Supplemental methods

Cell culture

MEFs were isolated from E13.5 embryos from timed pregnant C57B/L6 mice as described previously [1]. Briefly, the head and gut were removed from embryos while the remaining body parts were minced and rinsed in 1X PBS. The tissues were pelleted down and incubated in a 50ml conical tube containing 0.1% trypsin/0.1 mM EDTA solution (3ml per embryo) at 37^{o}C for 20min. Afterwards, to dissociate the tissue, a 10ml pipette was used to mix the solution and allowed to sit 5min at room temperature for debris to settle. The supernatant was transferred to a new 50ml conical tube and spun at 200G for 5min. The pellet was resuspended with fresh fibroblast media (10% α-MEM medium containing 1% Penicillin/Streptomycin, 1% glutamine, 1% non-essential amino acid, 1% sodium pyruvate and 10% FBS) and spun again. Cells were then counted and diluted to 1x10^{6} cells/ml. Ten million cells were plated per 10cm dish. The next day the medium was changed to remove the remaining floating cells. MEFs were frozen down at passages 2 and 3. J1 ES cells that had been adapted to feeder-free conditions were cultured on gelatin-coated plates in standard 15% D-MEM medium containing 15% FBS, 10μM 2-mercaptoethanol, 2mM L-glutamine, 1% of nucleoside mix (100X stock, Sigma), 1000U/ml leukemia inhibitory factor (LIF) and 1% Penicillin/Streptomycin as previously described [2, 3].

Reporter construction

The eight truncations of the CD45 promoter region positioned from -164 to +168 of the transcriptional start site were amplified using Ultra Phusion II with primers flanked with AttB recombination sequences. The PCR products were loaded on a 2% agarose gel then purified

using the Qiagen QIAquick Gel Extraction Kit. Gateway cloning was used to move promoters into the lentiviral vector backbone. The purified promoter fragments were cloned into a Gateway adapted entry vector using a BP reaction. Subsequently, the entry vectors underwent LR reactions with a Gateway adapted donor vector containing EGFP to yield the final lentiviral expression vector. For HM1 embryonic stem (ES) cell transduction, EGFP was replaced with mCherry to enable co-transduction with a HOXB4-GFP expression cassette.

Lentivirus production

The lentiviral vector used in these studies is equipped with a self-inactivating (SIN) LTR, woodchuck hepatitis virus post-transcriptional regulatory element (WPRE) and a central polypurine tract (cPPT) to increase biosafety, full-length RNA production and titer, respectively [4-7]. Three days prior to virus production, 293T cells were split on 2 consecutive days. Approximately 24hr before transfection, $5.5x10^6$ 293T cells were seeded in 15cm dish. The next day, 293T cells were co-transfected with 15mg of packaging plasmid, 5mg of VSV-G envelope encoding plasmid, and 20mg vector plasmid per 15cm plate using a calcium phosphate-based transfection method. At 48hr, lentivirus supernatants were harvested, passed through a 0.45μm filter, and concentrated by ultracentrifugation at 26,000 RPM for 90 minutes. The pellets were resuspended in 100ul OptiMEM and frozen down at -80^{o}C. To ensure quantitative measurement of viral transducing units (TU), real-time quantitative PCR (RT-qPCR) was used for obtaining titers, and this was compared by functional titering using CMV-EGFP transduction of HeLa cells as described elsewhere [6]. Viral RNA was isolated using the QIAamp Viral RNA mini Kit. RNA was then subjected to reverse transcription using the Promega GoScript Reverse Transcription system. Primers were designed to anneal to the WPRE region.

Gene expression and regulatory element identification

HSPC (c-Kit^{+}, Sca1^{+}, lineage$^{-/low}$), CMP, GMP, MEP, lymphoid, myeloid and erythroid cell gene expression data were obtained from published data employing Affymetrix high-density oligonucleotide arrays [8, 9]. CD45 gene conservation was obtained from multiple alignments of 30 vertebrate species and measurements of evolutionary conservation using two methods (phastCons and phyloP) derived from the PHAST package [10]. ChIP-Seq datasets were downloaded from GEO (www.ncbi.nlm.nih.gov/geo/, accession numbers: GSE36030, GSE22178, GSE21614, GSE32970, GSE31477), reads were aligned to the mouse (mm9) or human (hg19) reference genomes using Bowtie [11] and peaks were called with MACS [12]. DNAseI hypersensitivity followed by high-throughput sequencing (DNase-seq) [13] datasets were obtained from GEO accession numbers GSE37074 and GSE32970.

