Cell Reports, Volume *21*

Supplemental Information

Critical Modulation of Hematopoietic Lineage Fate by Hepatic Leukemia Factor

Martin Wahlestedt, Vasileios Ladopoulos, Isabel Hidalgo, Manuel Sanchez Castillo, Rebecca Hannah, Petter Säwén, Haixia Wan, Monika Dudenhöffer-Pfeifer, Mattias Magnusson, Gudmundur L. Norddahl, Berthold Göttgens, and David Bryder

SUPPLEMENTARY INFORMATION

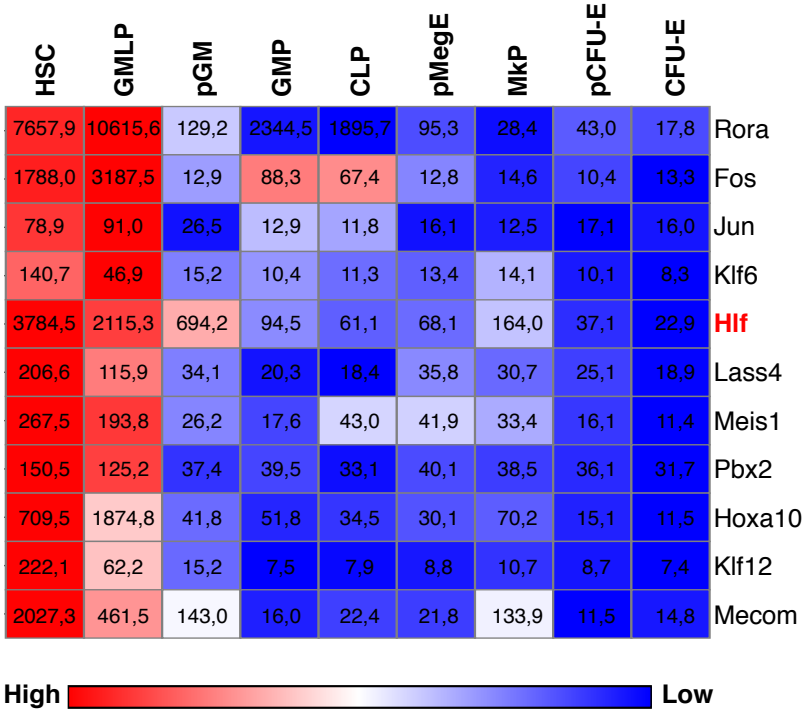


Figure S1

Figure S1. Extended heatmap of transcription factors with 2-fold or higher expression levels in HSCs and GMLPs compared to downstream progenitor subsets (related to Figure 1). The heatmap shown in Figure 1B is depicted along with the probe level expression values.

