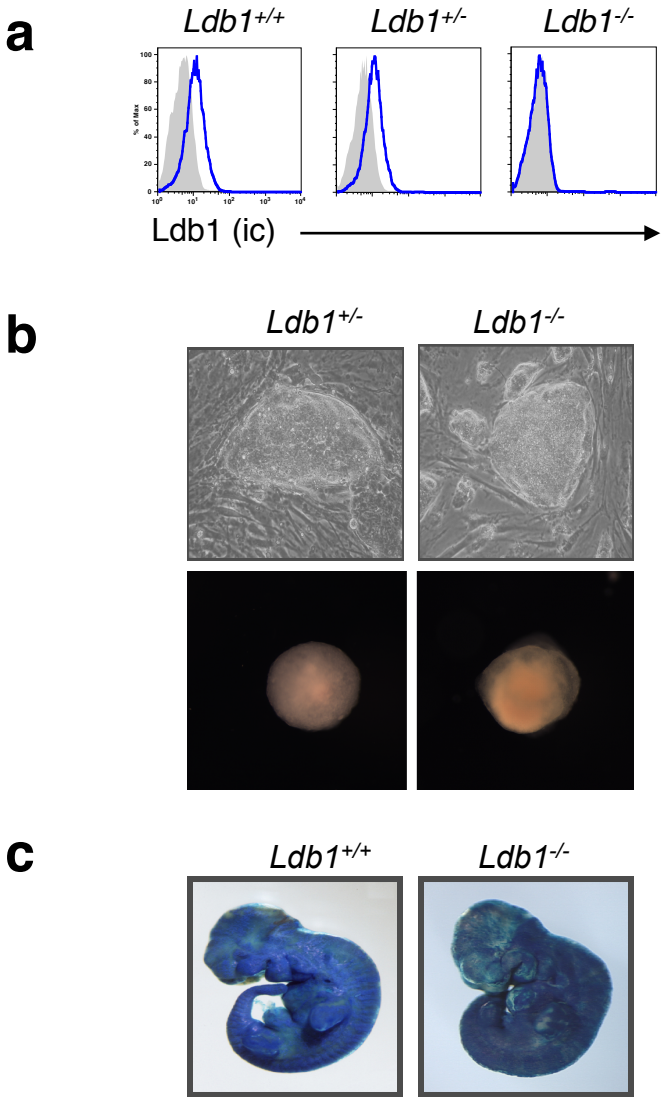
Nuclear adaptor Ldb1 regulates a transcriptional program essential for the maintenance of hematopoietic stem cells

LiQi Li[1], Raja Jothi[2], Kairong Cui[3], Jan Y. Lee[1], Tsadok Cohen[4], Marat Gorivodsky[4], Itai Tzchori[4], Yangu Zhao[4], Sandra M. Hayes[5], Emery H. Bresnick[6], Keji Zhao[3], Heiner Westphal[4], and Paul E. Love[1]*

[1]Section on Cellular & Developmental Biology, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892. [2]Biostatistics Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, North Carolina 27709. [3]Laboratory of Molecular Immunology, National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, Maryland 20892. [4]Section on Mammalian Molecular Genetics, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892. [5]SUNY Upstate Medical University, Department of Microbiology & Immunology, Syracuse, NY 13210. [6]University of Wisconsin School of Medicine and Public Health, Wisconsin Institutes for Medical Research, Paul Carbone Cancer Center, 1111 Highland Ave., Madison, WI 53705.

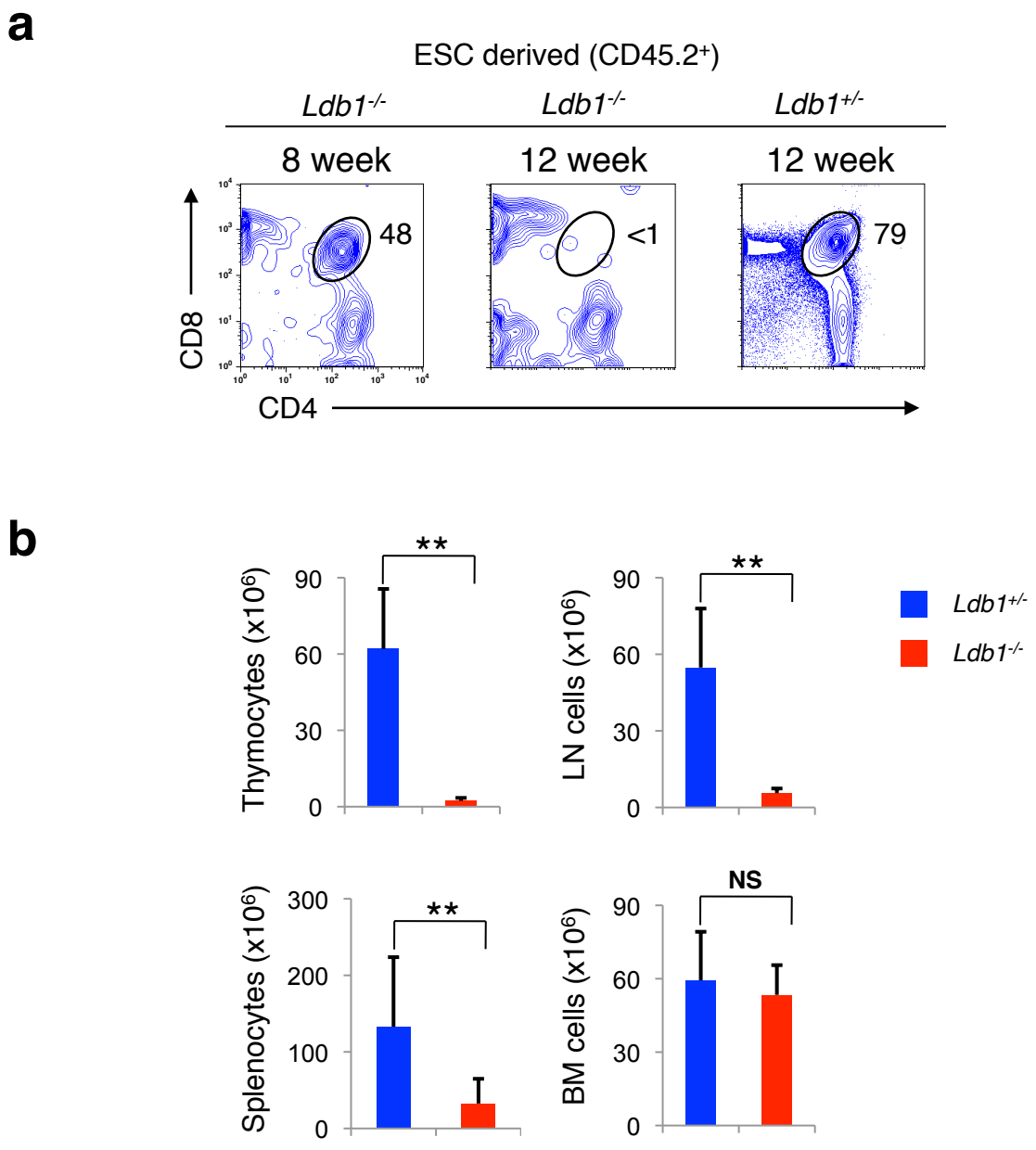
*Correspondence: email: lovep@mail.nih.gov
TEL: 301-402-4946
FAX: 301-480-6332

