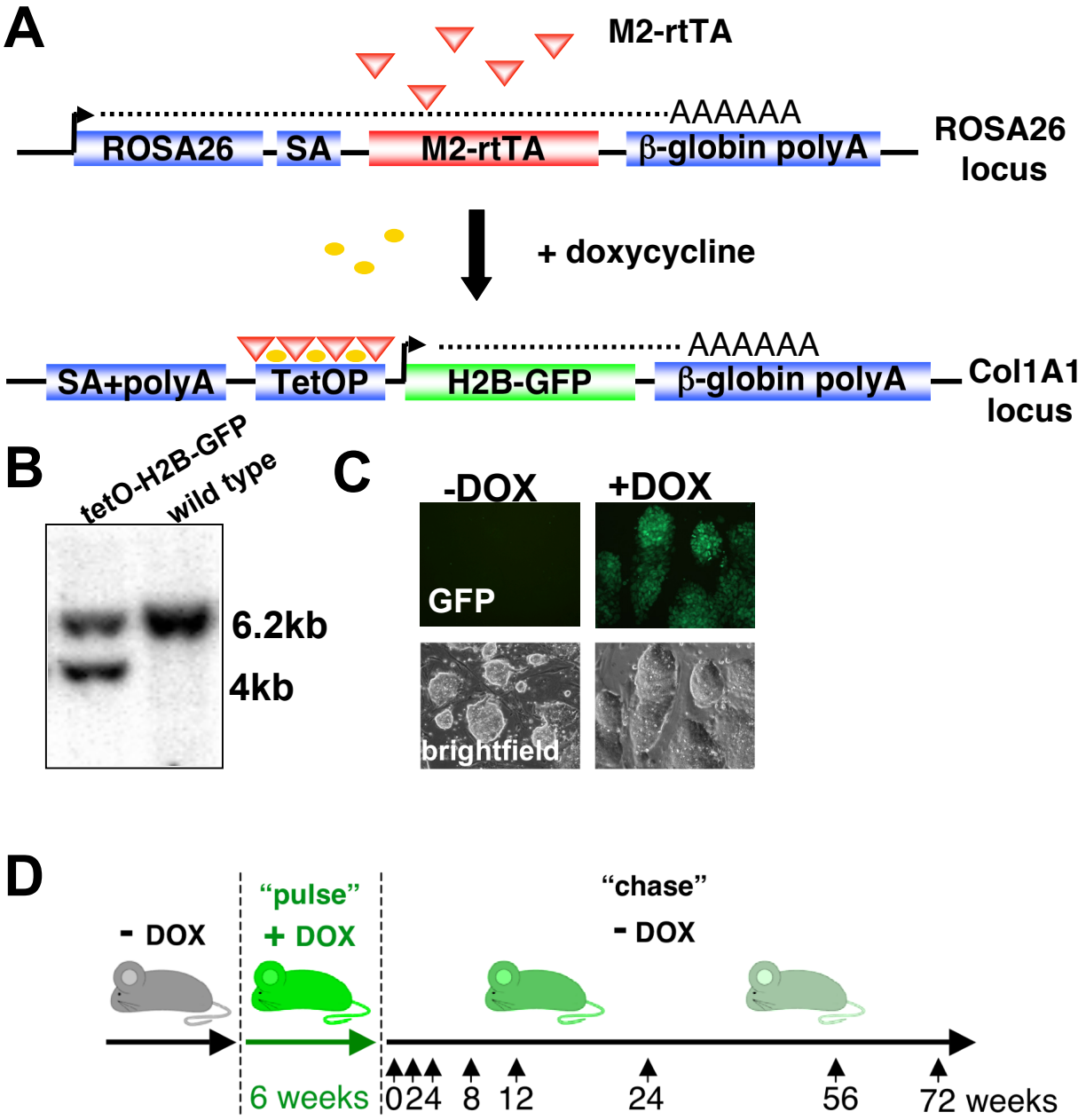
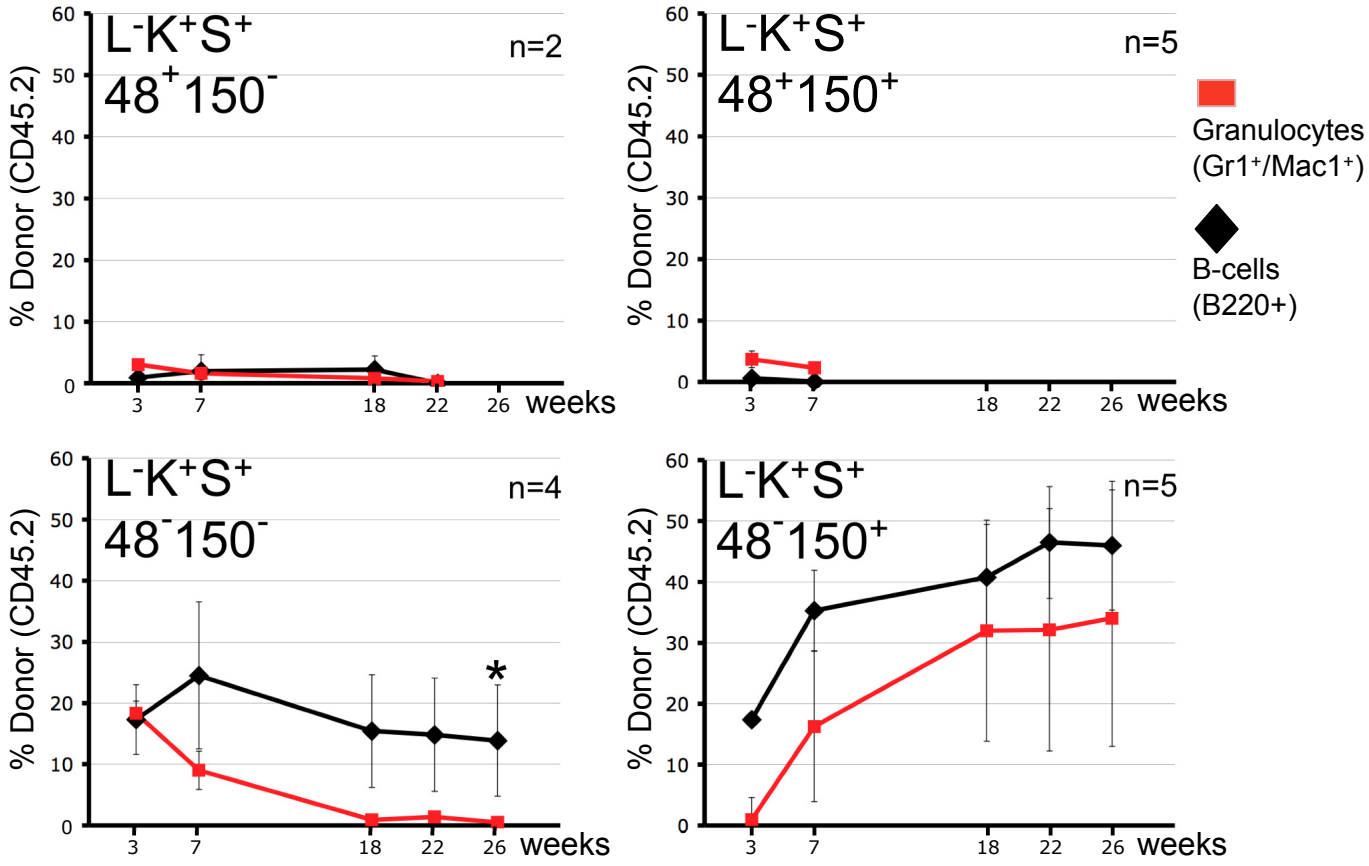
