Supplemental Materials

Supplemental Figure Legends

Figure S1. Sorting strategy for committed lymphoid and myeloid progenitors, Related to Figure 2.

(A) Sorting strategy for Pro/Pre B cells. Bone marrow cells were pre-gated through doublet discriminated, live (propidium iodide negative) cells. Pre-sort (top panel) and post-sort (lower panel) are shown.
(B) Sorting strategy for common myeloid progenitors (CMPs). Bone marrow cells were pre-gated through live (propidium iodide negative), doublet discriminated cells. Pre-sort (top panel) and post-sort (lower panel) are shown. GMP and MEP gates are also shown for reference. Lineage negative was defined as $Gr1^-$, $Mac1^-$, $B220^-$, $CD3^-$, $CD4^-$, $CD8^-$, $Ter119^-$.

Figure S2. Analysis of donor- and host-derived chimerism in the peripheral blood and bone marrow of reconstituted mice showing lineage contribution and absence of dox-independent expression of the viral reporter cassette, Related to Figures 2-6.

(A-C) Representative peripheral blood analysis of donor and host chimerism 16 weeks post-transplant for experiments in which Pro/Pre B cells were transduced with (A) 36-Factor cocktail, (B) 6-TF cocktail, or (C) 8-TF polycistronic cocktail. Donor and recipient cells were gated as indicated and stained with Mac1, B220, CD3, CD4, CD8 and Gr1. Donor and host cells were also analyzed for ZsGreen to assess dox-independent activation of viral transgenes (right panels
(D) Representative peripheral blood analysis of donor and host chimerism 16 weeks post-transplant for experiments in which myeloid effector cells were transduced with 8-TF polycistronic cocktail. Donor and recipient cells were gated as indicated and stained with Mac1, B220, CD3, CD4, CD8 and Gr1. Donor and host cells were also analyzed for ZsGreen to assess dox-independent activation of viral transgenes (right panels).
(E) Representative bone marrow analysis of a recipient mouse transplanted with Pro/Pre B-Cells transduced with 8-TF polycistronic cocktail analyzed 16 weeks post-transplant showing both host- and donor-derived contribution to the primitive LSK compartment pre-gated through live, doublet discriminated lineage- bone marrow cells, and then gated through ckit+, Sca1+ bone marrow cells and analyzed for CD34 and Flk2 expression. Donor- and host-derived cells from the bone marrow, spleen and thymus are also shown analyzed for ZsGreen to assess dox-independent expression of viral transgenes (right panels).

.

Figure S3. Identification of factors associated with myeloid colony forming potential *in vitro*, Related to Figure 4.

Summary of PCR-based screening data showing presence (yellow) or absence (black) of each of the indicated factors in the indicated colony type arising in methylcellulose from experiments in which Pro/Pre B cells or CMPs were

transduced the 36-factor cocktail. Data was generated by the PCR-based strategy outlined in Figure 1F, using primers indicated in Table S2. For Pro/Pre B cell experiments granulocyte, macrophage (GM), erythroid (BFU-E), and granulocyte, erythroid, macrophage, megakaryocyte (GEMM) colony types were scored whereas only GEMM colonies were scored in CMP experiments.

Figure S4. Host-derived contribution to the myeloid progenitor and primitive LSK compartment, and evidence that reprogrammed cells possess robust platelet and erythrocyte potential, Related to Figure 5.

(A-B) Bone marrow analysis showing host-derived contribution (CD45.2) to (A) myeloid progenitors, and (B) primitive multi-potent and hematopoietic stem cell compartments, of mice analyzed in Figure 5 that were transplanted with 8-TF, or 8-TF^{Poly} transduced Pro/Pre B cells (CD45.1) analyzed 18-20 weeks post-transplantation.
(C) Kaplan-Meier curve showing percent survival over a time course of 35 days of lethally irradiated mice transplanted with either 6.3 x 10^4 sorted donor-derived (CD45.1) MEPs (black line). Untransplanted (control) mice are also shown (grey line).
(D) Complete blood cell counts were performed in all surviving mice on day 20 post-irradiation and transplantation, and red blood cell (RBC), and platelet counts are shown for the mice described in (C). n=3 mice per condition.

Figure S5. Reprogrammed cells possess multi-lineage differentiation and self-renewal potential at the clonal level, Related to Figure 5.