Supplementary Figure 1



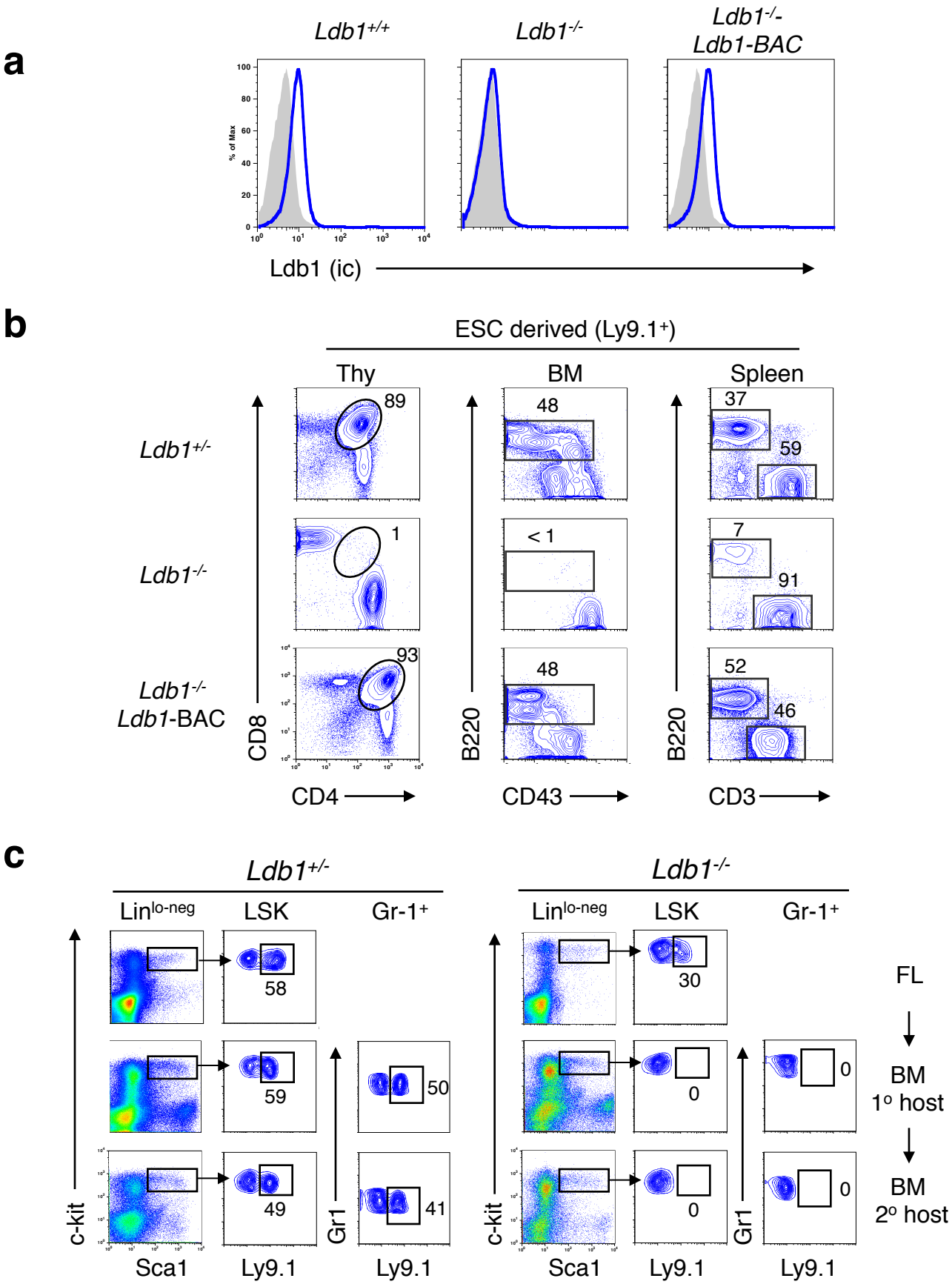
Supplementary Figure 1. *Ldb1* is not required for ESC maintenance.
(**a**) Intracellular staining of *Ldb1*$^{+/+}$, *Ldb1*$^{+/-}$, and *Ldb1*$^{-/-}$ ESC lines with anti-Ldb1. Shaded area represents control rabbit IgG staining; solid line represents Ldb1 staining. One representative of 10 experiments. (**b**) *Ldb1*$^{+/-}$ and *Ldb1*$^{-/-}$ ESC colonies (top panels) and embryoid bodies (bottom panels). (**c**) Chimeric mice were generated by injecting *Ldb1*$^{+/+}$, *RosaLacZ* ESC or *Ldb1*$^{-/-}$, *RosaLacZ* ESCs into C57/BL6 blastocysts. E10.5 embryos were fixed and stained with X-Gal to detect β-galactosidase activity. Images are representative of 6 *Ldb1*$^{+/+}$ and 9 *Ldb1*$^{-/-}$ ESC injected chimeric embryos obtained from 5 independent blastocyst injection experiments.

Supplementary Figure 2



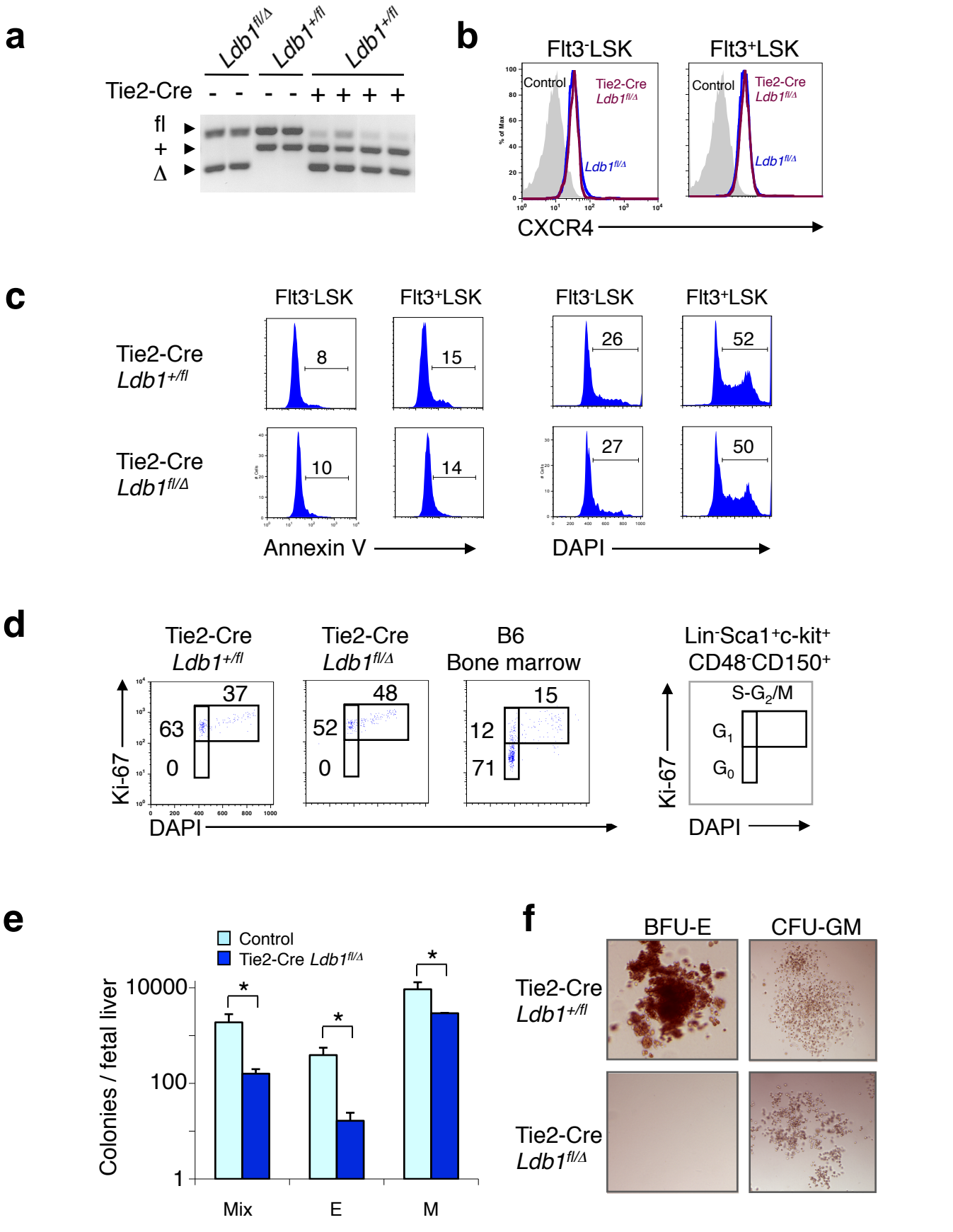
Supplementary Figure 2. *Ldb1*$^{-/-}$ ESC-derived CD4^{+}CD8^{+} thymocytes are present in the thymus of 8 week but not 12 week old adult chimeric mice. (**a**) Chimeric mice were generated by injection of *Ldb1*$^{+/-}$ or *Ldb1*$^{-/-}$ (CD45.2) ESCs into *Rag2*$^{-/-}$ (CD45.1) blastocysts. ESC derived thymocytes were identified by staining for CD45.2. Data are representative of 3 experiments. (**b**) Total number of CD45.2^{+} cells from thymus, lymph nodes (LN), spleen, and bone marrow (BM) of adult *Ldb1*$^{+/-}$ or *Ldb1*$^{-/-}$ ESC injected blastocyst chimeric mice. *Ldb1*$^{+/-}$ ESC chimeric mice, n=6; *Ldb1*$^{-/-}$ ESC chimeric mice, n=3. ** $p < 0.01$. NS, not significant.