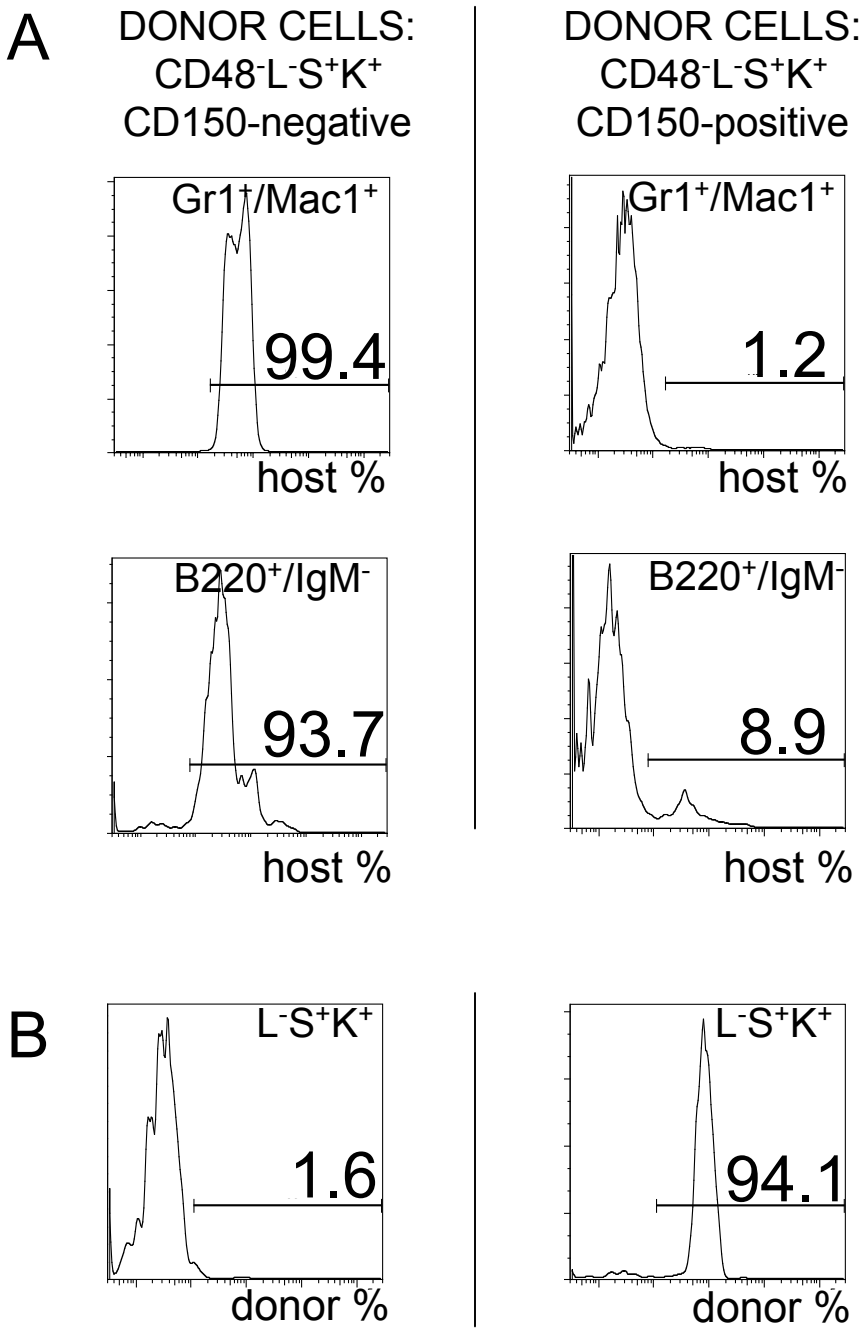
This supplementary file contains:



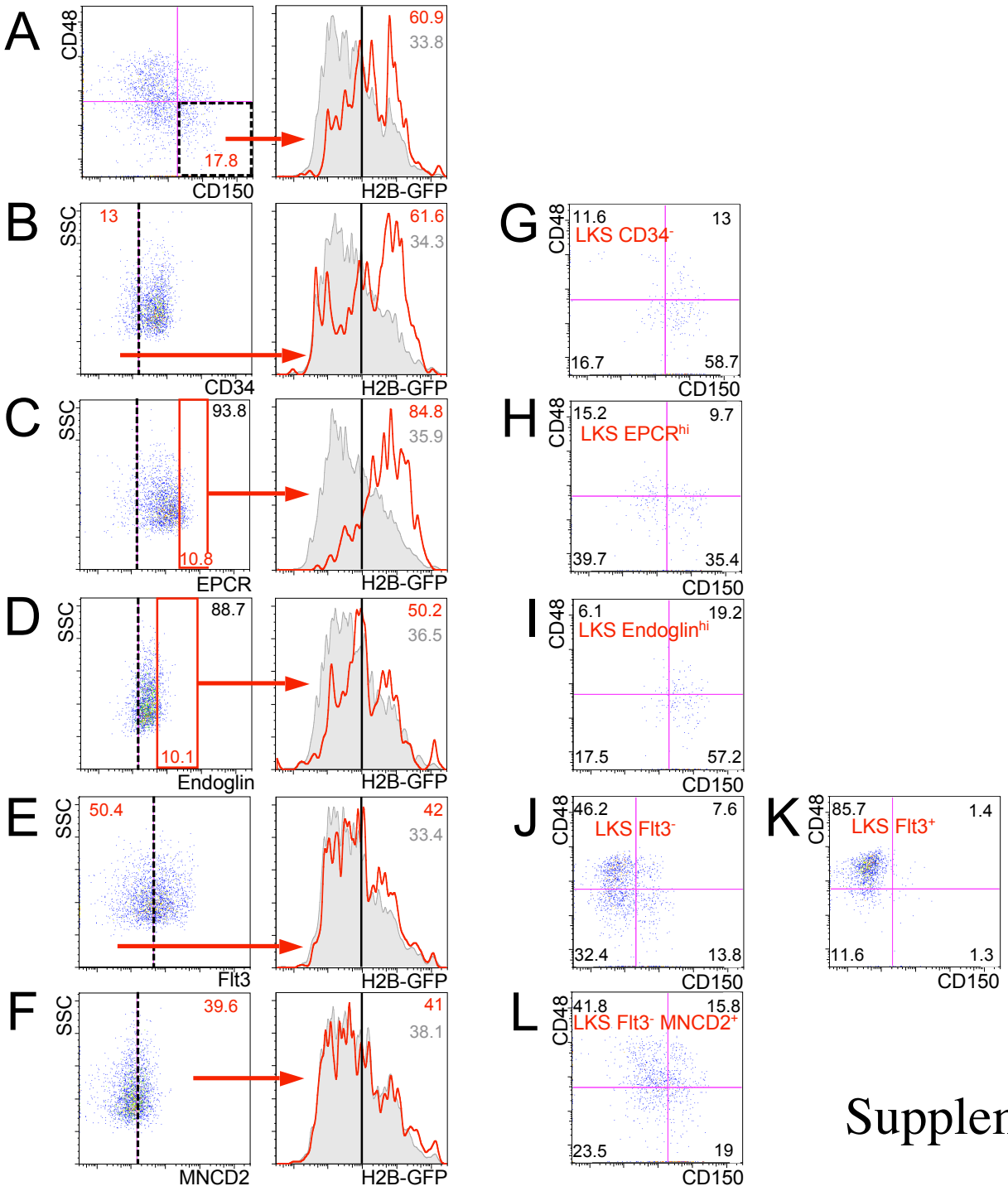
Supplementary Figure 1



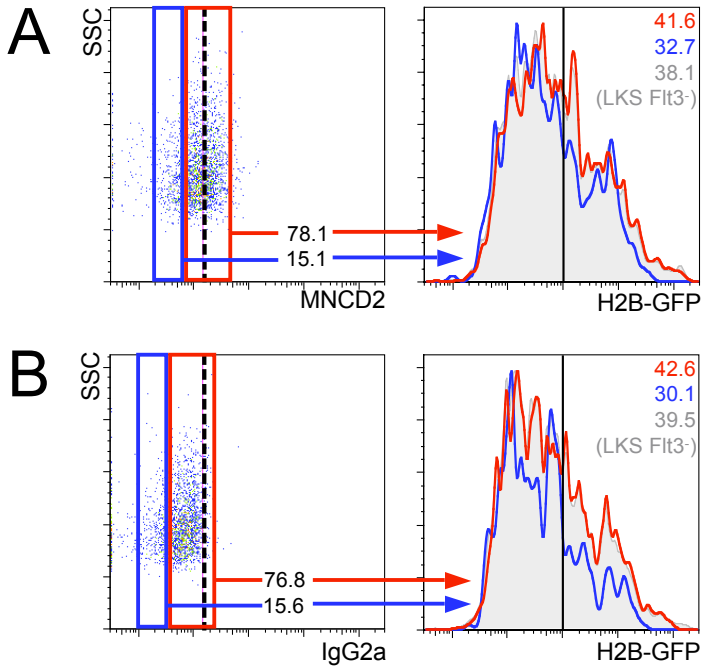
Supplementary Figure 2



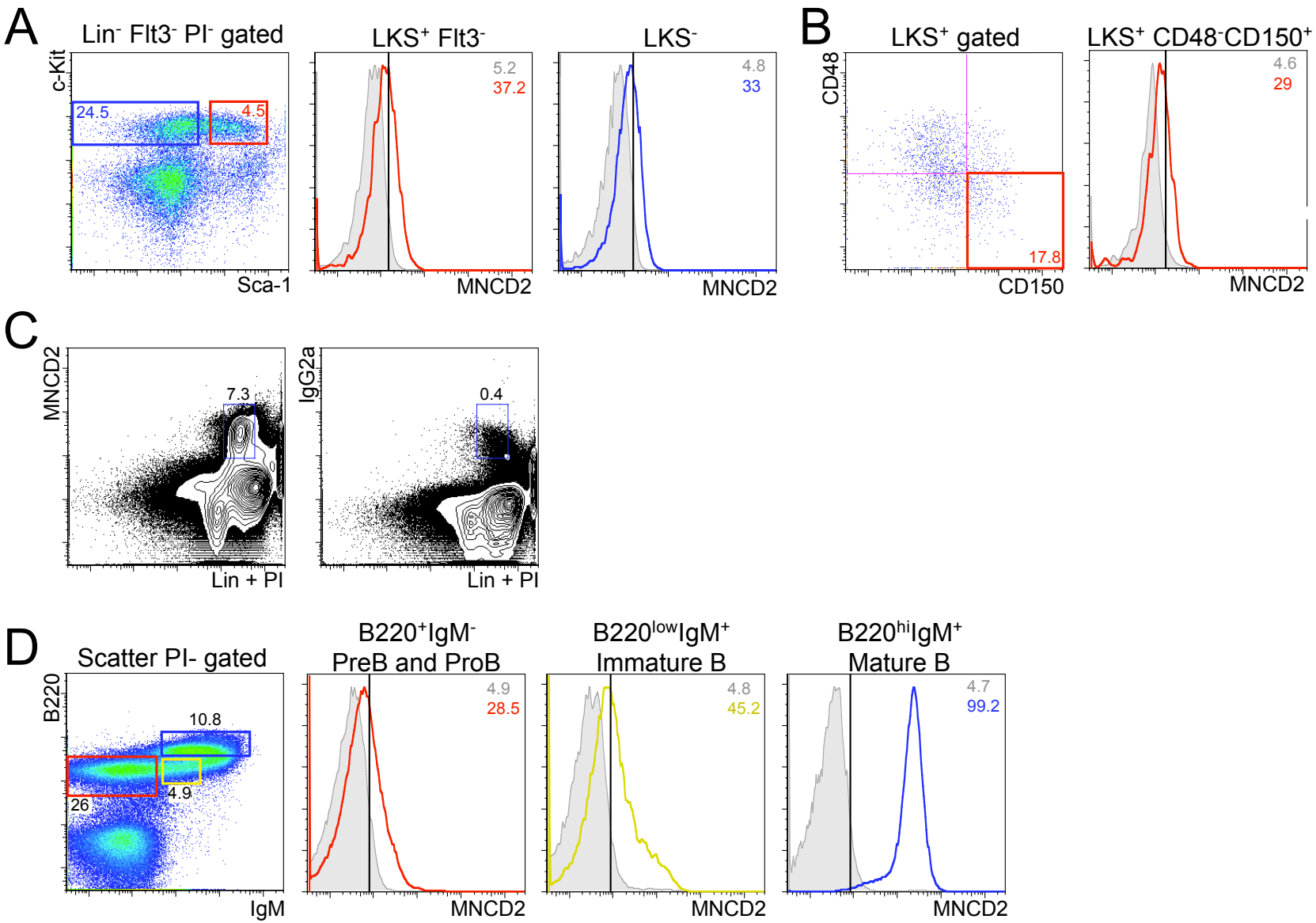
Supplementary Figure 3



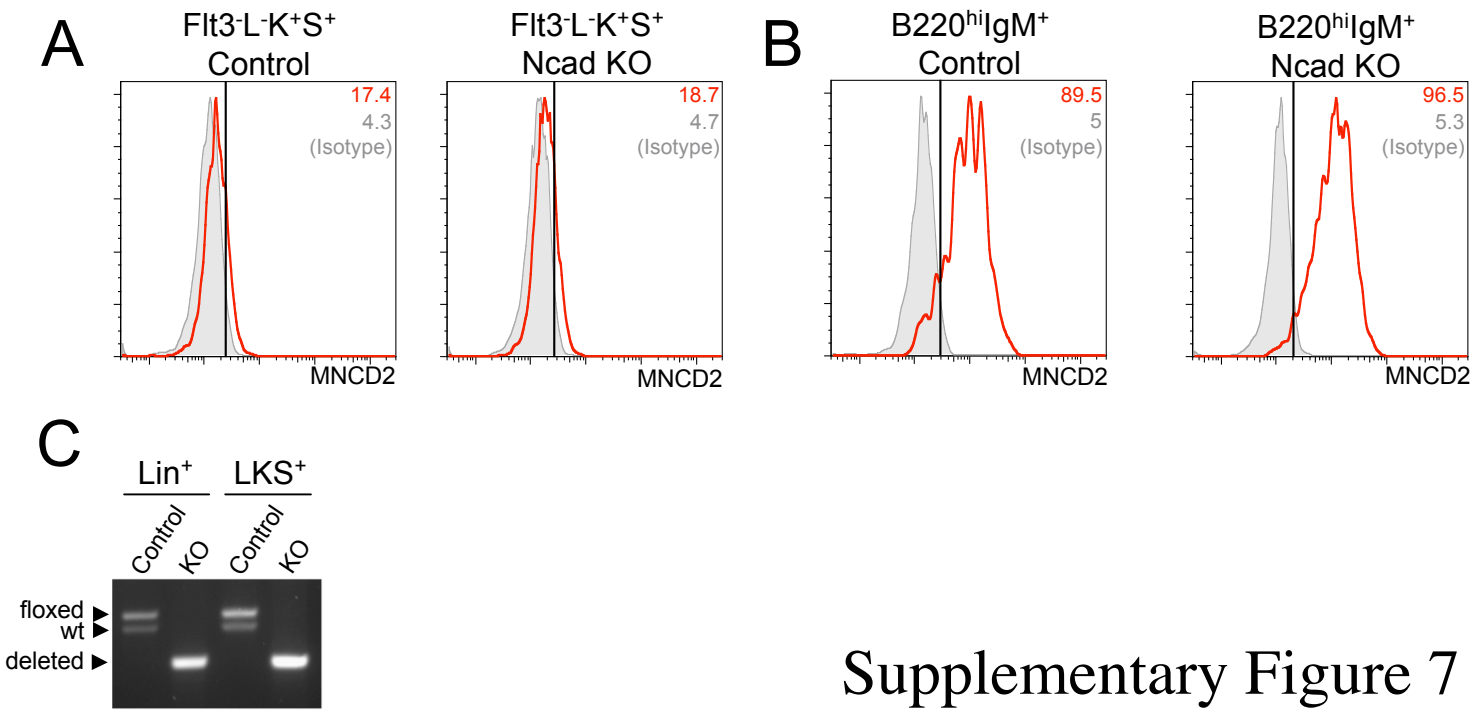
Supplementary Figure 4



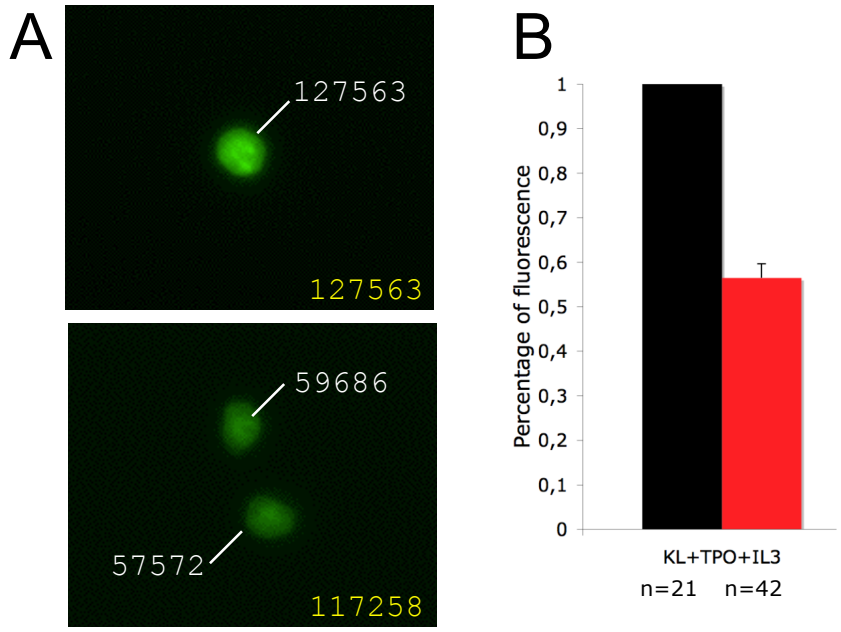
Supplementary Figure 5



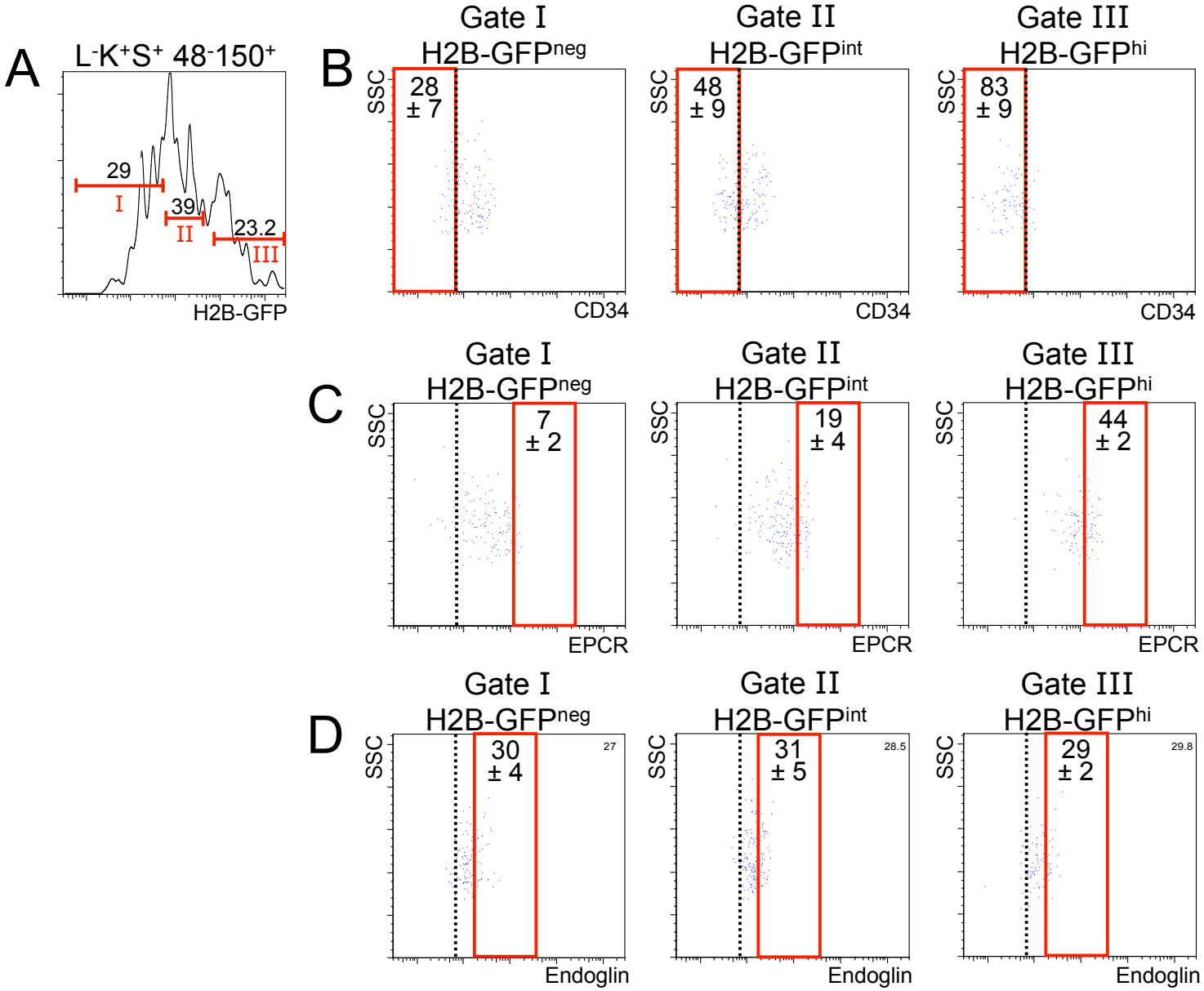
Supplementary Figure 6



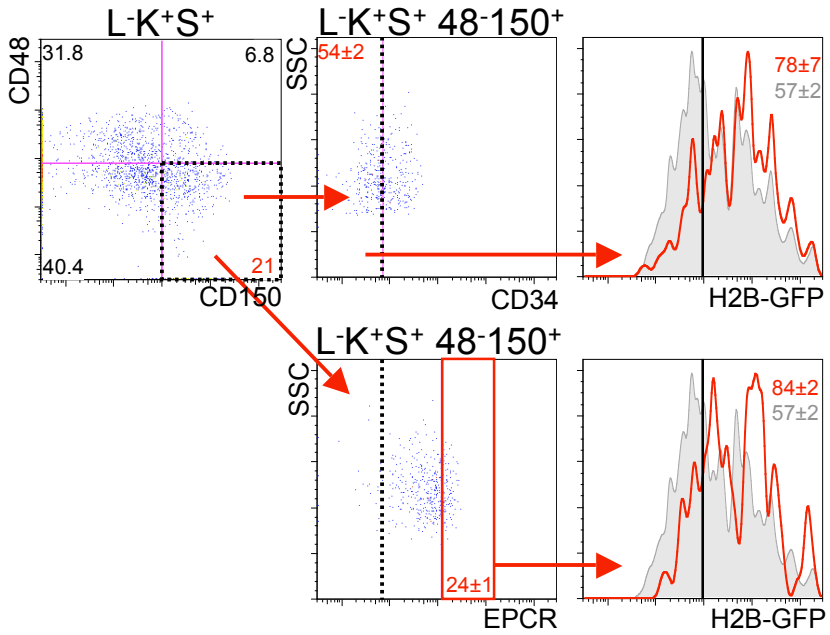
Supplementary Figure 7



Supplementary Figure 8



Supplementary Figure 9



Supplementary Figure 10

II) Supplementary Figure Legends

Supplementary Figure 1: Generation of ES cells and mice that permit inducible expression of H2B-GFP. (A) Alleles used to generate ES-cells and mice with doxycycline-inducible expression of H2B-GFP. (Top) The M2 reverse tetracycline transactivator (M2-rtTA) is driven by the constitutive ROSA26 promoter (R26::rtTA)[31, 32]. (Bottom) A previously characterized H2B-GFP cDNA[8] was inserted downstream of the Collagen (Col) 1A1 locus under the control of a tetracycline-dependent minimal CMV promoter by recombinase-mediated site-specific integration as previously described[31]. SA, splice acceptor; TetOP, tetracycline operator elements fused to CMV minimal promoter. (B) Southern blot analysis confirming integration of the H2B-GFP transgene downstream of Col1A1 locus. (C) ES-cells containing both the R26::rtTA and the TetOP-H2B-GFP allele turn green in the presence of the tetracycline analog doxycycline. These ES cells were used to produce chimeric animals and germline offspring (data not shown). (D) Experimental scheme: mice harboring M2-rtTA and H2B-GFP were subjected to doxycycline in their drinking water for 6 weeks to achieve widespread acquisition of green fluorescence. Subsequently, mice were checked for the retention of H2B-GFP in the absence of doxycycline at defined intervals for up to 72 weeks after the end of doxycycline administration.