(A) Sequencing analysis showing unique V(D)J junctions of the *Ig* heavy chain common to each of donor-derived B220$^+$ (B-Cells), CD3$^+$ (T-Cells), Mac1$^+$Gr1$^-$ (Myeloid) and Mac1$^+$Gr1$^+$ (Gran) cells. The cells were FACS sorted from primary recipients transplanted with 8-TF, or 8-TF^{Poly} transduced Pro/Pre B cells as indicated, and analyzed 20 weeks post-transplantation. Donor mouse ID is indicated. The identity of V_H, D_H and J_H genes was determined by IgBLAST.
(B) Sequencing analysis showing V(D)J junction of the *Ig* heavy chain common to donor-derived B220$^+$ (B-Cells), CD3$^+$ (T-Cells), and Mac1$^+$ (Myeloid) cells sorted from a secondary transplant recipient transplanted with whole bone marrow cells derived from Donor 1 shown in (A). The identity of V_H, D_H and J_H genes was determined by IgBLAST.

Figure S6. Transient expression of defined transcription factors in myeloid effector cells is sufficient instill them with long-term transplantation potential *in vivo,* and reprogrammed cells contribute to secondary hematopoietic organs, and reconstitute bone marrow progenitor compartments, Related to Figure 6.

(A) Flow cytometry showing percent donor (CD45.2) reconstitution (grey boxes) of recipient mice transplanted with 5x10^6 ZsGr, 6-TF^{Poly}, 8-TF or 8-TF^{Poly} transduced Mac1+cKit- myeloid effector cells 18-weeks post-transplantation.

(B) Donor reconstitution of peripheral blood (PB), bone marrow (BM), spleen, and thymus of mice transplanted with 6-TF^{Poly}, 8-TF, or 8-TF^{Poly} transduced Pro/Pre B cells 18-20 weeks post-transplantation.
(C) Bone marrow analysis showing donor reconstitution of the indicated populations of the mice described in (A).

Figure S7. Single cell gene expression analysis of iHSC and control cells, Related to Figure 7.

Heat map representation of single cell gene expression data in which 151 genes were analyzed in 271 individual cells. Control cell types analyzed included: HSCs (47 single cells), MPPs (48 cells), MEPs (18 cells), CMPs (21 cells), GMPs (23 cells), CLPs (26 cells), Pro/Pre B cells (6 cells), and host-derived HSCs (24 cells). Experimental cell analyzed included: iHSC 8-TF (23 cells), and iHSC 8-TF^{Poly} (35 cells). Expression levels are shown with high expression in red, and low expression in blue.

Supplemental Tables:

Table S1: Normalized microarray data of the 36 factors identified in this study showing Probeset ID, Gene Symbol, and expression values for each of the indicated populations with Gene Expression Omnibus identifying number indicated, Related to Figure 1.

Table S2: Primers used to identify integrated factors. PCR amplicon size is indicated, Related to Figure 4.

Table S3: Single cell expression data showing indicated population and log2 expression data for each of the indicted genes, Related to Figure 7.

Supplemental Procedures

Mice

B6.SJL-Ptprc^{a}/BoyAiTac1 (Taconic Farms; Hudson, NY) and C57BL/6N (Charles River Laboratories; Cambridge, MA) recipient mice and B6.CgGt(ROSA)26Sor$^{tm1(rtTA*M2)Jae}$/J donor mice on CD45.2, or CD45.1 background (Jackson, Bar Harbor, ME) were used. All mice were maintained according to protocols approved by Harvard Medical School Animal Facility and all procedures were performed with consent from the local ethics committees.

Pro/pre B-cell, CMP and HSC purification

Pro/Pre B-cells, CMPs and HSCs were purified as previously described (Akashi et al., 2000; Beerman et al., 2013; Derudder et al., 2009). Briefly, to isolate Pro/Pre B cells, a B220 enrichment was performed using a biotin-conjugated B220 antibody (BD Bioscience), followed by enrichment via streptavidin magnetic beads and column purification (Milteny). Cells were then stained with antibodies against, AA4.1, CD19, IgM, B220, CD25 and c-kit and purified by FACS. To isolate CMPs and HSCs, c-kit enrichment was performed on bone marrow cells