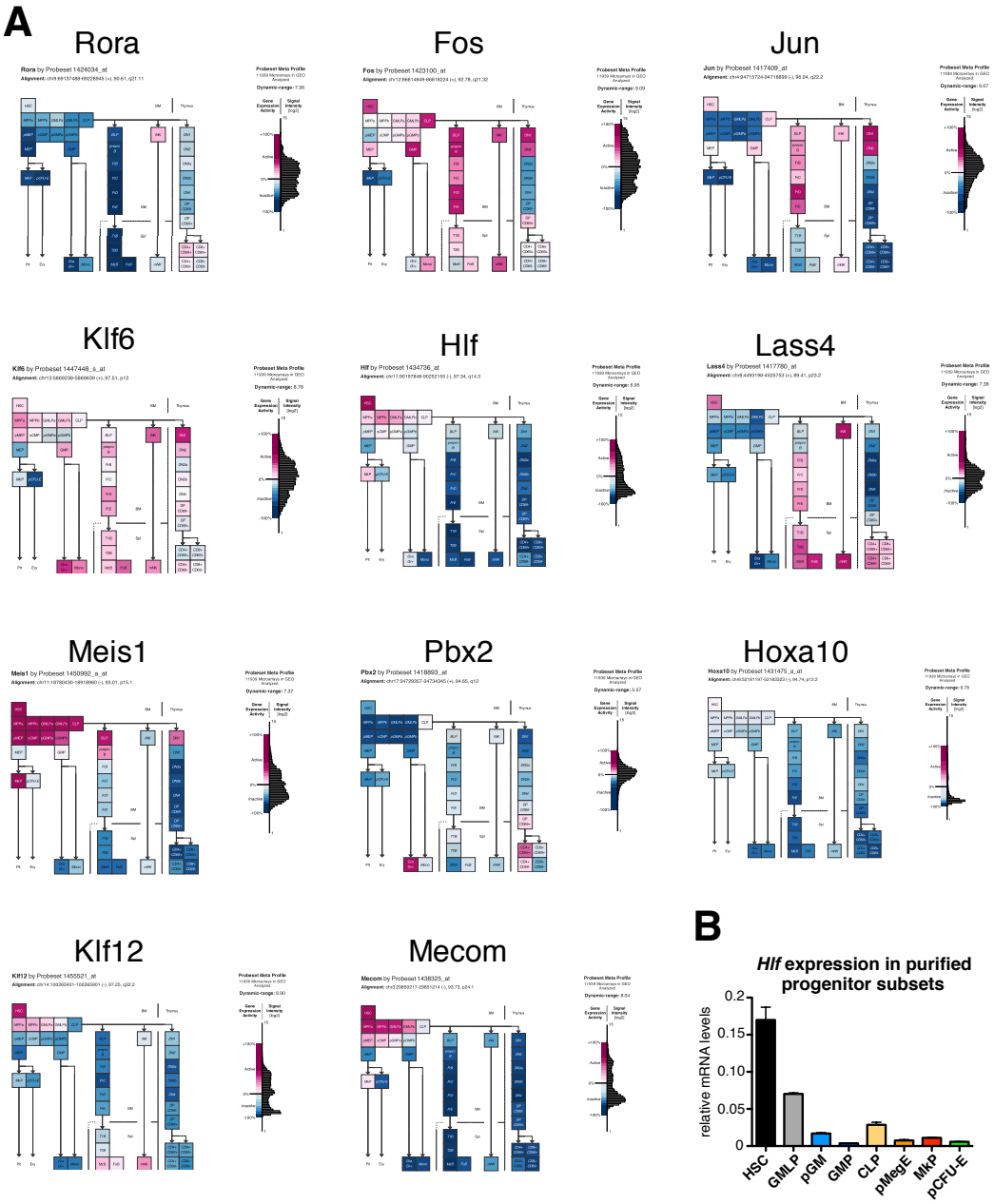


Figure S2. Expression profiles from Gene Expression Commons of the 11 identified HSC and GMLP specific transcription factors and qRT-PCR validation of Hlf expression levels in primary hematopoietic progenitors (related to Figure 1). (A) Hematopoietic gene expression profiles from Gene Expression commons (https://gexc.riken.jp) of identified transcription factors. (B) Hlf mRNA expression levels in the indicated stem and progenitor subsets as determined by qRT-PCR and expressed relative to Actb (3 replicates/cell type from one of two experiments with similar results). qRT-PCR; quantitative Real Time-Polymerase Chain Reaction.