Supplementary Figure 2: CD150 expression distinguishes long-term from short-term repopulating HSCs. Analysis of CD45-isotype expression in peripheral blood lymphocytes and granulocytes after transplantation of highly purified early bone marrow populations (donor, CD45.2) into irradiated hosts (CD45.1) together with a small number of support marrow cells (CD45.1) to ensure survival. The four plots show the proportions of donor-derived B-cells and granulocytes in groups of mice that had received 50 cells from sub-fractions of the Lineage$^-$ (L), c-Kit$^+$ (K), Sca1$^+$ (S)-population further resolved by SLAM markers CD48 (48) and CD150 (150) as indicated by the letter/number code in each panel (flow-cytometry of these populations is shown in Figure 1A, first panel on the right). Note: CD48-positive L$^-$K$^+$S$^+$-cells did not give rise to donor-contribution (upper panels). Within the CD48-negative fractions (lower panels), CD150-negative cells (left) gave rise to transient hematopoiesis and CD150-positive cells (right) gave rise to long-term engraftment. *Persistent blood lymphoid cells in lower left panel represent re-circulating B-cells as early bone marrow B-cells are not donor-derived in these mice (see Supplementary Figure S3 for bone marrow contribution analysis of selected mice from lower panels). Individual data points represent means from analysis of groups of mice, number of mice (n) given in each panel, error bars represent standard deviation in upper left panel and standard error of mean in all other panels.