Supplementary Figure 3



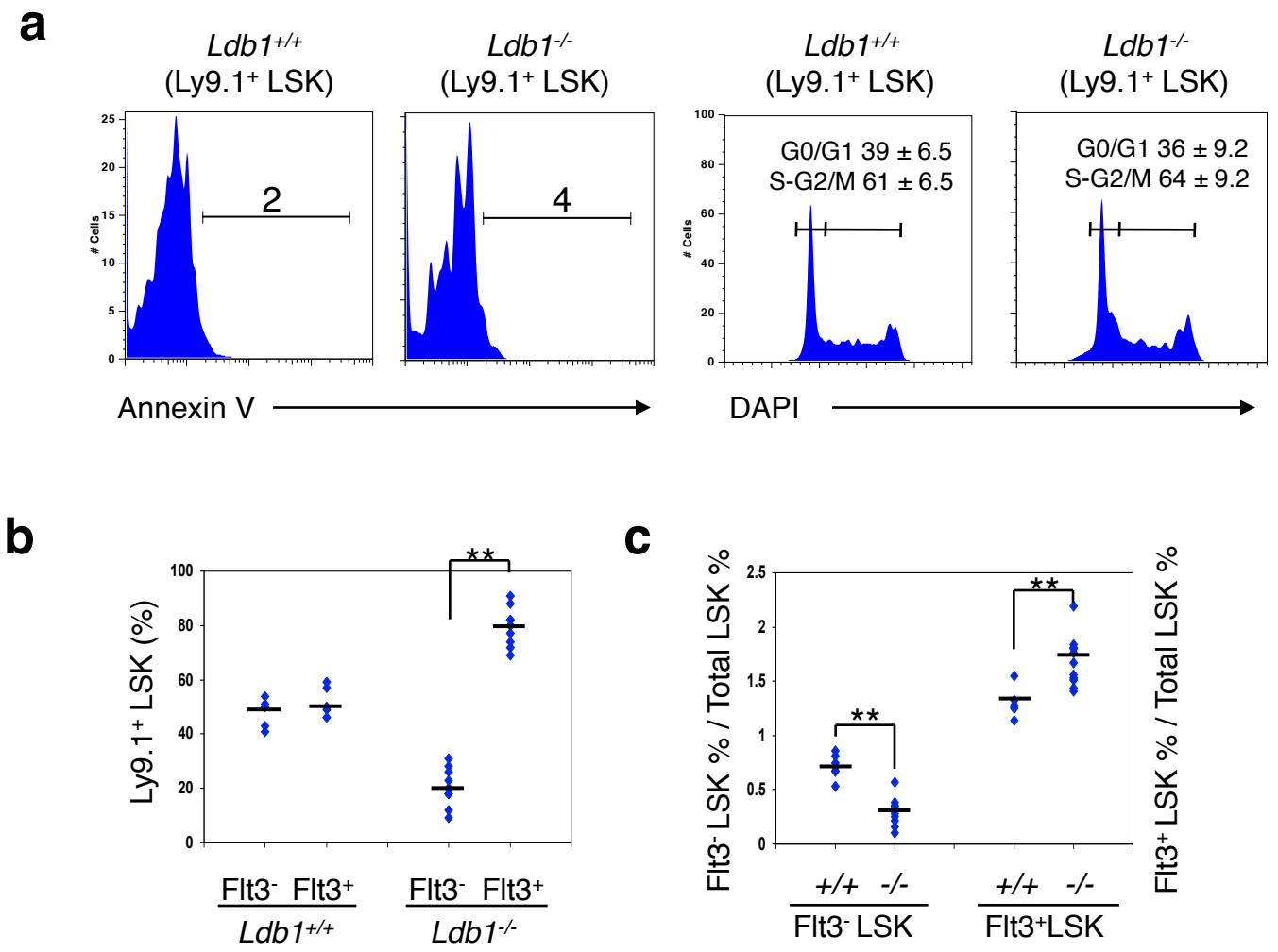
Supplementary Figure 3. **Reconstitution of Ldb1 expression in *Ldb1*$^{-/-}$ ESC restores long term repopulating HSC potential.** (**a**) Intracellular Ldb1 staining of *Ldb1*$^{+/+}$, *Ldb1*$^{-/-}$ and *Ldb1*-BAC reconstituted *Ldb1*$^{-/-}$ (*Ldb1*$^{-/-}$; *Ldb1*-BAC) ESC lines. Shaded area shows control rabbit IgG staining; solid line represents Ldb1 staining. One representative of 5 experiments from 2 independently generated clones. (**b**) *Ldb1*$^{+/-}$ (top row), *Ldb1*$^{-/-}$ (middle row) or *Ldb1*$^{-/-}$; *Ldb1*-BAC ESC (bottom row) were injected into *Rag2*$^{-/-}$ blastocysts to generate chimeric embryos. E15.5 fetal liver (FL) cells were harvested, and injected into irradiated *Rag2*$^{-/-}$ (Ly9.1^{-}) mice. Sixteen weeks later, thymus (Thy), bone marrow (BM), and spleen cells from recipient mice were analyzed for the presence of ESC derived (Ly9.1^{+}) hematopoietic lineage cells. Data are representative of four independent experiments (**c**) Confirmation of loss of HSCs after deletion of *Ldb1* in secondary BM transfers. Primary transfer experiments were performed with fetal liver cells from E15.5 blastocyst chimeras as described in (**b**). Sixteen weeks after transfer BM cells from recipient mice were analyzed by FACS and the remaining cells were transferred into irradiated secondary *Rag2*$^{-/-}$ (Ly9.1^{-}) hosts.

Supplementary Figure 4



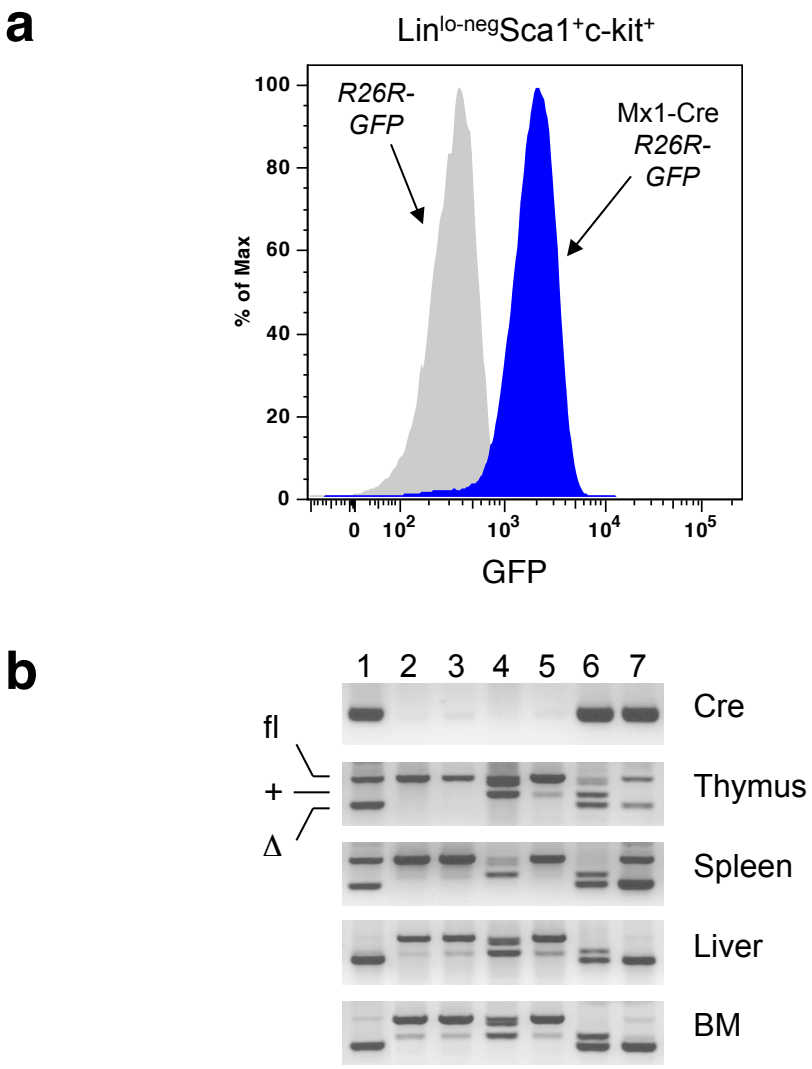
Supplementary Figure 4. Analysis of fetal liver cells in Tie2-Cre *Ldb1*$^{fl/\Delta}$ embryos. (**a**) PCR analysis of genomic DNA from total fetal liver cells from E12.5 *Ldb*$^{+/fl}$ ± Tie2-Cre and *Ldb*$^{fl/\Delta}$ embryos to assess *Ldb1* deletion efficiency. One representative of 2 experiments. Arrowheads indicate PCR products generated by the Floxed (fl), wild type (+) and deleted (Δ) alleles. Presence (+) or absence (-) of the Tie2-Cre transgene is indicated. (**b**) CXCR4 surface expression on Flt3$^-$ LSK and Flt3$^+$ LSK from E12.5 *Ldb*$^{fl/\Delta}$ (blue) and Tie2-Cre *Ldb*$^{fl/\Delta}$ (red) littermate mice. Shaded area represents staining with control antibody. Solid line represents staining with anti-CXCR4. One representative of 3 experiments. (**c**) Annexin V (columns one and two), and DAPI (columns three and four) staining profiles of fetal liver Flt3$^-$ LSKs and Flt3$^+$ LSKs from E12.5 littermate Tie2-Cre *Ldb*$^{+/fl}$ and Tie2-Cre *Ldb*$^{fl/\Delta}$ embryos. One representative of 3 experiments. (**d**) Cell cycle analysis of Lin$^-$Sca1$^+$c-kit$^+$CD48$^-$CD150$^+$ fetal liver HSC by staining with Ki-67 and DAPI. Gates show percentage of cells in G_0, G_1 and S-G_2/M phases. Identical staining of adult bone marrow cells is included as a positive control. One representative of two experiments. (**e**) Numbers of mixed (Mix), erythroid (E) and myeloid (M) colonies in day 8 methycellulose cultures from E12.5 control (Tie2-Cre *Ldb1*$^{+/fl}$) and Tie2-Cre *Ldb*$^{fl/\Delta}$ fetal livers are shown. One representative of 3 experiments. Colony numbers per fetal liver: Mixed colonies: Control: 1892 ± 915; Tie2-Cre *Ldb*$^{fl/\Delta}$: 159 ± 39; Erythroid (BFU-E): Control: 390 ± 166; Tie2-Cre *Ldb1*$^{fl/\Delta}$: 17 ± 8; Granulocyte/Macrophage: Control: 9308 ± 3759; Tie2-Cre *Ldb1*$^{fl/\Delta}$: 2931 ± 54. (**f**) Representative erythroid (BFU-E), and myeloid (CFU-GM) colonies from Tie2-Cre *Ldb*$^{+/fl}$ and Tie2-Cre *Ldb*$^{fl/\Delta}$ fetal livers at day 8 of culture. Note absence of BFU-E colonies in cultures from Tie2-Cre *Ldb*$^{fl/\Delta}$ mice. Images represent duplicate cultures from 3 litters. Original magnification 50x.