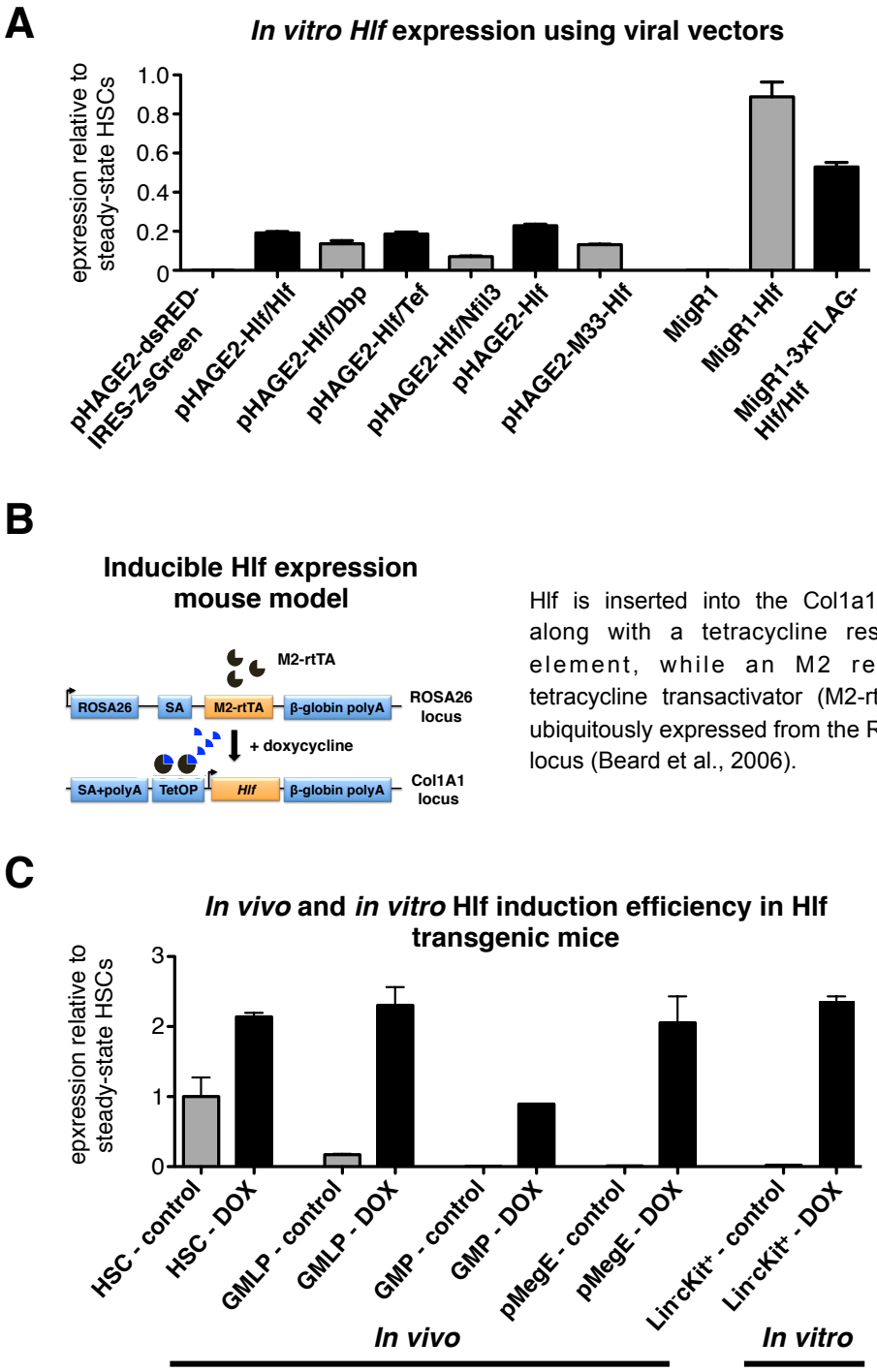


Figure S3. Validation of the Hlf-based viral constructs (*in vitro*) and schematic depicition and validation of the inducible Hlf mouse model (related to Figure 1). (A) Hlf expression levels in FACS sorted GFP+ transduced cells using the viral vectors in the work (3 replicates for each vector). (B) Schematic depiction of the Hlf conditional transgenic mouse model. A reverse tetracycline transactivator (M2-rtTA) is driven from the Rosa26 locus, which, upon the addition of DOX can bind to the TetOP and drive the expression of the single-copy Hlf transgene inserted into the Col1a1 locus. (C) Hlf expression levels in depicted cell subsets using the inducible Hlf transgenic mouse (3 replicates for each cell type and treatment). Error bars depict SEM. TetOP; Tetracycline Operon, SA; Splice Acceptor.