using directly conjugated magnetic beads (BD Bioscience) followed by column purification. Cells were then stained with antibodies against c-kit, Sca1, CD34, FcgR, and lineage (CD3, CD4, CD8, GR1, B220, Ter119, Mac1 and IL-7Ra) for CMPS, with inclusion of CD150, and Flk2 for HSCs. Antibodies to discriminate donor and recipient (CD45.1, CD45.2) were also used when purifying HSCs from transplant recipients. Antibodies used included: CD34, Sca1, c-kit, AA4.1 from eBioscience (San Diego, CA); FcgR from BD Bioscience (San Jose, CA); IgM Sigma Aldrich (St. Louis, MO); IL-7Ra, Ter119, CD45.1, CD45.2, Mac1, CD3, CD4, CD8, Gr1, CD150, CD19, CD25 and B220 from BioLegend (San Diego, CA). All cells were sorted on a FACS Aria II (Becton Dickinson) directly into sample media containing 2% FBS.

Virus Production and Transduction

Open reading frames (ORFs) from a cDNA library generated from purified HSCs were cloned into pHAGE2 lentivirus (Mostoslavsky et al., 2005). Lentiviruses (LV) were produced in 293T cells, concentrated, and titered in Jurkat cells. Transduction was carried out in 96 U-bottom plates. Doxycycline (Sigma-Aldrich) was added to 1ug/ml for factor induction.

High throughput single cell qPCR and computational analysis

Individual primer sets were pooled to a final concentration of 0.1µM for each primer. Individual cells were sorted directly into 96 well plates loaded with 5µL RT-PCR master mix (2.5µL CellsDirect reaction mix, Invitrogen; 0.5µL primer pool; 0.1µL RT/Taq enzyme, Invitrogen; 1.9µL nuclease free water) per well. Sorted plates were immediately frozen on dry ice. After brief centrifugation at 4°C, the plates were immediately placed on PCR machine. Cell lyses and sequence-specific reverse transcription were performed at 50°C for 60 minutes. Reverse transcriptase inactivation and Taq polymerase activation was achieved by heating to 95°C for 3 min. Subsequently, in the same tube, cDNA went through 20 cycles of sequence-specific amplification by denaturing at 95°C for 15 sec, annealing and elongation at 60°C for 15 min. After preamplification, plates were stored at -80°C. Pre-amplified products were diluted 5-fold prior to analysis. Amplified single cell samples were analyzed with Universal PCR Master Mix (Applied Biosystems), EvaGreen Binding Dye (Biotium) and individual qPCR primers using 96.96 Dynamic Arrays on a BioMark System (Fluidigm). Ct values were calculated using the BioMark Real-Time PCR Analysis software (Fluidigm). Gene expression levels were estimated by subtracting the background level of 28 by the Ct level, which approximates Log2 gene expression levels. Principal component analysis (PCA) was performed in Matlab. Unsupervised hierarchical clustering was done using R using the average linkage method and a correlation-based distance. Representative control cells were selected as those whose expression levels were closest to the median within each control cell-type based on Euclidean distance. Eight HSCs, eight Host HSCs, six Pro/Pre B-cells, and four from each of the remaining control cell types were selected. The dendrogram branches were color-coded by cell type, as in the PCA analysis. Violin plots and heatmaps were generated with Matlab. The master heatmap of

all the raw data was generated with MultiExperiment Viewer (MeV) (http://www.tm4.org/mev.html) using the default settings.

V(D)J Analysis

Ig Heavy chain recombination was determined using genomic DNA isolated from sorted host- or donor-derived cells using a semi-nested PCR strategy based on (Diss et al., 1994; Ramasamy et al., 1992) to amplify the framework regions of VH to specific sites of JH. Primers used FR: ACACGGCCGTGTATTACTG, JH1: TGAGGAGACGGTGAC and JH2: GTGACCAGGGTNCCTTGGCCCCAG. First round amplification of 25 cycles was performed with FR/JH1 (70°C annealing/ 20s extension). Second round amplification of 35 cycles with FR/JH2 (65°C annealing/ 30s extension). Light chain (kappa and lambda) recombination was tested as previously described (Cobaleda et al., 2007). For V(D)J junction sequencing, amplified products were run on a gel and amplicons of a similar size observed in all tested lineages were extracted, cloned, sequenced, and analyzed using IgBlast (Ye et al., 2013).

Supplemental References

Akashi, K., Traver, D., Miyamoto, T., and Weissman, I.L. (2000). A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. Nature *404*, 193-197.