Supplementary Figure 3: Host and donor contributions to bone marrow hematopoiesis after transplantation of CD150-positive and CD150-negative CD48$^-$L$^-$K$^+$S$^+$cells. Bone marrow analysis of selected mice from lower panels in Supplementary Fig. 2. Shown are donor (CD45.2) or host (CD45.1) contributions to bone marrow hematopoiesis after transplantation of CD150-negative (left panels) or CD150-positive (right panels) donor cells 26 weeks after transplantation. (A) Contribution to B-cells and myeloid cells. (B) Contribution to stem cells (LKS). Note: myeloid cells (Gr1$^+$/ Mac1$^+$;

upper panel), early B-cells ($B220^+/IgM^-$, middle panel), and $L^-S^+K^+$ -cells are derived from the donor only after transplantation of $CD150^+$ cells.

Supplementary Figure 4: H2B-GFP retention correlates with multiple alternative markers for long-term repopulating HSCs, but not with expression of MNCD-2. (A-E) Left column shows HSCs ($L^-K^+S^+$) stained with SLAM markers CD48/CD150[4], CD34[14], Endothelial Protein C Receptor (EPCR)[17], Endoglin[18], and Flt3[18], all of which predict long-term repopulation potential. (F) Left plot shows HSCs ($L^-K^+S^+Flt3^-$) stained with MNCD-2, an antibody that has been claimed to[20] recognize N-cadherin by flow-cytometry (see Supplementary Figure 7). Dotted lines in B-F indicate arbitrary threshold set to result in less than 5% events above threshold with isotype-matched control antibody. Right plots show H2B-GFP content of the total population (gray shaded area, $L^-K^+S^+$ in A-E, $L^-K^+S^+Flt3^-$ in F) or gated populations as indicated (red arrows and red curves). Shown is one representative out of three similar experiments