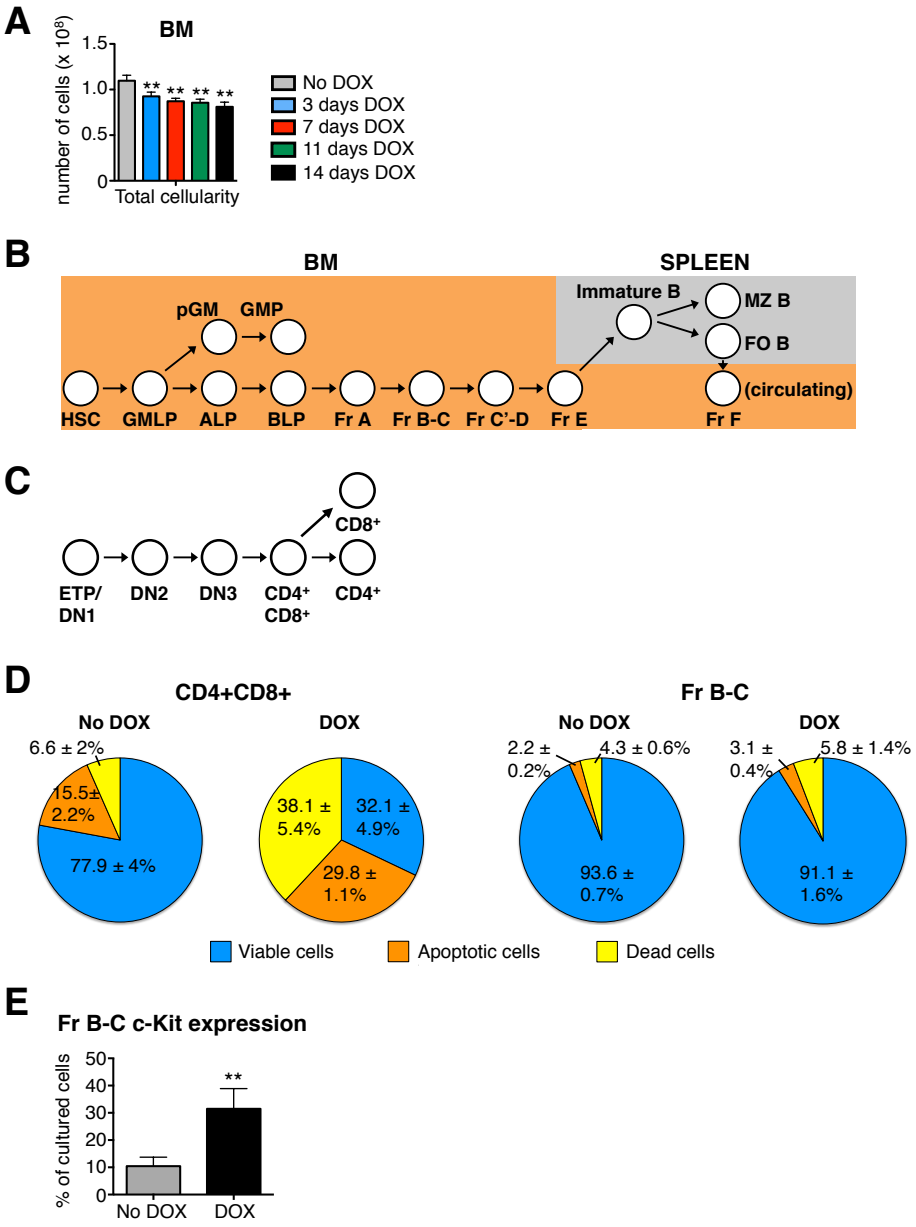
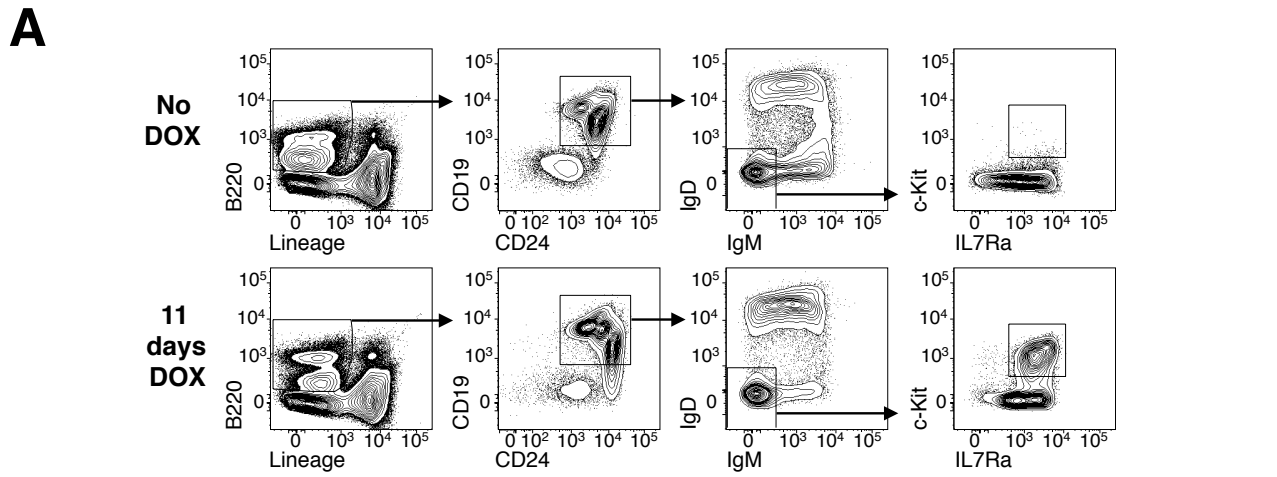