Supplementary Figure 5



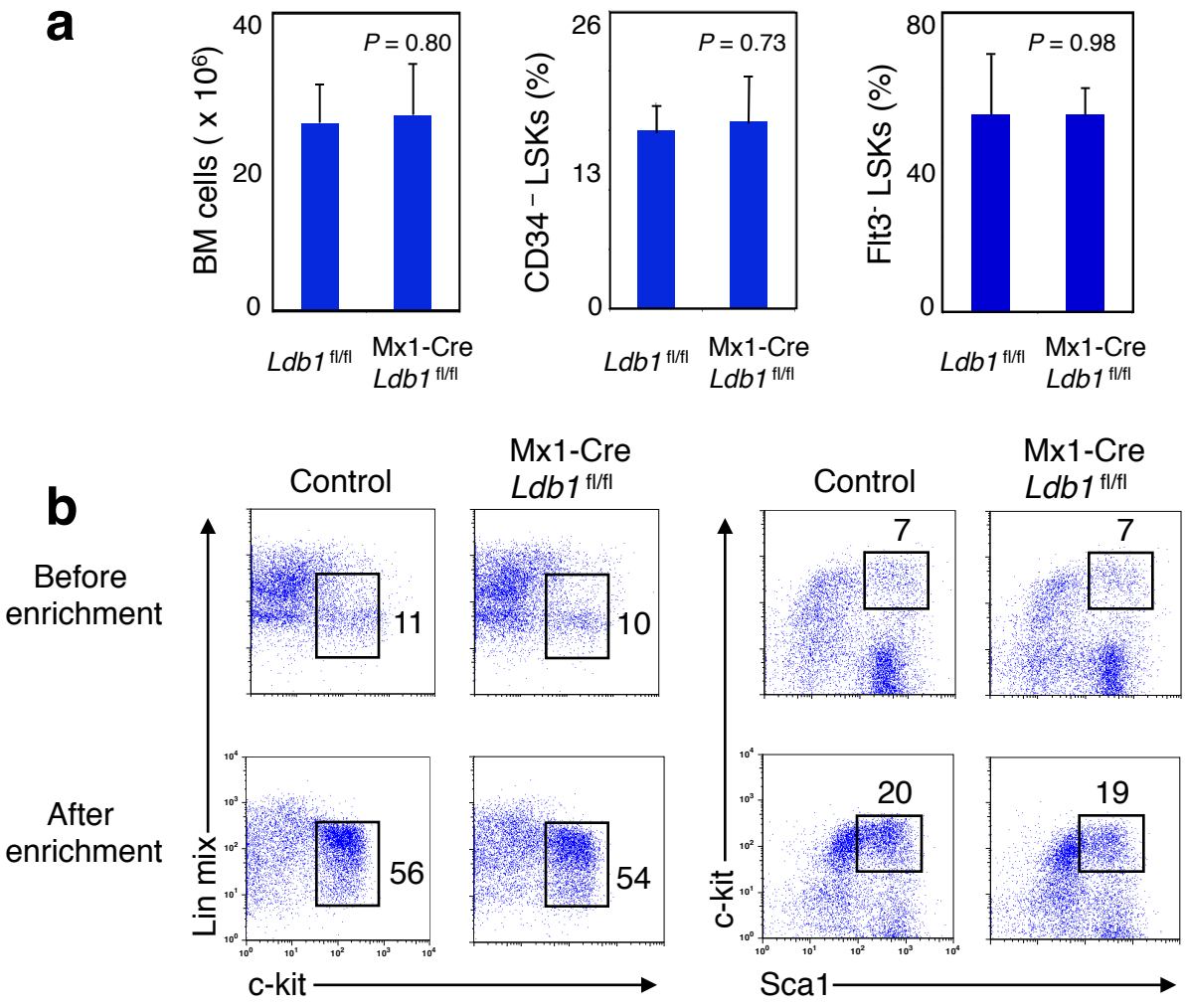
Supplementary Figure 5. Analysis of *Ldb*$^{-/-}$ ESC derived fetal liver LSK.
(**a**) Annexin V and DAPI staining profiles of E15.5 fetal liver LSKs derived from *Ldb*$^{+/+}$ ESCs (*Ly9.1*$^{+}$, *Ldb*$^{+/+}$) or *Ldb*$^{-/-}$ ESCs (*Ly9.1*$^{+}$, *Ldb*$^{-/-}$) from blastocyst chimeras. One representative of 3 experiments. (**b**) Summary of data from *Ldb1*$^{-/-}$ /B6 and control *Ldb1*$^{+/+}$/B6 chimeras. Percentages of ESC derived (Ly9.1^{+}) Flt3^{-} LSKs and Flt3^{+}LSKs are shown. Each symbol represents data from an individual chimeric embryo. Data shown are from E13.5-E16.5 embryos. (**c**) Relative to control *Ldb1*$^{+/+}$ ESC derived LSKs, *Ldb1*$^{-/-}$ ESC derived LSKs are composed of more Flt3^{+} MPP and fewer Flt3^{-} HSC. Data were calculated for each chimera by determining the percentage of Flt3^{+}LSKs (or Flt3^{-}LSKs) that are Ly9.1^{+} and dividing by the percentage of total LSKs that are Ly9.1^{+}. $^{**}p < 0.01$. $p = 2.4 \times 10^{-6}$ for Flt3^{-}LSK; $p = 2.0 \times 10^{-4}$ for Flt3^{+}LSK. Each symbol represents data from an individual chimeric embryo.

Supplementary Figure 6



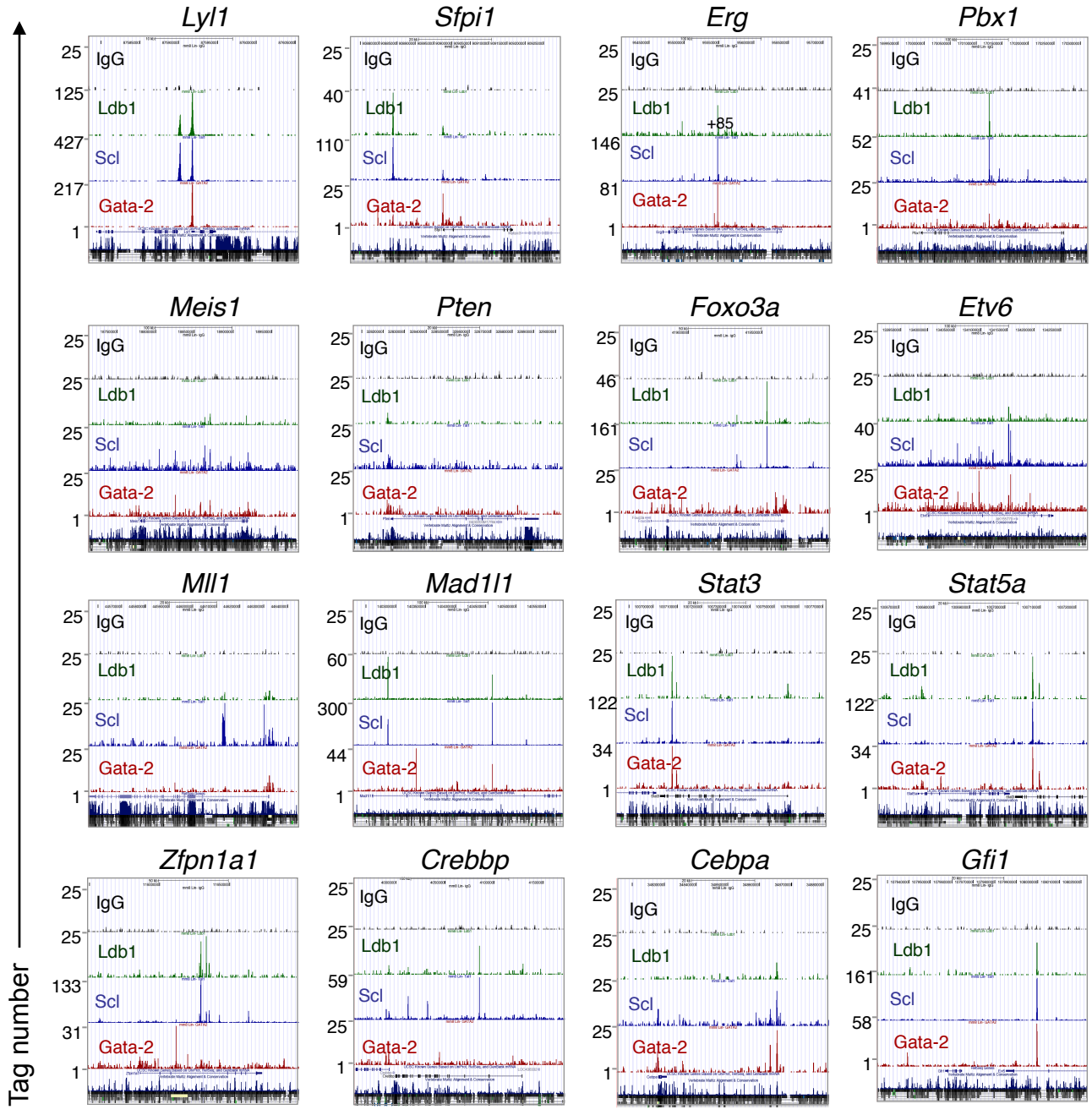
Supplementary Figure 6. Mx1-Cre transgene mediated activation and *Ldb1* deletion efficiency in bone marrow LSKs. (**a**) Expression of GFP in bone marrow LSKs from Mx1-Cre R26R-GFP reporter mice and control R26R-GFP mice after pI:pC treatment. Mice were injected with pI:pC every other day X 3 and bone marrow LSKs were analyzed by FACS 7 days after the last injection. Data shown are representative of those obtained with 3 control R26R-GFP mice and 4 experimental Mx1-Cre R26R-GFP mice analyzed. (**b**) Efficiency of *Ldb1* deletion in pI:pC treated mice. Mice were injected with pI:pC as described in (**a**). Lanes 1, 7: Mx1-Cre *Ldb1*$^{fl/fl}$, lanes 2, 3, 5: *Ldb1*$^{fl/fl}$, lane 4: *Ldb1*$^{+/fl}$, lane 6: Mx1-Cre *Ldb1*$^{+/fl}$. Data shown are from 2 experimental (lanes 1,7) and 5 control (lanes 2-6) mice.