Beerman, I., Bock, C., Garrison, B.S., Smith, Z.D., Gu, H., Meissner, A., and Rossi, D.J. (2013). Proliferation-dependent alterations of the DNA methylation landscape underlie hematopoietic stem cell aging. Cell Stem Cell *12*, 413-425.

Cobaleda, C., Jochum, W., and Busslinger, M. (2007). Conversion of mature B cells into T cells by dedifferentiation to uncommitted progenitors. Nature *449*, 473-477.

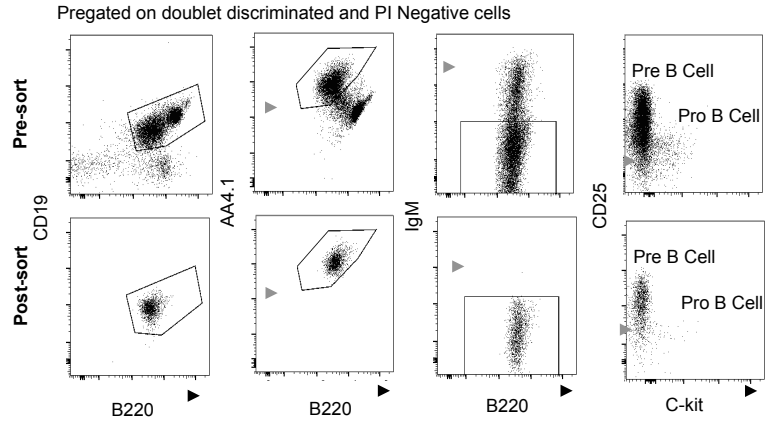
Derudder, E., Cadera, E.J., Vahl, J.C., Wang, J., Fox, C.J., Zha, S., van Loo, G., Pasparakis, M., Schlissel, M.S., Schmidt-Supprian, M.*, et al.* (2009). Development of immunoglobulin lambda-chain-positive B cells, but not editing of immunoglobulin kappa-chain, depends on NF-kappaB signals. Nat Immunol *10*, 647-654.

Diss, T.C., Pan, L., Peng, H., Wotherspoon, A.C., and Isaacson, P.G. (1994). Sources of DNA for detecting B cell monoclonality using PCR. J Clin Pathol *47*, 493-496.

Ramasamy, I., Brisco, M., and Morley, A. (1992). Improved PCR method for detecting monoclonal immunoglobulin heavy chain rearrangement in B cell neoplasms. J Clin Pathol *45*, 770-775.

Ye, J., Ma, N., Madden, T.L., and Ostell, J.M. (2013). IgBLAST: an immunoglobulin variable domain sequence analysis tool. Nucleic Acids Res *41*, W34-40.

Figure S1


Pregated on doublet discriminated and PI Negative cells
Pre-sort
Post-sort
CD19
B220
AA4.1
B220
IgM
B220
CD25
C-kit
Pre B Cell
Pro B Cell
Pre B Cell
Pro B Cell

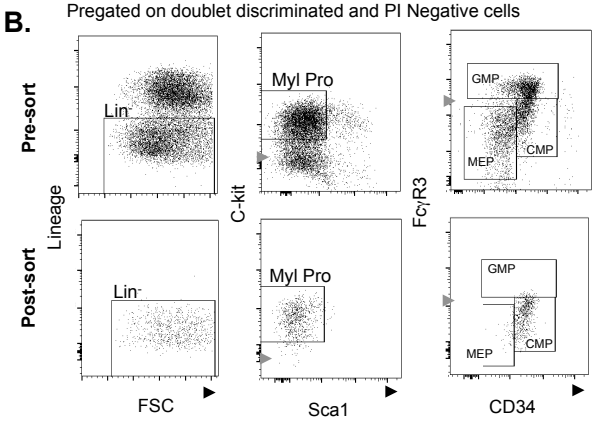

B.
Pregated on doublet discriminated and PI Negative cells
Pre-sort
Post-sort
Lineage
Lin-
FSC
C-kit
Myl Pro
Sca1
FcγR3
GMP
CMP
MEP
CD34