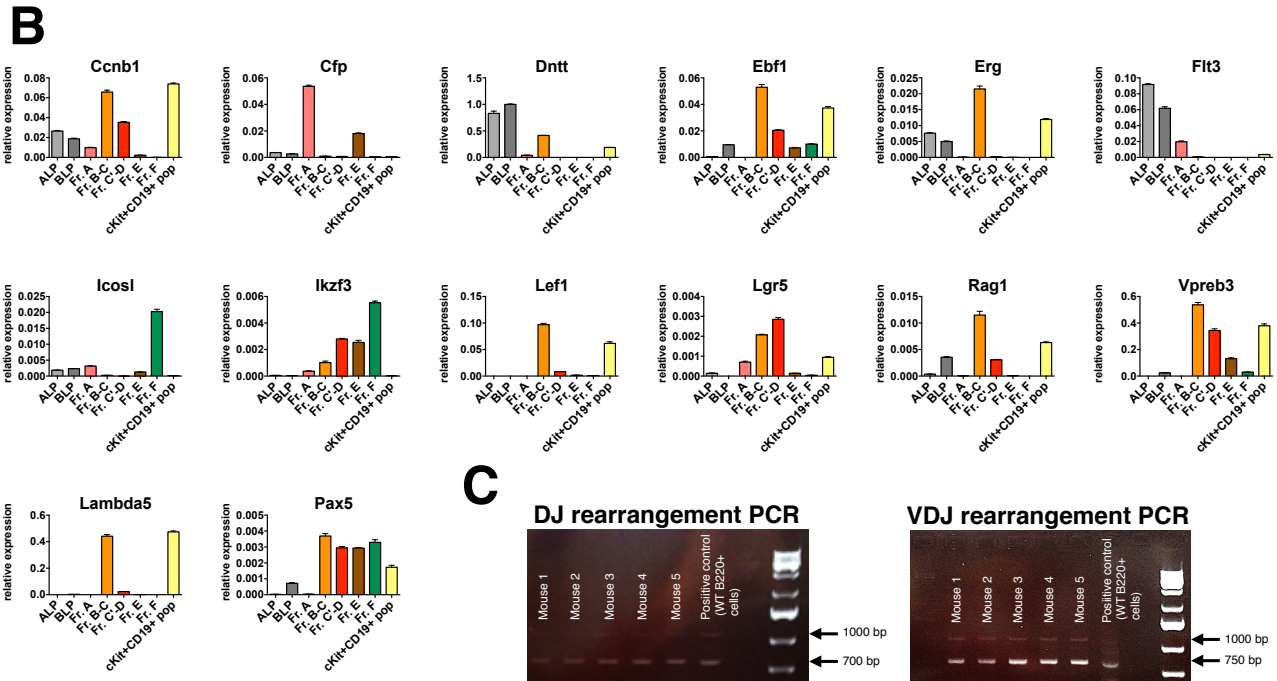