Supplementary Figure 7



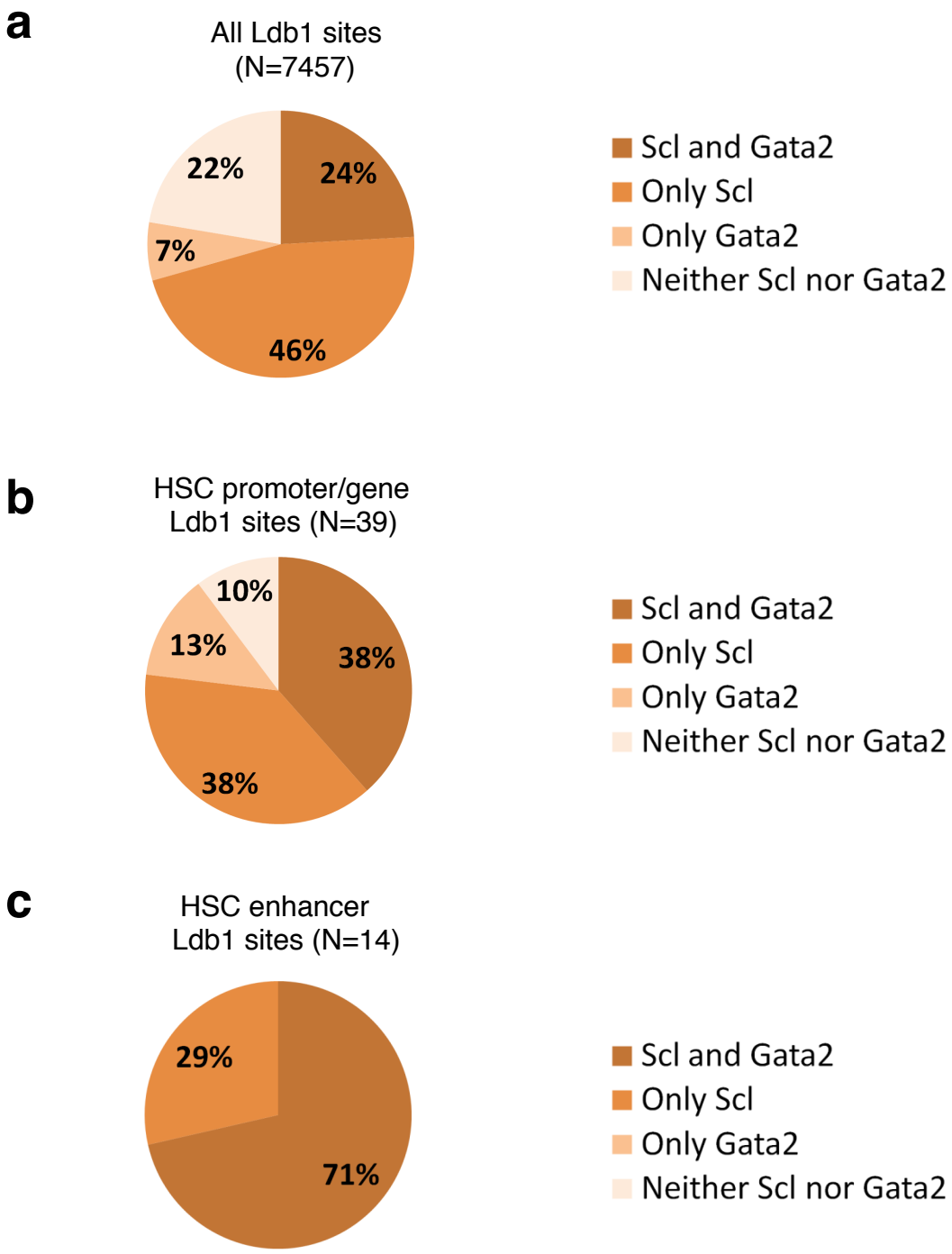
Supplementary Figure 7. Enrichment of BM hematopoietic progenitor cells from pI:pC injected mice for real-time PCR and ChIP-seq analysis. Control (Mx1-Cre *Ldb1*$^{+/+}$) and experimental Mx1-Cre *Ldb1*$^{fl/fl}$ mice were injected with pI:pC on two consecutive days and sacrificed on day 3. (**a**) Assessment of total number of bone marrow cells, percentage of CD34$^-$LSKs and percentage of Flt3$^+$LSKs before hematopoietic progenitor cell enrichment by lineage$^+$ cell depletion. Results shown were pooled cells obtained from 5 Control (Mx1-Cre *Ldb1*$^{+/+}$) or 6 experimental (Mx1-Cre *Ldb1*$^{fl/fl}$) mice. Total bone marrow cell counts/mouse: Control, 25±5 x 10^6, Mx1-Cre *Ldb1*$^{f/f}$, 26±7 x 10^6; p=0.795. Percent CD34$^-$ LSKs: Control, (15±2)%, Mx1-Cre *Ldb1*$^{f/f}$, (16±4)%; p=0.730. Percent Flt3$^+$ LSKs: Control, (47±16)%, Mx1-Cre *Ldb1*$^{f/f}$, (47±7)%; p=0.977. (**b**) FACS analysis of BM cells before and after enrichment of hematopoietic progenitor cells. The yield was 1-2 x10^6 cells for each group after subtraction of cells for FACS validation.

Supplementary Figure 8



Supplementary Figure 8. Genome browser images of Ldb1 complex binding sites at or near HSC maintenance genes or within known enhancer elements. Shown are binding sites identified by anti-Ldb1, anti-Scl or anti-Gata2 ChIP-seq on lineage depleted bone marrow cells. Binding sites at known distal regulatory elements are also designated where applicable. The y-axis denotes normalized ChIP-Seq tag densities. Also shown are binding sites at putative regulatory sites near *Cebpa* and *Gfi1*. The major Ldb1 complex binding site fragments contained conserved E-box and GATA motifs and conserved binding sites for Ets, Myb, Gfi1, Runx1 and Meis1. Locations of the prominent Ldb1 complex binding sites are: *Cebpa*, chr7:34,865,270-34,865,649; *Gfi1*, chr5:107,999,505-107,999,827.

Supplementary Figure 9



Supplementary Figure 9. Co-localization of Ldb1, Scl and Gata2 binding sites. Percentage of fragments identified as Ldb1 complex binding sites in lineage depleted bone marrow cells by anti-Ldb1 ChIP-seq that bind Scl, Gata2, both or neither. (**a**) Analysis of the 7457 total genomic Ldb1 complex sites. (**b**) Analysis of the 39 Ldb1 complex binding sites within the promoter/body of HSC maintenance genes. (**c**) Analysis of the 14 Ldb1 complex binding sites within known enhancer elements of HSC maintenance genes.

Supplementary Table 1. Ldb1 complex binding sites within HSC genes (promoter +/- 5Kb and gene body).