Figure S2

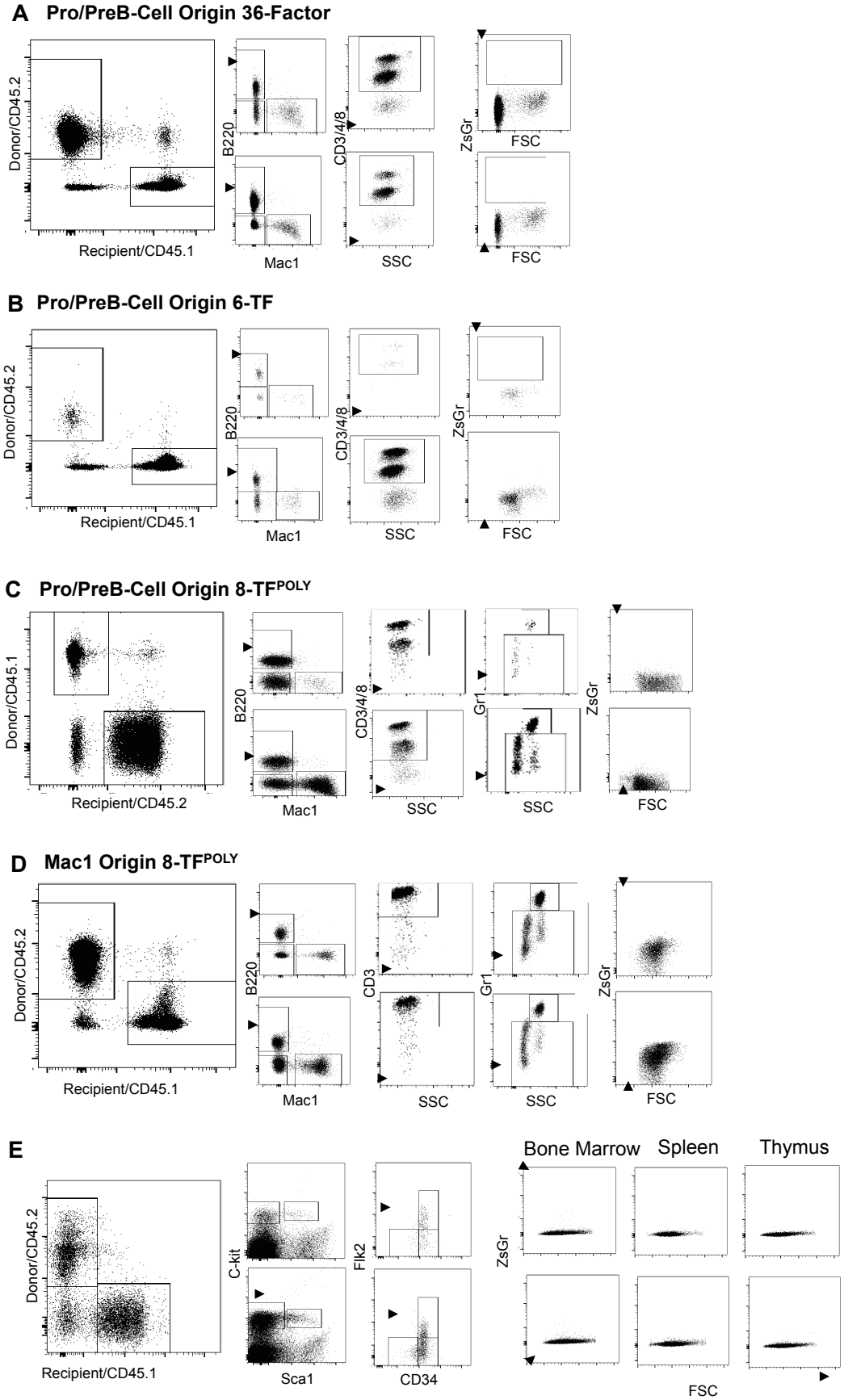

A Pro/PreB-Cell Origin 36-Factor
Donor/CD45.2
Recipient/CD45.1
B220
Mac1
CD3/4/8
SSC
ZsGr
FSC
B Pro/PreB-Cell Origin 6-TF
Donor/CD45.2
Recipient/CD45.1
B220
Mac1
CD3/4/8
SSC
ZsGr
FSC
C Pro/PreB-Cell Origin 8-TFPOLY
Donor/CD45.1
Recipient/CD45.2
B220
Mac1
CD3/4/8
SSC
Gr1
SSC
ZsGr
FSC
D Mac1 Origin 8-TFPOLY
Donor/CD45.2
Recipient/CD45.1
B220
Mac1
CD3
SSC
Gr1
SSC
ZsGr
FSC
E
Donor/CD45.2
Recipient/CD45.1
C-kit
Sca1
Flk2
CD34
Bone Marrow
Spleen
Thymus
ZsGr
FSC