performed after 16 weeks of chase (see Supplementary Fig. 1D). (G-L) SLAM-marker profiles of HSCs identified by alternative markers (B-F). Previous gating strategies are indicated on plots in red (as shown in B-F). (G) The majority of CD34-negative HSCs are $CD48^-CD150^+$. (H) $EPCR^{hi}$ HSCs are predominantly CD48-negative. (I) $Endoglin^{hi}$ HSCs are predominantly $CD48^-CD150^+$. (J, K) $Flt3^-$ HSCs contain the majority of the $CD48^-CD150^+$ cells. (L) $MNCD\text{-}2^+$ HSCs are not biased toward a particular SLAM-marker profile.

Supplementary Figure 5: H2B-GFP retention of MNCD-2 or isotype control stained HSCs using alternative gating strategy. (A) H2B-GFP retention of MNCD-2 stained HSCs as shown in Supplementary Figure 4F, but with alternative gates. While we could not discern distinct expression levels of MNCD-2 as reported[20] because of the low intensity of the stain, these gates represent our best approximation of "low" (blue frame) and "intermediate" (red frame) expression levels. Note that there is a small difference in H2B-GFP retention but the same is seen with the isotype control (B).

Supplementary Figure 6: Characterization of MNCD-2 staining in the adult bone marrow. (A, B) Dim staining with MNCD-2 is detectable on HSCs ($Flt3^-L^-K^+S^+$ (A), $CD150^+CD48^-L^-K^+S^+$ (B), red curves) and myeloid progenitors ((A), blue curve) compared to isotype control (gray shaded areas). (C) Bright staining with MNCD-2 (left panel) but not isotype control (right panel) on lineage positive cells. (D) Progressive up-regulation of MNCD-2 during B-cell maturation in the bone marrow. Note that this staining profile does not represent N-cadherin expression (shown in Supplementary Figure 7).