Number	Gene	Gene coordinates	Strand	Ldb1 binding site location	Fold (ChIP) enrichment	*p-value*	Notes*
1	*Ikaros (Zfpn1a1)*	chr11:11586215-11672929	+	chr11: 11628341-11628541	8.02	4.7E-04	within GB
				chr11: 11629321-11629521	27.62	6.3E-05	within GB
				chr11: 11633321-11633521	24.95	8.1E-05	within GB
				chr11: 11635291-11635491	13.37	2.3E-04	within GB
2	*Erg*	chr16:95469001-95639163	-	chr16: 95497211-95497411	7.13	5.8E-04	within GB
				chr16: 95548121-95548321	18.71	1.4E-04	within GB
3	*Lmo2*	chr2:103758833-103782710	+	chr2:103757771-103757971	41.88	2.7E-05	at Promoter
			+	chr2:103770321-103770521	8.02	4.7E-04	within GB
4	*Scl (Tal1)*	chr4:114557458-114569687	+	chr4: 114554511-114554711	104.26	6.0E-06	at Promoter
				chr4: 114554831-114555031	9.80	4.0E-04	at Promoter
5	*Lyl1*	chr8:87591561-87594821	+	chr8: 87590201-87590401	81.09	1.1E-05	at Promoter
				chr8: 87591671-87591871	163.97	1.0E-06	at Promoter
6	*Gata2*	chr6:88164304-88172671	+	chr6: 88168741-88168941	24.95	8.1E-05	within GB
7	*Pu.1 (Sfpi1)*	chr2:90897635-90916595	+	chr2: 90897471-90897671	14.26	2.1E-04	at Promoter
8	*Cbp (Crebbp)*	chr16:3999275-4128632	-	chr16: 4084881-4085081	16.93	1.8E-04	within GB
9	*Etv6*	chr6:134001399-134235840	+	chr6: 134152101-134152301	10.69	3.3E-04	within GB
10	*Myb*	chr10:20814345-20850400	-	chr10: 20835891-20836091	9.80	4.0E-04	within GB
11	*Meis1*	chr11:18780430-18918683	-	chr11: 18870091-18870291	8.02	4.7E-04	within GB
12	*Mll1*	chr9:44554349-44632269	-	chr9: 44632391-44632591	7.13	5.8E-04	at Promoter
13	*Foxo3a*	chr10:41874200-41965157	-	chr10: 41932891-41933091	11.58	3.0E-04	within GB
				chr10: 41953511-41953711	52.58	1.7E-05	within GB
				chr10: 41966481-41966681	8.91	4.6E-04	at Promoter
14	*Stat5a*	chr11:100675440-100701258	+	chr11: 100676491-100676691	9.80	4.0E-04	at Promoter
15	*E2a (Tcfe2a)*	chr10:79812293-79836782	-	chr10: 79824971-79825171	17.82	1.6E-04	within GB
				chr10: 79836771-79836971	9.80	4.0E-04	at Promoter
16	*Runx1*	chr16:92490324-92714606	-	chr16: 92604491-92604691	51.68	1.9E-05	within GB
				chr16: 92653711-92653911	7.13	5.8E-04	within GB
				chr16: 92690521-92690721	66.83	1.4E-05	within GB
				chr16: 92711001-92711201	25.84	7.5E-05	within GB
17	*Mad1l1*	chr5:140261162-140574026	-	chr5: 140286911-140287111	86.44	1.1E-05	within GB
				chr5: 140461271-140461471	24.95	8.1E-05	within GB
				chr5: 140574061-140574261	7.13	5.8E-04	at Promoter
18	*Pten*	chr19:32823573-32892157	+	chr19: 32822051-32822251	7.13	5.8E-04	at Promoter
19	*Stat3*	chr11:100702899-100755601	-	chr11: 100707851-100708051	24.06	8.1E-05	within GB
				chr11: 100709581-100709781	8.91	4.6E-04	within GB
				chr11: 100755531-100755731	10.69	3.3E-04	at Promoter
20	*Pbx1*	chr1:169956038-170268933	-	chr1: 170124971-170125171	57.03	1.5E-05	within GB
				chr1: 170125231-170125431	9.80	4.0E-04	within GB
				chr1: 170125941-170126141	16.04	1.8E-04	within GB

GB, Gene body.

Supplementary Table 2. Conserved transcription factor binding motifs within Ldb1 complex binding sites.

Gene	Binding site location	GATA	CANNTG	Paired motif(s)[a]	Loc.[b]	Transcription factor binding motifs[c]	Conservation[d]
Ikaros	chr11: 11628341-8541	Yes	Yes	No	GB	My, Nk	M,R
(Zfpn1a1)	chr11: 11629321-9521	Yes	No	Yes	GB	E, A1, Sp	M,R,H
	chr11: 11633321-3521	Yes	Yes	Yes	GB	E, A1, NF, G, Me, My, Sp, C	M,R,H
	chr11: 11635291-5491	Yes	Yes	Yes	GB	E, Sp, A2	M,R,H
Erg	chr16: 95497211-7411	No	Yes	Yes	GB	C, G, My, Nk	M,R,H,C
	chr16: 95548121-8321	Yes	Yes	Yes	GB	E, C, Sp, My, G, F, Sm	M,R,H,C,Md
Lmo2	chr02:103757771-7971	Yes	Yes	Yes	P	E, Me, Sm	M,R,H
	chr02:103770321-0521	No	No	No	GB	E, Sp, A2	M,R,H,C
Scl (Tal1)	chr04: 114554511-4711	Yes	Yes	Yes	P	A2, Sp, C	M,R,H,C
	chr04: 114554831-5031	No	No	No	P	E, P, C, Sp	M,R,H,C
Lyl1	chr08: 87590201-0401	No	Yes	No	P	E, A2, C	M,R,H
	chr08: 87591671-1871	Yes	No	Yes	P	E, A2, H, Sp	M,R,H,C
Gata2	chr06: 88168741-8941	Yes	Yes	Yes	GB	E, C, Sp, O, Mf, G	M,R,H,C,Md
Pu.1 (Sfpi1)	chr02: 90897471-7671	Yes	Yes	Yes	P	E, Sp, A2, O, G, Nk, C	M,R,H,C,Md
Crebbp	chr16: 4084881-5081	No	Yes	Yes	GB	E, Sp	M,R,H
Etv6	chr06: 134152101-2301	Yes	Yes	Yes	GB	E, C, My, F, R, Sm	M,R,H,C,Md,G
Myb	chr10: 20835891-6091	Yes	Yes	Yes	GB	E, Sp, My	M,R,H,C,Md
Meis1	chr11: 18870091-0291	Yes	No	No	GB	E, Sp, C, F	M,R,H,C
Mll1	chr09: 44632391-2591	Yes	Yes	No	P	E, Sp A2, C	M,H,Md
Foxo3a	chr10: 41932891-3091	Yes	Yes	No	GB	E, Sp, G, Me	M,H,C
	chr10: 41953511-3711	Yes	Yes	Yes	GB	Sp, My, G, C	M,R,H,C
	chr10: 41966481-6681	Yes	Yes	Yes	P	E, Sp, F	M,R,H
Stat5a	chr11: 100676491-6691	Yes	Yes	No	P	E, Sp, A2, My	M,R,H,C
E2a	chr10: 79824971-5171	Yes	No	Yes	GB	Sp	M,R,H,C
(Tcfe2a)	chr10: 79836771-6971	Yes	Yes	No	P	E, C, Sp	M,H,C
Runx1	chr16: 92604491-4691	Yes	Yes	Yes	GB	E, Sp, My, C	M,R,H,C
	chr16: 92653711-3911	Yes	No	Yes	GB	E, R, G, Sp, C, F	M,R,H
	chr16: 92690521-0721	Yes	Yes	Yes	GB	E, Sp, R, My, G, Me	M,R,H,C
	chr16: 92711001-1201	Yes	Yes	Yes	GB	E, R, F	M,R,H,C,Md,G,X
Mad1l1	chr05: 140286911-7111	Yes	Yes	Yes	GB	E, Sp	M,R,H
	chr05: 140461271-1471	No	Yes	Yes	GB	E, Sp	M,R,H
	chr05: 140574061-4261	No	No	No	P	Sp	M,R,H,C
Pten	chr19: 32822051-2251	No	Yes	Yes	P	E, Sp, A2, Me, Nk, F, G	M,R,H
Stat3	chr11: 100707851-8051	Yes	Yes	Yes	GB	E, G, Sm, F	M,R,H,C
	chr11: 100709581-9781	Yes	Yes	No	GB	E, Sp, C, My	M,R,H,C
	chr11: 100755531-5731	No	Yes	Yes	P	E, Sp, A2, My, Sm, C	M,R,H,C
Pbx1	chr01: 170124971-5171	Yes	Yes	Yes	GB	E, Sp, My, Mf	M,R,H,C,Md,G
	chr01: 170125231-5431	Yes	Yes	Yes	GB	E, Sp, Sm	M,R,H,C,Md
	chr01: 170125941-6141	Yes	Yes	No	GB	Sp	M,R,H,C

a. Presence of one or more conserved (CANN)TG-$N_{7\text{-}10}$-GATA or (CANN)TG-$N_{7\text{-}10}$-CANNTG paired motifs within the Ldb1 complex binding site.
b. Location of Ldb1 complex binding site. P, promoter; GB, gene body.
c. Presence of conserved transcription factor binding motifs within Ldb1 biding sites (+/- 200bp). Abbreviations: My, Myb; Nk, Nkx, E, ETS; A1, AP-1; Sp, Sp1; NF, NF-E2; A2, AP-2; C, c/EBP; G, Gfi-1; F, FOXO; Sm, Smad; Me, Meis1; P, Pu.1; H, HOXA10; O, Oct; Mf, Mef2c; R, Runx.
d. Conservation of transcription factor binding motifs. Species designations, M: Mus musculus, H: Homo sapiens, R:Rattus norvegicus, C: Canis familiaris, Md: Monodelphis domestica, G: Gallus gallus, X: Xenopus tropicalis.

Supplementary Table 3. Ldb1 complex binding sites within known HSC gene regulatory elements.