Figure S3

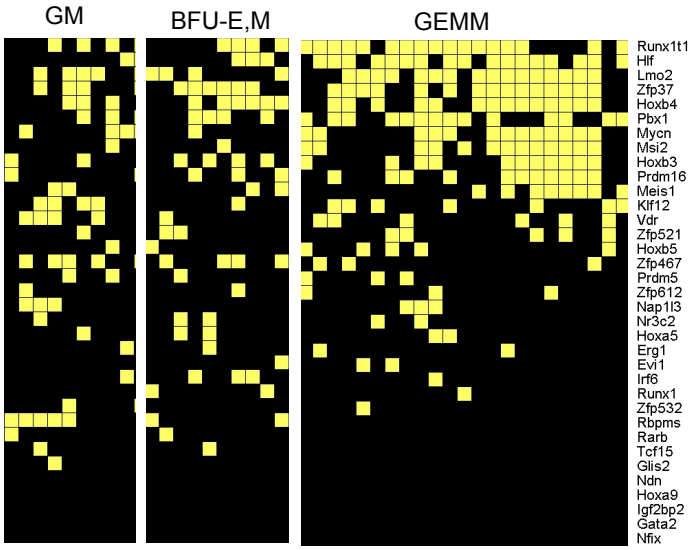

GM
BFU-E,M
GEMM
Runx1t1
Hlf
Lmo2
Zfp37
Hoxb4
Pbx1
Mycn
Msi2
Hoxb3
Prdm16
Meis1
Klf12
Vdr
Zfp521
Hoxb5
Zfp467
Prdm5
Zfp612
Nap1l3
Nr3c2
Hoxa5
Erg1
Evi1
Irf6
Runx1
Zfp532
Rbpms
Rarb
Tcf15
Glis2
Ndn
Hoxa9
Igf2bp2
Gata2
Nfix