Figure S4. Hematopoietic effects of in vivo Hlf induction and specific investigation into the Hlf-associated loss of immature lymphocytes (related to Figure 2). (A) Bar charts showing the overall BM cellularity in Hlf induced mice. (B) Overview of the various analyzed cell fractions and their proposed developmental relationships. (C) Overview of the various analyzed cell fractions in the thymus and their proposed developmental relationships. (D) $CD4^+CD8^+$ thymocytes and Fraction B-C B-lymphocytes were cultured for 48h in the absence or presence of DOX on OP9-DL1 and OP9 stromal cell respectively before staining with Annexin V and Propidium Iodide. Pie charts show the degree of cell viability (average ± SEM), apoptosis and death for both cell types (n = 6 replicates per group, from one of two experiments with similar results). (E) Fraction B-C B-lymphocytes treated as in (D) were investigated for c-Kit expression using FACS (n = 6, one experiment). Bar graph shows the average percentage of $c\text{-}Kit^+$ cells in Fraction B-C cultures. Error bars denote SEM.


A
No DOX
11 days DOX
B220
Lineage
CD19
CD24
IgD
IgM
c-Kit
IL7Ra


B
Ccnb1
Cfp
Dntt
Ebf1
Erg
Flt3
Icosl
Ikzf3
Lef1
Lgr5
Rag1
Vpreb3
Lambda5
Pax5
relative expression
ALP BLP Fr. A Fr. B-C Fr. C'-D Fr. E Fr. F cKit+CD19+ pop

C
DJ rearrangement PCR
VDJ rearrangement PCR
Mouse 1
Mouse 2
Mouse 3
Mouse 4
Mouse 5
Positive control (WT B220+ cells)
1000 bp
700 bp
333 bp
750 bp
350 bp

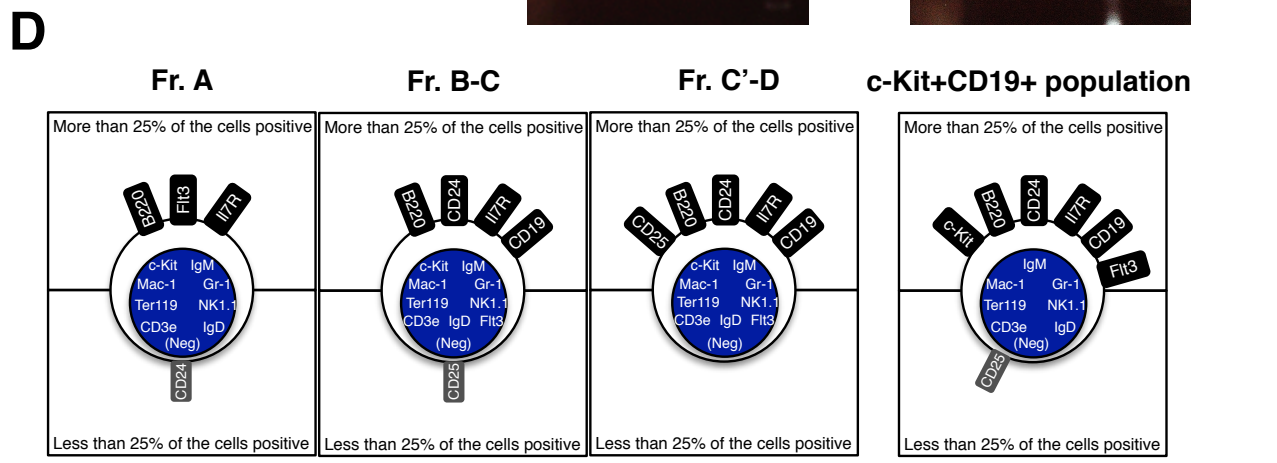
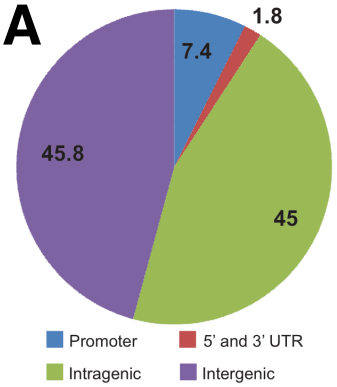