Number	RE Name**	RE Coordinates	Ldb1 binding site location	Fold (ChIP) enrichment	*p-value*
1	*Erg+85*	chr16:95548099-95548301	chr16: 95548121-95548321	18.71	1.40E-04
2	*Lmo2-75*	chr2:103694619-103694920	chr2: 103694411-103694611	8.91	4.60E-04
			chr2: 103694691-103694891	70.40	1.30E-05
3	*Lmo2-12*	chr2:103757770-103757982	chr2: 103757771-103757971	41.88	2.70E-05
4	*Scl+40*	chr4:114593666-114594084	chr4: 114593481-114593681	16.93	1.80E-04
			chr4: 114593741-114593941	109.61	4.00E-06
5	*Lyl1-P1*	chr8:87589965-87590364	chr8 : 87590191-87590411	81.09	1.10E-05
6	*Lyl1-P2*	chr8:87591444-87592090	chr8: 87591671-87591871	163.97	1.00E-06
7	*Gata2-77*	chr6:88081954-88082945	chr6: 88082091-88082291	16.04	1.80E-04
			chr6: 88082541-88082741	16.93	1.80E-04
8	*Gata2+9.5*	chr6:88168700-88168939	chr6: 88168741-88168941	24.95	8.10E-05
9	*Pu.1 (Sfpi1)-14*	chr2:90883769-90884084	chr2: 90883811-90884011	47.23	2.20E-05
10	*Myb-68*	chr10:20917924-20918363	chr10: 20917991-20918191	41.88	2.70E-05
11	*Runx1+23*	chr16:92690492-92690781	chr16: 92690521-92690721	66.83	1.40E-05

** RE, regulatory element. Suffix denotes Kbp up or downstream (- or +) of transcription start site.

Supplementary Table 4. Conserved transcription factor binding motifs within known regulatory elements.

RE Name[a]	Binding site location	GATA	CANNTG	Paired motifs[b]	Ref.	Transcription factor binding motifs[c]	Conservation[d]
Erg +85	chr16: 95548121-8321	Yes	Yes	Yes	1	E, C, Sp, My, G, F	M,R,H,C,Md
Lmo2 -75	chr2: 103694411-4611	Yes	No	No	2	E, Nk G	M,R
	chr2: 103694691-4891	Yes	Yes	Yes		E, My, A2, Sp	M,R,H
Lmo2 -12	chr2: 103757771-7971	Yes	Yes	Yes	2	E, C, Me, Sm, Nk	M,R
Scl +40	chr4: 114593481-3681	Yes	Yes	No	3	E, Sp	M,R,H
	chr4: 114593741-3941	Yes	Yes	Yes		E, Sp, Sm	M,R,H,C
Lyl1-P1	chr8 : 87590201-0401	No	Yes	No	4	E, A2, C	M,R,H
Lyl1-P2	chr8: 87591671-1871	Yes	No	Yes	4	E, A2, H, Sp	M,R,H,C
Gata2 -77	chr6: 88082091-2291	Yes	No	No	5	E, G,	M,R,H,C
	chr6: 88082541-2741	Yes	Yes	Yes		E, A2, Sp, Me, R, C, Nk, My	M,R,H,C,Md
Gata2 +9.5	chr6: 88168741-8941	Yes	Yes	Yes	6	E, C, Sp, O, Mf, G	M,R,H,C,Md
Pu.1 -14	chr2: 90883811-4011	No	Yes	Yes	7	E, R, A2, Sp, C, My	M,R,H,C,Md,G
Myb -68	chr10: 20917991-8191	Yes	Yes	Yes	1	E, G, C, My, Sm	M,R,H,C
Runx1 +23	chr16: 92690521-0721	Yes	Yes	Yes	8	E, Sp, R, My, G, Me	M,R,H,C
Scl +19	none	N/A	N/A	N/A	9	N/A	N/A

a. Location of Ldb1 complex binding sites. Suffix denotes Kbp up or downstream (- or +) of transcription start site. RE, regulatory element.
b. Presence of one or more conserved (CANN)TG-$N_{7\text{-}10}$-GATA or (CANN)TG-$N_{7\text{-}10}$-CANNTG paired motifs within the Ldb1 complex binding site.
c. Presence of conserved transcription factor binding motifs within Ldb1 biding sites (+/- 200bp). Abbreviations: Myb; Nk, Nkx, E, ETS; A1, AP1; Sp, Sp1; NF, NF-E2; A2, AP-2; C, c/EBP; G, Gfi-1; F, FOXO; Sm, Smad; Me, Meis1; P, Pu.1; H, HOXA10; O, Oct; Mf, Mef2c; R, Runx.
d. Conservation of transcription factor binding motifs. Species designations, M: Mus musculus, H: Homo sapiens, R:Rattus norvegicus, C: Canis familiari, Md: Monodelphis domestica, G: Gallus gallus, X: Xenopus tropicalis

Supplementary Table 5. Primer sequences for RT-PCR.

Target gene	5'-Primer	3'-Primer	Reference*
Ldb1	5'-TGCTGAAGTGCCACGTCTTT-3'	5'-CACGCTACTTCCGAAGCATTT-3'	
Lyl1	5'-AGATGAGGAAACGCCCTGTA-3'	5'-AGCCACTGCAAGTAGCCTGT-3'	4
Scl	5'-CTCACTAGGCAGTGGGTTCTTT-3'	5'-GACCATCAGAAATCTCCATCTCAT-3'	10
Tcfe2a	5'-AAGAGGACAAGAAGGACCTGAA-3'	5'-TTATTGGCCATACGCCTCTC-3'	
Lmo2	5'-GAGAGACTATCTCAGGCTTTTTGG-3'	5'-TTGAAACACTCCAGGTGATACACT-3'	10
Gata2	5'-GGCTCTCTTGGTGTCTCTTACTCT-3'	5'-GTCCACTACTGTGTCTTGGGAAC-3'	10
Erg	5'- AGGAGCTGTGCAAGATGACA-3'	5'- TCAGATGTGGAAGGGGAGTC-3'	
Pbx1	5'-GACAACTCAGTGGAGCATTCC-3'	5'-CATCACGTGGGTGGTGAAT-3'	
Meis1	5'-GTGGCACTACATGTAACCTTCATC-3'	5'-GTATAACTCGGCTGTCCCATACTC-3'	10
Runx1	5'- CTCCGTGCTACCCACTCACT -3'	5'- ATGACGGTGACCAGAGTGC -3'	
Cebpa	5'- AGGACACGGGGACCATTAG-3'	5'- TAGACGTGCACACTGCCATT-3'	
c-Myb	5'-CAGAAGAGGAGGACAGAATCATTT-3'	5'-TTCCAGTGGTTCTTGATAGCATTA-3'	10
Pten	5'- AGGCCAACCGATACTTCTCTC-3'	5'- CATCTGGAGTCACAGAAGTTGAA-3'	
Mll-1	5'-AAGGAAGACTGTGAAGCAGAAAAT-3'	5'-AACACACTTGGCAATACACAAACT-3'	10
Foxo3a	5'- GCTAAGCAGGCCTCATCTCA-3'	5'- TTCCGTCAGTTTGAGGGTCT-3'	
Rae28	5'-GAGGAGGTCTACGAGTTTATTGCT-3'	5'-TTAAGTAATAAAAGGGCCTGTCCA-3'	10
Id1	5'- GCGAGATCAGTGCCTTGG-3'	5'- TTTTTCCTCTTGCCTCCTGA-3'	
Irf2	5'-TGATGAAGAGAACGCAGAGG-3'	5'-TTGTTGGAGGTGACAAAGGTC-3'	
Ep300	5'- CTGGGTCCTTTGCCAACA-3'	5'- GTGGGAAATATGGCTTGAACA-3'	
Mad1l1	5'- CACAGAGGAGCTGGAAGGTC-3'	5'- GCTCTGCAGCTTTGCAAGTA-3'	
Stat3	5'-GCTGGACACACGCTACCTG-3'	5'-GACTCTTTGCTGGCTGCAT-3'	
Stat5a	5'- TCATCATCCAGTACCAGGAGAG-3'	5'- GGTCTCCAGGGACACTTGC-3'	
Bmi1	5'-GCCTAAGGAAGAGGTGAATGATAA-3'	5'-ATCTGGAAAGTATTGGGTATGTCC-3'	
Etv6	5'-TGCTCTATGAACTCCTTCAGCATA-3'	5'-TCTGGACACAGTTATCTTCGTCAT-3'	10
Zfpn1a1	5'- TGTGGTTATCACAGCCAGGA-3'	5'- CCGGGAAGGAGGCATAAC-3'	
Zfx	5'-GCAGTGCATGAACAGCAAGT-3'	5'-GCAAGGTGTTCAGGATGGTT-3'	
Crebbp	5'-ATAATCGTAAAACGTCCCGTGTAT-3'	5'-GTCTGTGGGGAGAACTCATACTTT-3'	
Sfpi1	5'- CAACGTCCAATGCATGACTACT-3'	5'- TCTCAAACTCGTTGTTGTGGAC-3'	
Gfi1	5'-CAGCTTTGACTGTAAGATCTGTGG-3'	5'-TCTTCATATCTGACTTCTGGTGGA-3'	
β-actin	5'-ATGAAGATCCTGACCGAGCG-3'	5'-TACTTGCGCTCAGGAGGAGC-3'	11
Nanog	5'- AGGGTCTGCTACTGAGATGCTC-3'	5'-CAACCACTGGTTTTTCTGCCAC-3'	
Rex1	5'- AAAGTGAGATTAGCCCCGAG-3'	5'-TCCCCATCCCCTTCAATAGCA-3'	
Oct4	5'-TTGGGCTCCCTTCTTGCT-3'	5'- AATGGGAACAGGGAAACAT-3'	

*Un-referenced primer sets were obtained from Roche and were selected from the Roche universal probe library real-time PCR Assay Design Center.