In vitro B cell differentiation assay

Bone marrow from 3 age and sex matched mice were harvested using B1 buffer (1X PBS containing 0.5% FBS, and 2mM EDTA). Afterwards, the cell pellet was collected and resuspended in 5ml 1X BD Pharmlyse for 4.5min at RT. The cells were rinsed and washed with 30ml B1 buffer. Subsequently, the pellet was resuspended in 8ml B1 buffer and incubated with an in house monoclonal streptavidin conjugated antibody cocktail (Ter119, Ly6C, Gr-1, IgM, CD5, and B220). The cell and antibody mixture was incubated on a nutator at 4^{o}C for 30 min. Immediately after, cells were spun down and washed with 10ml of B1 buffer at 1,500 rpm for 5min. Next, 0.5ml of washed Dynal Sheep Anti-Rat Beads were added to the cells, and incubated at 4^{o}C for 30 min before magnetic immuno-selection. Cells remaining in the supernatant were spun down and the lineage depleted hematopoietic stem and progenitor cells were resuspended in

an HSPC stimulating cocktail (StemSpan SFEM medium with 50ng/ml mSCF, 100 ng/ml Flt3L, 20 ng/ml IL-3, 50ng/ml IL-11, and 10ng/ml TPO) overnight at 37^{o}C. After 20hr, cells were counted and transduced with an MOI of 30 in the presence of 8μg/ml polybrene. The next day transduced HSPCs were transferred to 12-well tissue culture plate coated with OP9 stromal cells in Opti-MEM medium containing 5ng/ml IL-7, 10ng/ml Flt3, and 10ng/ml mSCF. Flow cytometry was employed to detect GFP, CD19 (PE-Cy7) and B220 (APC) following 10 and 24 days of culture.

ES cell differentiation

Antibodies used for day 10 staining of ES differentiated hematopoietic cells: c-kit-PerCP-eFluor 710 (eBioscience; cat #: 46117180), Sca-1-Brilliant Violet (BV) 421, (Biolegend; cat #: 108127), CD45-APC (eBioscience; cat #: 17045182), Gr-1/Mac-1-BV 650 (Biolegend; cat#: 108441 and 117339).

Antibody staining and flow cytometry

For the in vitro B cells differentiation assay, cells were Fc blocked for 20min to preclude non-specific antigen binding, followed by two rinses (PBS + 3% FBS) and then incubated 30min with antibodies (CD19-PE.Cy7, B220-APC, Mac-1-PE, and Gr-1-PE) and then rinsed twice being processed through a LSR Violet II flow cytometer. For staining of peripheral blood and bone marrow from transplants, cells were incubated in 2ml 1X RBC lysis buffer for 5 min at RT. Subsequently the lysed samples were diluted and washed twice in 8ml of 3%FBS/ PBS. The cells were transferred to a 96 well plate and blocked with 1μg Fc block for 20min before washing and

incubation with the following conjugated antibodies for 30 min: CD45.1-BV421, Mac-1-PE.Cy7, Gr-1-PE, Ter119-APC, and Cd3ε-Fluor710. FlowJo 8.6 was used for analysis of FACS files.

Supplemental references:

1. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. **Cell**. 2006;126:663-676.

2. Wang J, Rao S, Chu J et al. A protein interaction network for pluripotency of embryonic stem cells. **Nature**. 2006;444:364-368.

3. Levasseur DN, Wang J, Dorschner MO et al. Oct4 dependence of chromatin structure within the extended Nanog locus in ES cells. **Genes Dev**. 2008;22:575-580.

4. Levasseur DN, Ryan TM, Pawlik KM et al. Correction of a mouse model of sickle cell disease: lentiviral/antisickling beta-globin gene transduction of unmobilized, purified hematopoietic stem cells. **Blood**. 2003;102:4312-4319.

5. Zufferey R, Nagy D, Mandel RJ et al. Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. **Nat Biotechnol**. 1997;15:871-875.

6. Zufferey R, Dull T, Mandel RJ et al. Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. **J Virol**. 1998;72:9873-9880.

7. Zufferey R, Donello JE, Trono D et al. Woodchuck hepatitis virus posttranscriptional regulatory element enhances expression of transgenes delivered by retroviral vectors. **J Virol**. 1999;73:2886-2892.