Supplementary Figure 7: Analysis of MNCD-2 binding in the adult bone marrow cells two months after pIpC mediated disruption of a conditional N-cadherin allele using Mx-Cre reveals that MNCD-2 staining is independent of N-cadherin expression. (A, B) MNCD-2 staining of HSCs ($Flt3^-L^-K^+S^+$) (A) and mature B-cells (B) in pIpC-treated control (left panels) and conditional N-cadherin knockout bone marrow[33] (right panels). Note that the MNCD-2 stain was not affected by disruption of N-cadherin. (C) Conditional N-cadherin was excised in sorted HSCs (LKS^+) and sorted

lineage marker$^+$ cells from conditional knockout mice (Genotype: Mx-Cre, N-cadherin floxed/floxed) but not pIpC treated controls (Genotype: No Cre, N-cadherin floxed/ wildtype) (primers as described[33]). Note: These data demonstrate that MNCD-2 does not recognize N-cadherin on hematopoietic cells when used in flow-cytometry. Thus the ligand for MNCD-2 in hematopoiesis is dubious, even if previous studies have shown that MNCD-2 binds to N-cadherin in Western blot analysis of neuronal tissues[34, 35].

Supplementary Figure 8: H2B-GFP of label-retaining HSCs is distributed equally to daughter cells at the time of division. Highly purified HSCs ($CD150^+CD48^-L^-K^+S^+$) from H2B-GFP mice, 20 weeks after the pulse, were single-cell sorted into individual culture wells containing stem cell factor, thrombopoetin, and interleukin-3 to induce division[36, 37]. (A) Fluorescence was quantified before and after the first division. Pixel intensity of individual cells is shown in white numbers, pixel intensity of all cell is shown in yellow. (B) Bar graphs show mean fluorescence of daughter cells (right bar) as the percentage of the fluorescence of the parental cells (left bar, set at 1); number of cells analyzed and standard deviations are given below.

Supplementary Figure 9: CD34 expression and loss of EPCR correlate with loss of H2B-GFP in $CD150^+CD48^-L^-K^+S^+$ HSCs. HSCs ($CD150^+CD48^-L^-K^+S^+$) from bone marrow of H2B-GFP mice 16 weeks after a doxycyline pulse were analyzed for H2B-GFP expression (A) and the expression of CD34[14] (B), EPCR[17] (C), and Endoglin[18] (D) was correlated with absence of H2B-GFP (left panels), intermediate levels of H2B-GFP (middle panels), and high levels of H2B-GFP (right panels). Representative panels from three identical experiments are shown, numbers are averages ± standard deviations, n=3 for all panels. Note: there in a striking inverse correlation between CD34 expression and H2B-GFP retention, and H2B-GFP retention correlates with high EPCR levels. Endoglin shows no correlation with H2B-GFP expression.

Supplementary Figure 10: CD34 and EPCR enhance the correlation with H2B-GFP retention in $CD150^+CD48^-L^-S^+K^+$ HSCs. HSCs ($CD150^+CD48^-L^-K^+S^+$) from bone marrow of H2B-GFP mice 16 weeks after a doxycyline pulse were analyzed for H2B-GFP expression in combination with CD34[14] (upper panels) or EPCR[17] (lower panels). Grey shaded areas and grey numbers in right panels show fluorescence of the entire $CD150^+CD48^-L^-K^+S^+$-population; red curves and numbers show fluorescence of CD34-negative (upper) or $EPCR^{hi}$ HSCs ($CD150^+CD48^-L^-K^+S^+$). Numbers are averages ± standard deviations (n=3).

III) Full Methods:

Mice

TetOP-H2B-GFP mice were generated following our previously described method [31]. R26::rtTA mice and Gfi-1KO mice were previously published [32, 38]. p21$^{Cip1/WAF1KO}$ mice (003263, B6;129Sv), C57bl/6, and B6.SJL (002014) mice were purchased from Jackson Laboratories. Mice were cared for according to the guidelines and supervision of the Massachusetts General Hospital Subcommittee on Animal Research (SRAC). For transgene expression, doxycycline (Sigma D9891, 2 mg/ml, supplemented with sucrose at 10 mg/ml) was added to drinking water for 6 weeks (4 – 8 weeks old mice at the beginning of treatment).

Flow-cytometric analysis and cell sorting

A modified FACS ARIA flow cytometer with five lasers (UV 300 nm, violet 405 nm, blue 488 nm, green 532 nm, red 633 nm) was used for analysis and sorting of HSCs. Immunostains were performed as described [25, 39]. Antibody conjugates were obtained from BD Biosciences or eBioscience, unless otherwise indicated. HSC-stains: lineage markers CD4 (RM4-5), CD8 (53-6), CD3ε (145-2C11), TCRβ (H54-597), B220 (RA3-6B2), Gr1 (RB6-8C5), Mac-1 (M1/70), Ter119 (all Cy-Chrome or PE-Cy5), CD117 (c-Kit, 2B8, APC, Alexa 750), CD135 (Flt3, A2F10, PE), Sca-1 (E13-161, FITC, PE), Sca-1 (D7, Pacific Blue, PE-CY7, Alexa 488; Biolegend) CD48 (HM48.1, APC), CD150 (TC15-12F12.2, PE, Biolegend), CD34 (RAM34, FITC, Pacific Blue), EPCR (RMEPCR1560, biotin), CD105/Endoglin (MJ7/18, biotin), streptavidin-Pacific Blue (Molecular Probes), propodium iodide (1 μg/ml; Molecular Probes). CD45 isotype analysis: B220 (RA3-6B2, APC), Gr1 (RB6-8C5, APC), Mac1 (M1/70, PE), CD45.1 (104, FITC), CD45.2 (A20, PE, FITC). For analysis of DNA content, bone marrow cells were stained with Hoechst 33342 (5 μg/ml) for 1 hour at 37 °C in Hoechst Staining Buffer (HSB) (Hanks Balanced Salt Solution (Invitrogen 14025-076)/2% FCS/ 10 mM Hepes / 1% Penicillin/ Streptomycin) with frequent shaking, protected from light. To prevent dye efflux from HSCs, Reserpine (5 μM) was added 5 minutes prior to Hoechst and kept in all buffers until analysis. After Hoechst incubation, cells were washed in ice-cold HSB and stained with fluorochrome-conjugated antibodies. For cell sorting experiments, lineage-positive cells were depleted prior to flow-cytometry using streptavidin-magnetic beads (M-280; Invitrogen) after labeling with antibody cocktail consisting of biotinylated anti-Gr-1, Mac-1, B220, CD4, CD8, and Ter-119 monoclonal antibodies. Subsequently, cells were stained with appropriate fluorchrome-conjugated HSC markers, including lineage cocktail. Residual biotinylated antibodies were detected with streptavidin-PE-Cy5 and cells were subjected to two rounds of FACS-sorting to achieve virtual purity.