SUPPLEMENTARY METHODS

Flow Cytometry and Cell Cycle Analysis

The lineage marker (Lin) mixture for fetal liver cells included the following biotinylated antibodies: CD3ε (145-2C11), CD4 (GK1.5), CD8α (53-6.7), CD8β (53-5.8), TCRβ (H57-597), TCRγδ (GL3), CD19 (1D3), B220 (RA3-6B2), Gr1 (RB6-8C5), Ter119, CD49b (Dx5), NK1.1 (PK136). Lin mixture for adult bone marrow cells included all of the above plus Mac-1 (M1/70). Other conjugated antibodies used for surface staining included: CD43 (S7), CD45.1 (A20), CD45.2 (A104), CD48 (HM48-1), CD150 (9D1), c-kit (2B8), Sca1 (D7), Flt3 (A2F10.1), Flk1 (Avas 12α1), IgD (11-26), IgM (II/41), Ly-9.1 (30C7), CXCR4 (2B11). Biotinylated primary antibodies were detected by incubation of antibody coated cells with streptavidin-PerCP-Cy5.5 or APC-Cy7 in a two step staining procedure. Percent apoptotic cells was determined by Annexin V (BD Biosciences) staining according to the manufacture's instructions. For cell cycle analysis, cells were surface stained, fixed in 2% formaldehyde, permeabilized in 0.1% NP-40, then stained with DAPI (Molecular Probes, Eugene, OR) and Ki-67 (BD Biosciences). Data were analyzed with FlowJo software (Tree Star Inc.).

Ldb1 Antibody and Ldb1 Intracellular Staining

Polyclonal rabbit anti-sera were raised against Ldb1 peptide MLDRDVGPTPMYPPTYLEPC-amide (Biosource, Hopkinton, MA), affinity purified, and conjugated with Alexa-647 (Molecular Probes). Purified Rabbit IgG

conjugated in the same way served as background control. Single cell suspensions were surface stained, fixed with 2% formaldehyde, permeabilized with 0.5% triton X-100 in PBS, stained with Alexa 647 conjugated rabbit anti-Ldb1 or rabbit IgG and analyzed by flow cytometry.

Generation of Rosa-*lacZ*$^{+}$ ESC

For tracking ESC contribution in chimeric mice, the *lacZ* gene was inserted into the *Rosa26* locus in *ldb1*$^{-/-}$ ESC by homologous recombination using the Rosaβgeo vector (gift of P. Soriano). Correct targeting was confirmed by Southern blot hybridization.

Genotyping and Deletion Analysis of Tie2-Cre *Ldb*$^{fl/\Delta}$ Mice and Mx1-Cre *Ldb*$^{fl/fl}$ Mice

PCR genotyping of Tie2-Cre *Ldb*$^{fl/\Delta}$ mice and Mx1-Cre *Ldb*$^{fl/fl}$ mice was performed on DNA isolated from mouse tail. Primers used for detection of the Cre transgene:

5'-CGATGCAACGAGTGATGAGG-3' and 5'-GACTTGCTGTCACTTGGTCGT-3'.

The wild type *Ldb1*, the *Ldb1*^{fl} and the Cre-mediated deleted (*Ldb*$^{\Delta}$) alleles were detected using a combination of 3 oligonucleotide primers:

Ldb1A 5'-TCAGGCTGGCCTTTAAACCTAA-3',

Ldb1B: 5'-TGGGACTACAAGGCTGAGAACA-3', and

Ldb1C: 5'-TGGCTGAGCTTATGTGACCA-3'.

Ldb1B-Ldb1C amplified a 458 bp wild type allele fragment and a 534 bp floxed allele fragment. Ldb1A-LdbC amplified a 342 bp Ldb1 deleted allele fragment.

RT-PCR and Real Time Quantitative RT-PCR

Primer sequences are summarized in supplementary Table 5.

Template preparation for ChIP-Seq analysis

Total bone marrow cells were isolated from B6 mice by flushing bone marrows with serum free medium (Mediatech, Inc.). Progenitor-enriched cells were obtained from total bone marrow by removal of lineage positive cells with magnetic beads (Miltenyi Biotech). The lineage cocktail included anti-CD3, CD4, CD8, CD11b, CD19, B220, Gr1, Ter119, NK1.1, DX5. Cells were crosslinked with formaldhyde, and genomic DNA was fragmented by sonication and subjected to chromatin immunoprecipitation (ChIP) with antibodies specific for Ldb1 (Santa Cruz, sc-11198), SCL (Santa Cruz, sc-12984), GATA-2 [5] or control IgG. The ChIP-Seq analysis was performed as described [12]. Briefly, approximately 200 ng of ChIP DNA were end-repaired using the Epicentre DNA END-Repair kit followed by treatment with Taq polymerase to generate a protruding 3' A base used for adaptor ligation. Following ligation of a pair of Solexa adaptors to the repaired ends, the ChIP DNA was amplified using the adaptor primers for 17 cycles and the fragments of approximately 220 bp (mononucleosome + adaptors) were isolated from an agarose gel. The purified

DNA was used directly for cluster generation and sequence analysis using the Solexa 1G Genome Analyzer according to the manufacturer's protocols.

ChIP-Seq data analysis

Twenty five base pair reads obtained from the sequencing were mapped to build mm8 of the mouse genome. Only those reads that mapped to unique genomic locations with at most two mismatches were retained for further analysis. This resulted in 6.0, 11.1, 7.1 and 5.2 million reads for Ldb1, Scl, Gata2 and control IgG samples, respectively, which were processed further using the SISSRs algorithm [13] to identify Ldb1 binding sites. MEME [14] was used to identify sequence motifs within Ldb1 binding sites. Summary windows, displaying the number of ChIP-ed DNA fragments mapped to 200 bp windows, were used for viewing the ChIP-seq data on the UCSC genome browser, and to generate screenshots.

Colony assays

Progenitor assays were performed by culturing cells in methylcellulose-based IMDM medium containing recombinant Epo, SCF, IL-3 and IL-6 according to the manufacturer's instructions (SemCell Technologies, Vancouver, BC, Canada). For cultures of fetal liver cells and bone marrow cells, 2.5×10^4 to 2.5×10^5 cells were plated in duplicate and scored for colony formation at day 8.

SUPPLEMENTARY REFERENCES

1. Wilson, N.K. *et al.* The transcriptional program controlled by the stem cell leukemia gene Scl/Tal1 during early embryonic hematopoietic development. *Blood* **113**, 5456-5465 (2009).
2. Landry, J.R. *et al.* Fli1, Elf1, and Ets1 regulate the proximal promoter of the LMO2 gene in endothelial cells. *Blood* **106**, 2680-2687 (2005).
3. Ogilvy, S. *et al.* The SCL +40 enhancer targets the midbrain together with primitive and definitive hematopoiesis and is regulated by SCL and GATA proteins. *Mol Cell Biol* **27**, 7206-7219 (2007).
4. Chan, W.Y. *et al.* The paralogous hematopoietic regulators Lyl1 and Scl are coregulated by Ets and GATA factors, but Lyl1 cannot rescue the early Scl-/- phenotype. *Blood* **109**, 1908-1916 (2007).
5. Grass, J.A. *et al.* Distinct functions of dispersed GATA factor complexes at an endogenous gene locus. *Mol Cell Biol* **26**, 7056-7067 (2006).
6. Khandekar, M. *et al.* A Gata2 intronic enhancer confers its pan-endothelia-specific regulation. *Development* **134**, 1703-1712 (2007).
7. Huang, G. *et al.* PU.1 is a major downstream target of AML1 (RUNX1) in adult mouse hematopoiesis. *Nat Genet* **40**, 51-60 (2008).
8. Nottingham, W.T. *et al.* Runx1-mediated hematopoietic stem-cell emergence is controlled by a Gata/Ets/SCL-regulated enhancer. *Blood* **110**, 4188-4197 (2007).
9. Gottgens, B. *et al.* Establishing the transcriptional programme for blood: the SCL stem cell enhancer is regulated by a multiprotein complex containing Ets and GATA factors. *Embo J* **21**, 3039-3050 (2002).
10. Kim, I., Saunders, T.L. & Morrison, S.J. Sox17 dependence distinguishes the transcriptional regulation of fetal from adult hematopoietic stem cells. *Cell* **130**, 470-483 (2007).
11. Carotta, S. *et al.* Directed differentiation and mass cultivation of pure erythroid progenitors from mouse embryonic stem cells. *Blood* **104**, 1873-1880 (2004).
12. Barski, A. *et al.* High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823-837 (2007).
13. Jothi, R., Cuddapah, S., Barski, A., Cui, K. & Zhao, K. Genome-wide identification of in vivo protein-DNA binding sites from ChIP-Seq data. *Nucleic Acids Res* **36**, 5221-5231 (2008).
14. Bailey, T.L., Williams, N., Misleh, C. & Li, W.W. MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res* **34**, W369-373 (2006).