Figure S4

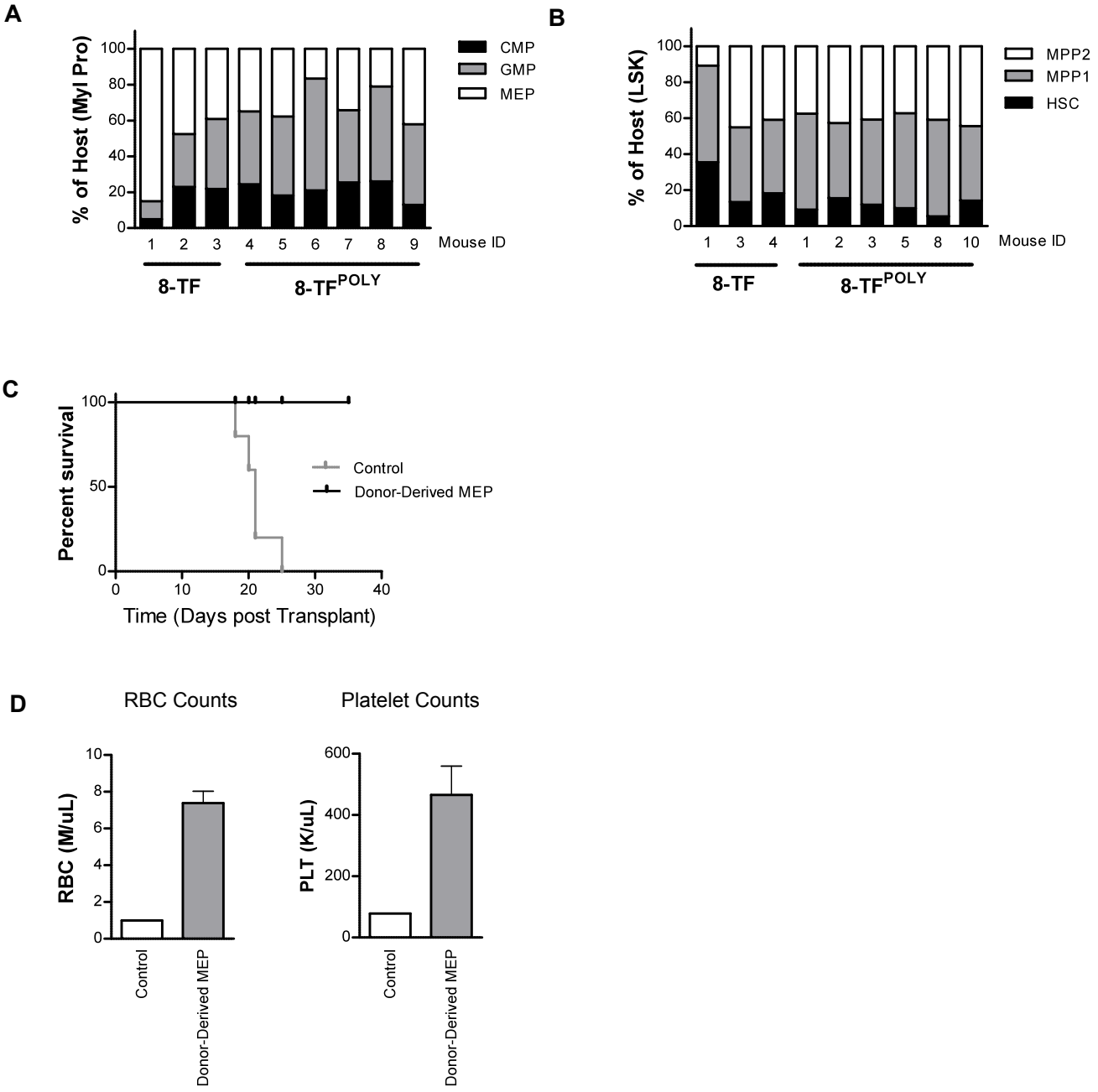

A
% of Host (Myl Pro)
100
80
60
40
20
0
CMP
GMP
MEP
1 2 3 4 5 6 7 8 9
Mouse ID
8-TF
8-TF^POLY
B
% of Host (LSK)
100
80
60
40
20
0
MPP2
MPP1
HSC
1 3 4 1 2 3 5 8 10
Mouse ID
8-TF
8-TF^POLY
C
Percent survival
100
50
0
Control
Donor-Derived MEP
0 10 20 30 40
Time (Days post Transplant)
D
RBC Counts
RBC (M/uL)
10
8
6
4
2
0
Control
Donor-Derived MEP
Platelet Counts
PLT (K/uL)
600
400
200
0
Control
Donor-Derived MEP

Figure S5

A Primary transplantation

Viral cocktail Donor ID	VDJ ID	Cell Type	V_H	D_H	J_H
8-TF Donor 4	V1-54/D5/J2-1	B-Cell	atggttagggcaag	g**caacactc**tcac	agtctcctcaggtga
		T-Cell	ctgatggttagggcaag	g**cacaactc**tcac	agtctcctcaggtga
		Myeloid	ctgatggttagggcaag	g**cacaactc**tcac	agtctcctcaggtga
		Gran	ctgatggttagggcaag	g**cacaactct**cac	agtctcctcaggtga
8-TF^{POLY} Donor 1	V5-17/D2-3/J2-1	B-Cell	acacggccgtgtattactgtgcaagg	cc**tgatggtta**ggg	caaggcaccactct
		T-Cell	cacggccgtgtattactgtgcaagg	cc**tgatggtta**ggg	caaggcaccactct
		Myeloid	acacggccgtgtattactgtgcaagg	cc**tgatggtta**ggg	caaggcaccactct
		Gran	cacggccgtgtattactgtgcaagg	cc**tgatggtta**ggg	caaggcaccactct
8-TF^{POLY} Donor 8	V5-9/D1-1/J2-1	B-Cell	tgtattactgtgcaaggga	ga**agctac**acc	actggtacttcgat
		T-Cell	tgtattactgtgcaaggga	ga**agctac**acc	actggtacttcgat
		Myeloid	tgtattactgtgcaaggga	ga**agctac**acc	actggtacttcgat
		Gran	tgtattactgtgcaaggga	ga**agctac**acc	actggtacttcgat