8. Su AI, Wiltshire T, Batalov S et al. A gene atlas of the mouse and human protein-encoding transcriptomes. **P Natl Acad Sci USA**. 2004;101:6062-6067.

9. Lattin JE, Schroder K, Su AI et al. Expression analysis of G Protein-Coupled Receptors in mouse macrophages. **Immunome research**. 2008;4:5.

10. Blanchette M, Kent WJ, Riemer C et al. Aligning multiple genomic sequences with the threaded blockset aligner. **Genome Res**. 2004;14:708-715.

11. Langmead B, Trapnell C, Pop M et al. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. **Genome Biol**. 2009;10:R25.

12. Zhang Y, Liu T, Meyer CA et al. Model-based analysis of ChIP-Seq (MACS). **Genome Biol**. 2008;9:R137.

13. John S, Sabo PJ, Thurman RE et al. Chromatin accessibility pre-determines glucocorticoid receptor binding patterns. **Nat Genet**. 2011;43:264-268.

14. Pinto do OP, Kolterud A, Carlsson L. Expression of the LIM-homeobox gene LH2 generates immortalized steel factor-dependent multipotent hematopoietic precursors. **EMBO J**. 1998;17:5744-5756.

Supplementary Figures

Supplemental Figure 1. Human CD45 expression levels are high in multiple blood cell lineages. Affymetrix high-density oligonucleotide arrays [8, 9] were used to measure CD45 levels in the human blood cell lineages and cell lines indicated. Values shown are from duplicate arrays.

Supplemental Figure 2. Mouse CD45 regulatory element discovery using transcription factor chromatin occupancy and nuclease sensitivity mapping. Chromatin immunoprecipitation followed by DNA sequencing (ChIP-seq) and nuclease accessibility (DNase-seq) analyses were used to identify regulatory elements throughout the murine CD45 gene body (top) and a region confined to the proximal promoter and portion of intron 2 (bottom). Transcription factor ChIP-seq data were obtained from the multipotent hematopoietic progenitor cell line 7 (HPC-7) [14]. H3K4me3 enrichment and associated HSs downstream of exon 2 are observed in the thymus, bone marrow derived macrophages (BMDM), spleen and the erythroid MEL cell line.

Supplemental Figure 3. Human CD45 regulatory element discovery using chromatin profiling and nuclease sensitivity mapping. Chromatin immunoprecipitation followed by DNA sequencing (ChIP-seq) and nuclease accessibility (DNase-seq) analyses were used to identify regulatory elements throughout the CD45 gene body (top) and a region confined to the proximal promoter and portion of intron 2 (bottom). The human erythroid progenitor cell line K562 (eryth), acute promyelocytic leukemia cell line NB4 (promyelo), EBV transformed B lymphoblastoid cell line GM12864 (B lympho), promyelocytic leukemia cell line HL-60, T

lymphocyte line Jurkat and primary human cell lineages queried for ChIP-seq and DNase-seq are indicated.

Supplemental Figure 4. CD45 reporters efficiently mark mouse lymphoid, erythroid, and multipotent progenitor blood cell lines. (A): Histograms of a murine nucleated erythroid progenitor cell line transduced with the CD45 reporter constructs indicated. GFP percentage and mean fluorescence intensity (geometric means) were calculated for each construct. **(B):** Percent GFP positivity for B cell, **(C):** erythroid MEL cells, and **(D):** a hematopoietic progenitor cell (HPC) line transduced with the indicated CD45 reporters. FACS measured analyses were done in duplicate and repeated a minimum of 3 times.

Supplemental Figure 5. CD45 expression and its promoter activity in murine cell lines. (A): Flow images of murine lymphoid (CH12LX) (A) and HPC (EML) (B) cell lines staining for CD45 without (top panel) or with CD45 promoter transduction (lower panel).

Supplemental Figure 6. CD45 promoter is active in human hematopoietic cells. Histograms of a human T cell line that was transduced with the CD45 reporter constructs indicated. GFP percentage and mean fluorescent intensity (geometric mean) were calculated for each construct. FACS measured analyses were done in duplicate and repeated a minimum of 3 times.