Bone marrow transplantation

Bone marrow transplantations were performed by intravenous injection of sorted donor cells (C57/bl6, CD45.2) and 2 x10^5 competitor bone marrow (2 x 10^5 cells (B6.SJL) into recipients (B6.SJL, CD45.1) that had received 1200 cGy of radiation as a split dose as previously described [25, 39].

***In vitro* culture of single HSCs and quantification of fluorescence**

Rosa26-rtTA/tetO-H2BGFP $Lin^-Sca\text{-}1^+c\text{-}Kit^+$ $CD48^-CD150^+$ HSCs were single cell sorted into Terasaki plates (Nunc 60 MicroWell™, Kamstrup, Denmark) and cultured in 5% CO_2 at 37 °C in 10 µl of StemSpan® Serum-Free Expansion Medium (StemCell Technologies, Vancouver, Canada) supplemented with 50 ng/ml mSCF, 20 ng/ml hTPO, and 10 ng/ml mIL-3 (R&D Systems, USA). Cells were monitored every day for divisions and fresh medium containing cytokines was added every 2 days. Dividing cells were imaged using an inverted Leica DMI4000B fluorescence microscope before and after they had undergone divisions. Images were quantified using ImageJ software (Research Services Branch, NIH), and values were recorded as pixel intensity.

Mathematical modeling

A model for H2B-GFP loss in a stem cell population continuously turning over and maintained at steady state levels was generated by conceptualizing the data underlying Fig. 3 as a variation on the familiar Kaplan-Meier (KM) curve for survival analysis [40]. In the classical approach to statistical modeling of event-time data, data on times (T) elapsed from a common origin to an event of interest are collected, and the KM curve relates time t elapsed on the x axis to S(t) = Pr(T>t) on the y axis. At the origin, S(t) = 1, and as t approaches infinity, S(t) approaches 0. The data presented in Fig. 3 fit this framework if we consider the time to event to be the time elapsed during the succession of divisions leading to loss of H2B-GFP positivity. However, instead of modeling cell lifetimes in the KM framework, we are modeling the duration of time from onset of chase to the event of the K^{th} division, where K denotes the number of divisions for which H2B-GFP can be detected above an arbitrary threshold. If the cell lifetimes are independently and identically exponentially distributed with mean $1/\lambda$, then, owing to the reproductivity property of the exponential model [41] GFP-positivity times will have mean K/λ. We can transform a model for the observed GFP-positivity times (the sum of K lifetimes) to a model for cell lifetimes and thus to an estimate of the cell turnover rate under the assumptions of a fixed K and a common cell lifetime distribution.

Formally, H2B-GFP content in stem cells are measured as proportions $\tilde{\pi}$ of cells gating positive for GFP after various numbers $\tilde{\tau}$ of days of chase (see Fig. 1C, $\tilde{\pi}$ are plotted as open circles in Fig. 3). We assume that each stem cell's lifetime (time elapsed from division of its ancestor to point of its own division) follows an exponential probability distribution, specified by a scalar rate parameter to be denoted λ (given in Fig. 3 as % cells cycling per day), and that at the end of the lifetime, the cell is replaced by a daughter bearing approximately half the mother's GFP content. We also assume that each cell in a lineage is GFP positive up to the K^{th} division, at which point all cells in the future of the lineage are GFP-negative. This threshold division count K is common to all cells. Under these assumptions, with a generational origin determined by experiment for all lineages, the probability that cells at time origin+t have undergone fewer than K divisions is:

$$\pi(\tau;\lambda,K)=\sum_{m=0}^{K-1}\frac{e^{-\lambda\tau}(\lambda\tau)^m}{m!}$$

The foregoing homogeneous model is extended with a two-component model in which one subpopulation of stem cells divides at rate λ_1 and another divides at rate λ_2. If $p \in (0,1)$ is the proportion of cells in the first subpopulation, the extended model is

$$\pi^*(\tau; p, \lambda_1, \lambda_2, K) = p\pi(\tau; \lambda_1, K) + (1-p)\pi(\tau; \lambda_2, K)$$

For various fixed values of K, parameters of models π and π^* are estimated using non-linear least squares on the basis of data values $\tilde{\pi}$ and $\tilde{\tau}$ [42].

IV) Supplementary references

31. Beard, C., Hochedlinger, K., Plath, K., Wutz, A. & Jaenisch, R. Efficient method to generate single-copy transgenic mice by site-specific integration in embryonic stem cells. Genesis 44, 23-8 (2006).
32. Hochedlinger, K., Yamada, Y., Beard, C. & Jaenisch, R. Ectopic expression of Oct-4 blocks progenitor-cell differentiation and causes dysplasia in epithelial tissues. Cell 121, 465-77 (2005).
33. Kostetskii, I. et al. Induced deletion of the N-cadherin gene in the heart leads to dissolution of the intercalated disc structure. Circ Res 96, 346-54 (2005).
34. Matsunami, H. & Takeichi, M. Fetal brain subdivisions defined by R- and E-cadherin expressions: evidence for the role of cadherin activity in region-specific, cell-cell adhesion. Dev Biol 172, 466-78 (1995).
35. Radice, G. L. et al. Developmental defects in mouse embryos lacking N-cadherin. Dev Biol 181, 64-78 (1997).
36. Ema, H., Takano, H., Sudo, K. & Nakauchi, H. In vitro self-renewal division of hematopoietic stem cells. J Exp Med 192, 1281-8 (2000).
37. Takano, H., Ema, H., Sudo, K. & Nakauchi, H. Asymmetric division and lineage commitment at the level of hematopoietic stem cells: inference from differentiation in daughter cell and granddaughter cell pairs. J Exp Med 199, 295-302 (2004).
38. Hock, H. et al. Intrinsic requirement for zinc finger transcription factor Gfi-1 in neutrophil differentiation. Immunity 18, 109-20 (2003).
39. Hock, H. et al. Tel/Etv6 is an essential and selective regulator of adult hematopoietic stem cell survival. Genes Dev 18, 2336-41 (2004).
40. Miller, R. Survival Analysis. NY: Wiley (1981).
41. Johnson, N. L., Kotz, S. & Balakrishnan, N. Continuous Uniform Distributions 2e. NY: Wiley 1, p415 (1994).
42. Bates, D. M. & Watts, D. G. Nonlinear regression analysis and its application. NY: Wiley (1988).