B Secondary transplantation

Viral cocktail Donor ID	VDJ ID	Cell Type	V_H	D_H	J_H
8-TF^{POLY} 1°- Donor 1	V5-17/D2-3/J2-1	B-Cell	acacggccgtgtattactgtgcaagg	cc**tgatggtta**ggg	caaggcaccactct
		T-Cell	acacggccgtgtattactgtgcaagg	cc**tgatggtta**ggg	caaggcaccactct
		Myeloid	acacggccgtgtattactgtgcaagg	cc**tgatggtta**ggg	caaggcaccactct

Supplemental Figure 6

Figure S6

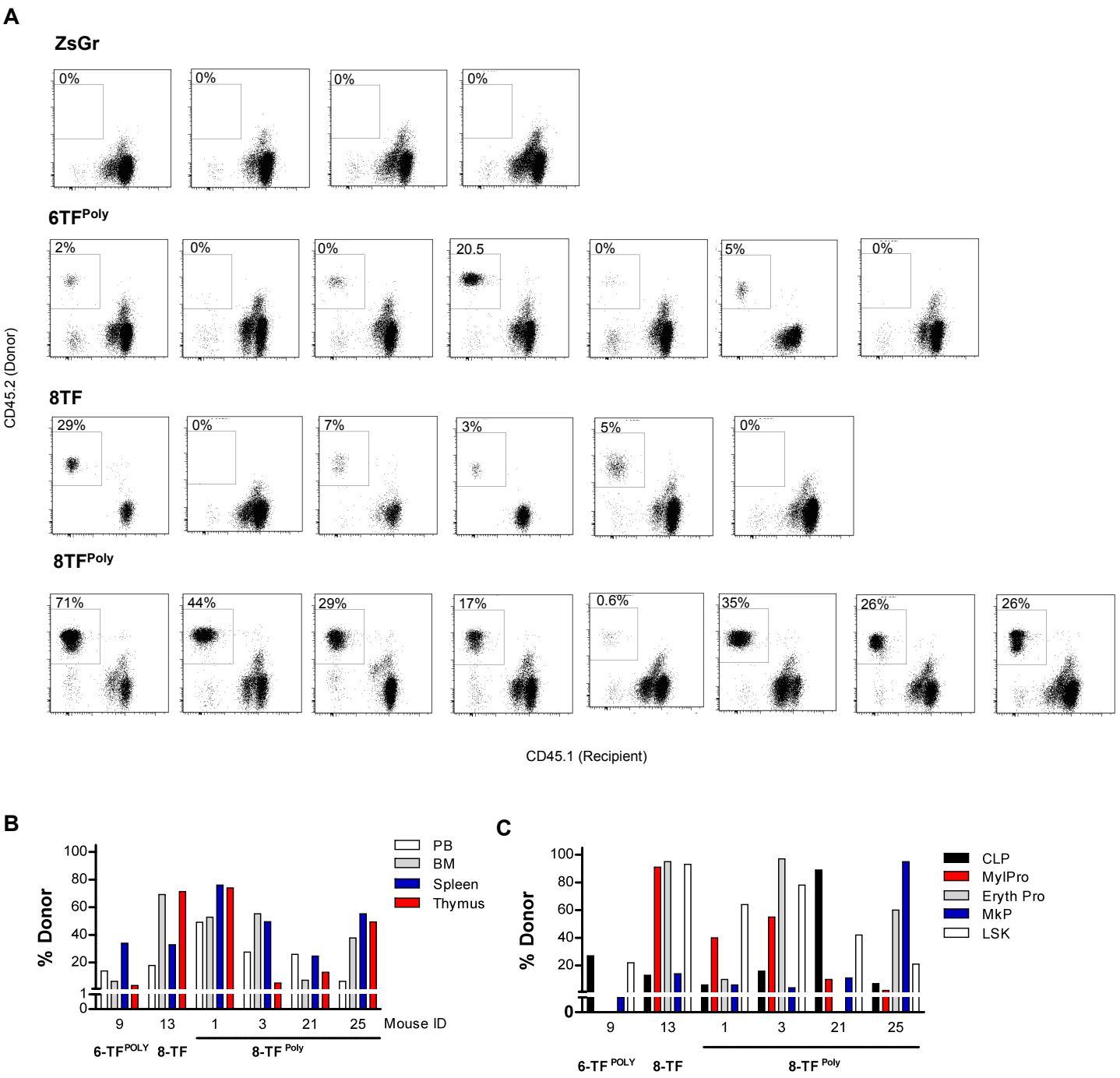


Figure S7