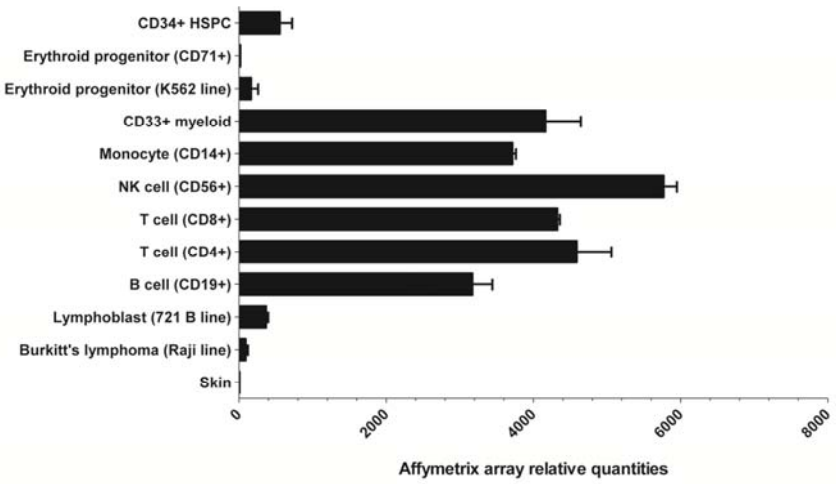
Supplemental Figure 7. CD45 reporters efficiently mark human lymphoid, erythroid and multipotent progenitor blood cell lines. (A): Percent GFP positivity for B cell, (B): erythroid progenitor, (C): T cell and (D): a hematopoietic progenitor cell (HPC) line transduced with the

indicated mouse CD45 reporters. FACS measured analyses were done in duplicate and repeated a minimum of 3 times. **(E):** Human CD45 reporter was used to transduce the HPC and **(F):** B cell lines. The mock control transduction is the red histogram.

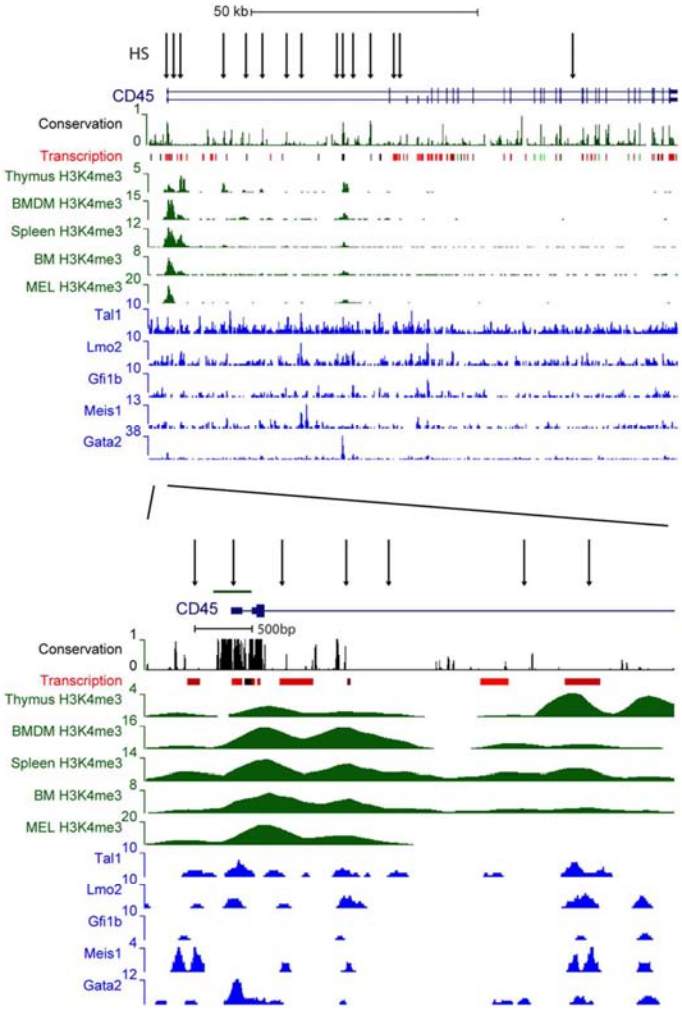
Supplemental Figure 8. High GFP marking following HSPC transplantation without supportive bone marrow. (A): Lineage depleted bone marrow was subjected to a mock or CD45 reporter transduction for 24 hours. The percentage of c-Kit, Sca-1 double positive, lineage depleted (KSL) HSPCs was determined in cells that were unstained (A, left) or stained (B, right) with antibodies directed against c-Kit and Sca-1. **(B):** CD45 reporter transduction results in high-level GFP marking efficiency in the KSL fraction (B, right panel). **(C):** Peripheral blood staining at 16 weeks post transplantation in CD45.2 mice showing two independent transplants (TP #1 and #2) that received CD45 reporter transduced HSPCs without supportive bone marrow. Note the virtually complete marking (95%) of donor transduced reporter cells in transplant 2 (C, bottom left panel).

Supplemental Figure 9. Persistent and robust GFP marking in secondary transplanted recipients. Peripheral blood staining at 22 weeks post transplantation in CD45.2 mice that received bone marrow cells from primary transplant mice. Data displayed is representative from 5 mice/group with an average of 50% chimerism. Approximately 98% of CD45.1 cells are GFP positive compared to mock controls. The reporter efficiently marks the myeloid (Gr-1/Mac-1), nucleated erythroid (Ter-119) and lymphoid (CD3ε, B220) compartments.

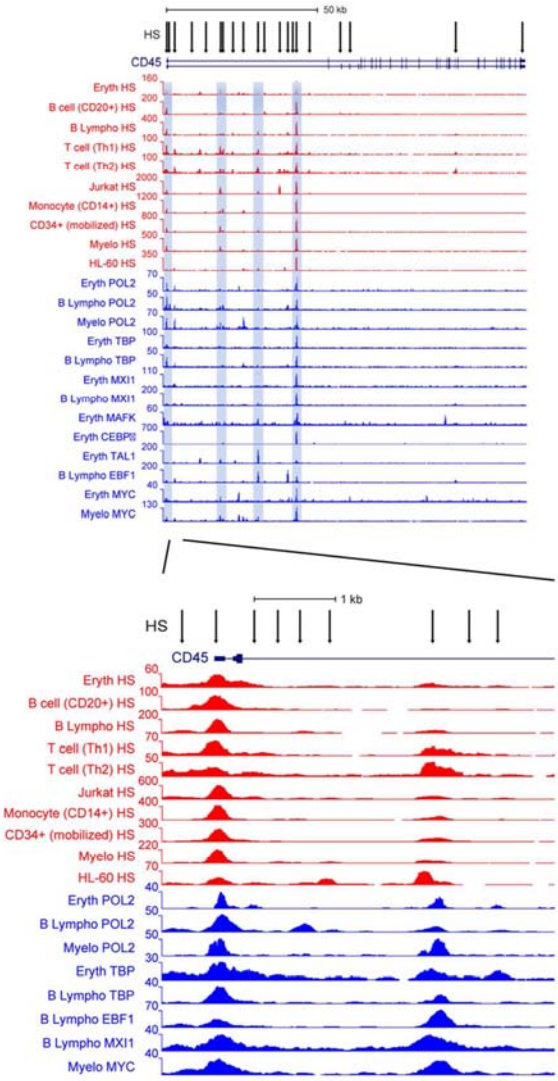
Supplementary Figure 1



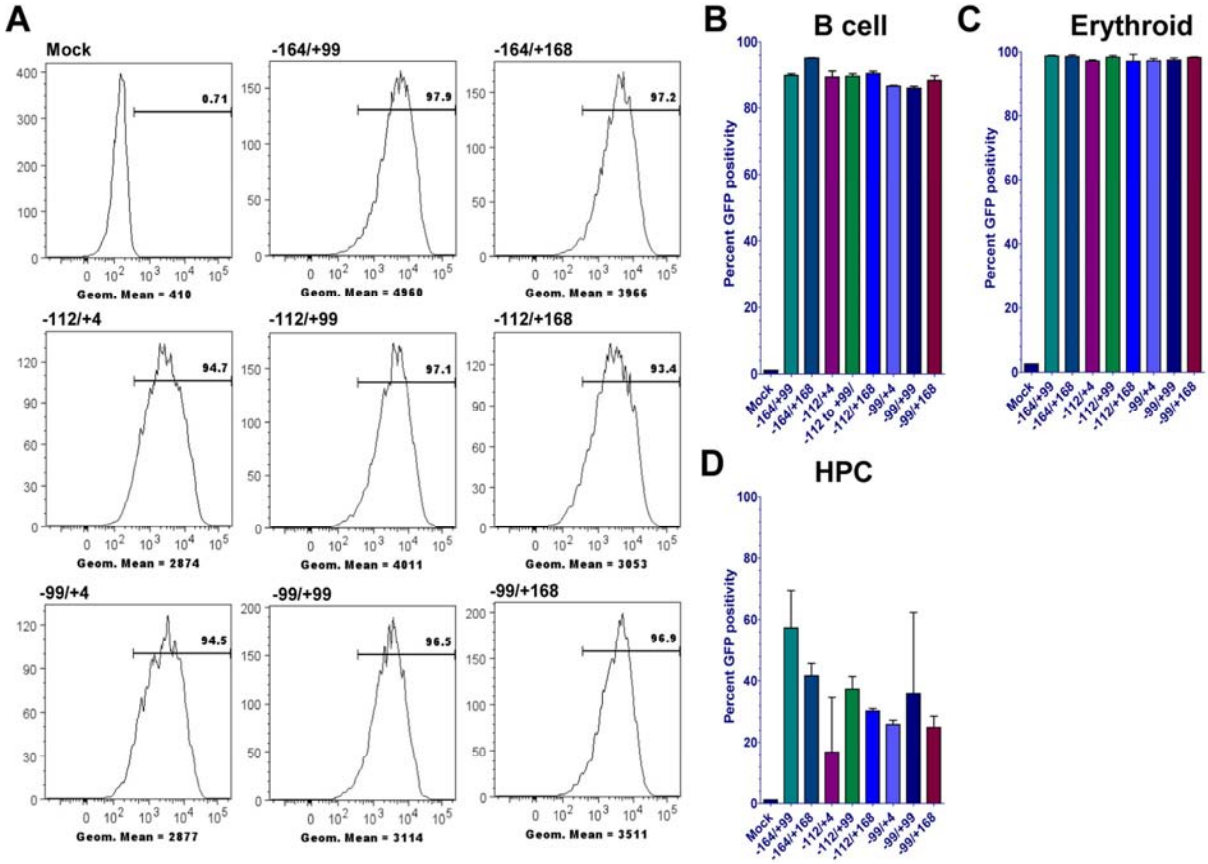
Supplementary Figure 2



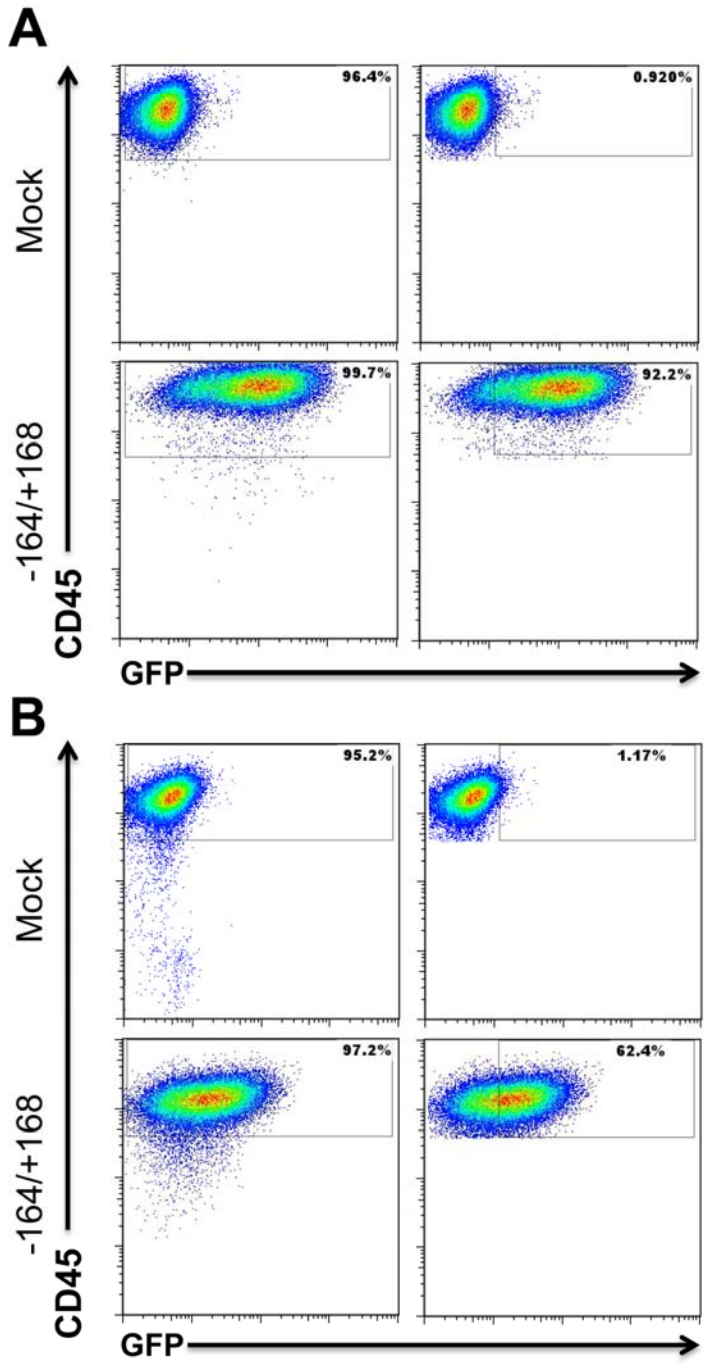
Supplementary Figure 3



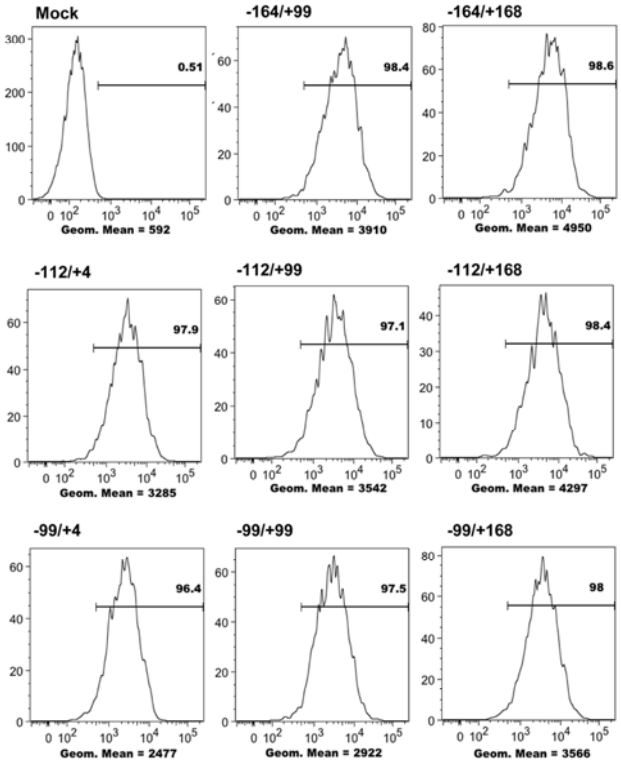
Supplementary Figure 4



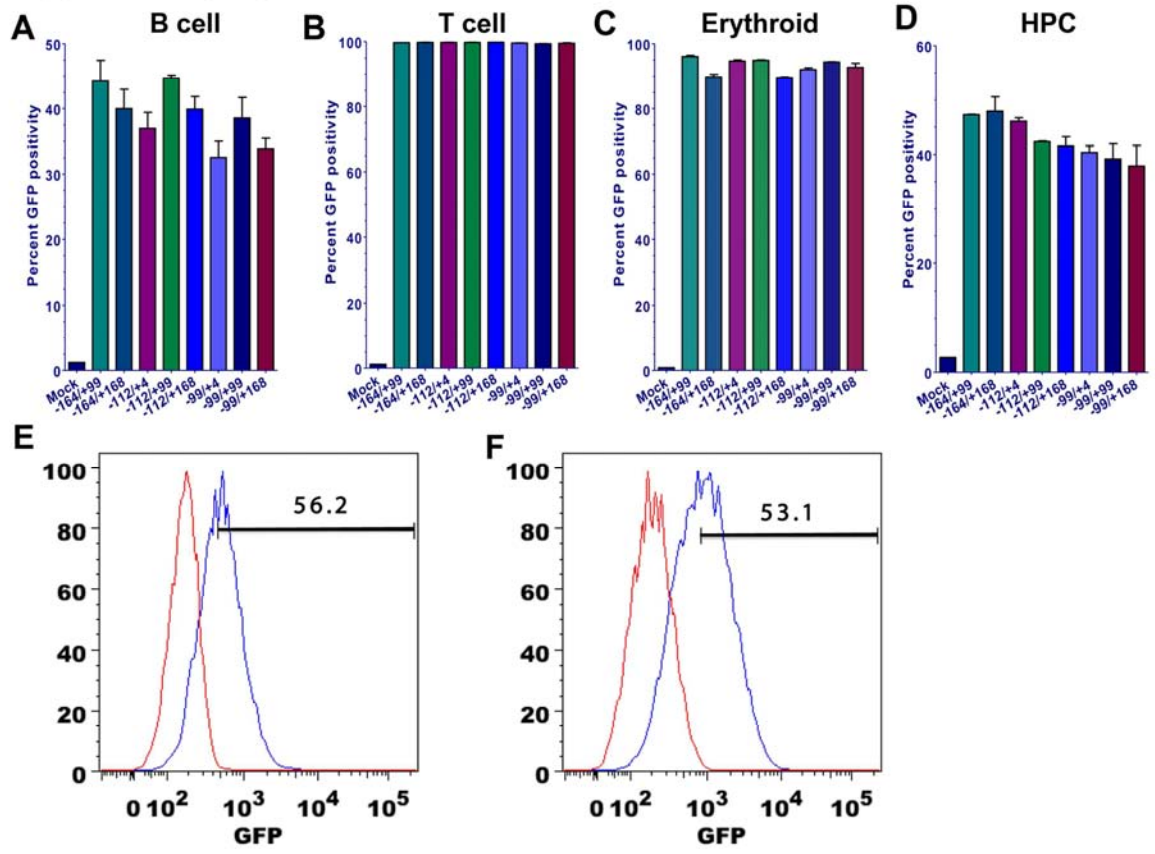
Supplementary Figure 5



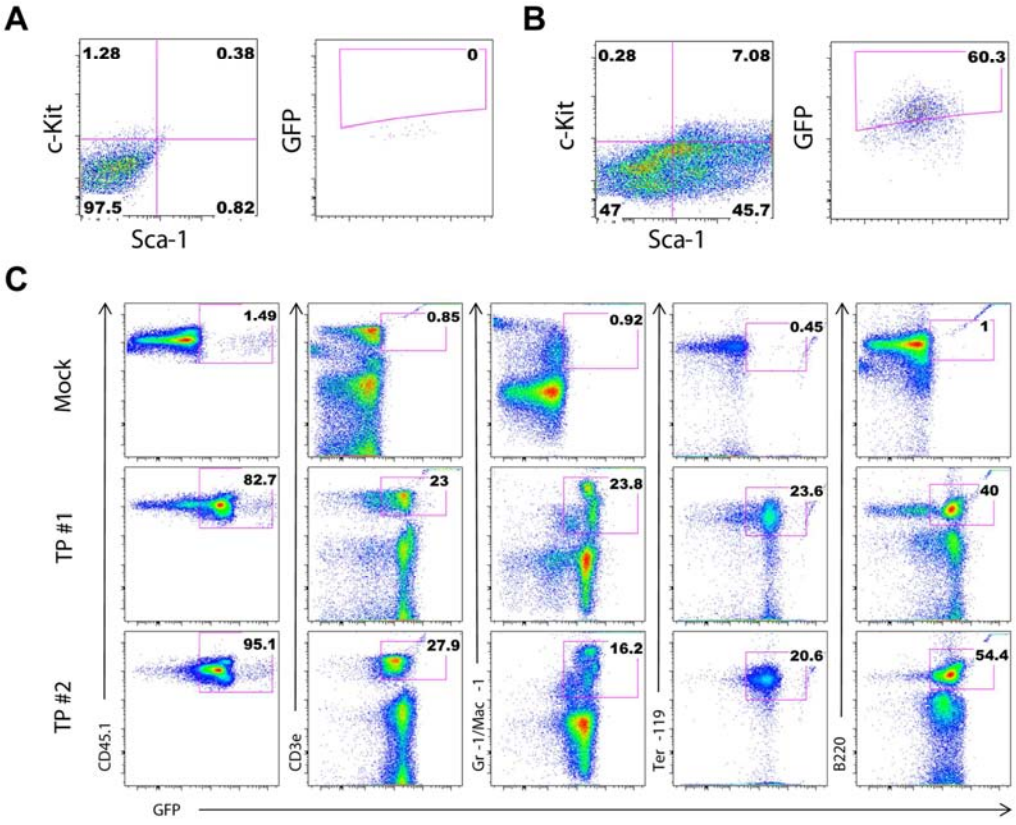
Supplementary Figure 5



Supplementary Figure 6



Supplementary Figure 7



Supplementary Figure 8