D
Fr. A
Fr. B-C
Fr. C'-D
c-Kit+CD19+ population
More than 25% of the cells positive
Less than 25% of the cells positive
B220
Flt3
Il7R
CD24
CD19
CD25
c-Kit
IgM
Mac-1
Gr-1
Ter119
NK1.1
CD3e
IgD
(Neg)

Figure S5 (related to Figure 2). Characterization of Hlf induced Lin$^-$ B220$^+$CD19$^+$CD24$^+$IgM$^-$IgD$^-$c-Kit$^+$IL7Ra$^+$ cells. (A) Representative FACS plots showing the appearance of B220$^+$CD19$^+$CD24^{+l}L7Ra$^+$c-Kit$^+$ cells in the BM following enforced Hlf expression. (B) The expression levels of Ccnb1, Cfp, Dntt, Ebf1, Erg, Flt3, Icosl, Ikzf3, Lambda5, Lef1, Lgr5, Pax5, Rag1 and Vpreb3 were determined in Lin$^-$ B220$^+$CD19$^+$CD24$^+$IgM$^-$IgD$^-$c-Kit$^+$IL7Ra$^+$ cells and compared to those in ALPs, BLPs, Fr. A, Fr. B-C, Fr. C'-D, Fr. E and Fr. F cells (n = 3 replicates per cell type and gene, from one experiment). (C) (TOP) Heavy chain DJ rearrangement and (BOTTOM) heavy chain VDJ rearrangement measured by PCR in Lin$^-$B220$^+$CD19$^+$CD24$^+$IgM$^-$IgD$^-$c-Kit$^+$IL7Ra$^+$ cells (from 5 individual mice) and compared to BM resident B220+ cells in WT mice. Arrows depict the expected bands. (D) Schematic diagrams showing the cell surface marker phenotype of Lin$^-$B220$^+$CD19$^+$CD24$^+$IgM$^-$IgD$^-$c-Kit$^+$IL7Ra$^+$ cells compared to Fr. A, Fr. B-C and Fr. C'-D. Error bars denote SEM.

A

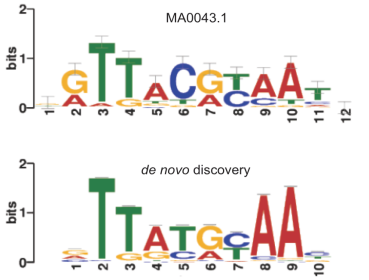


B

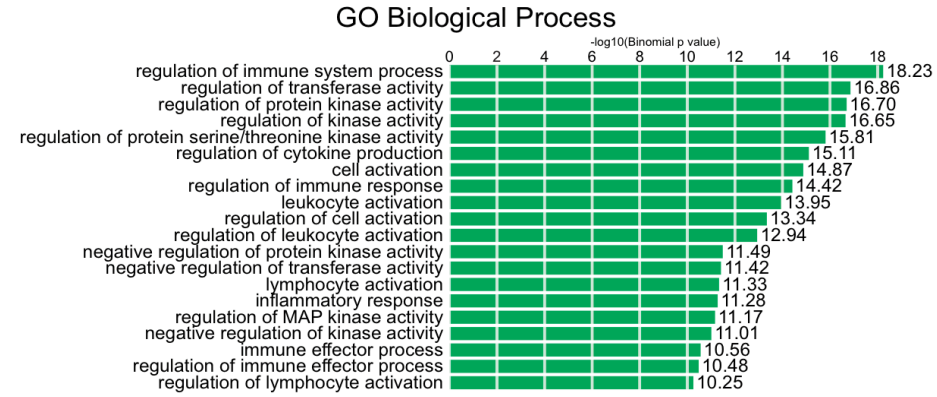
Summary

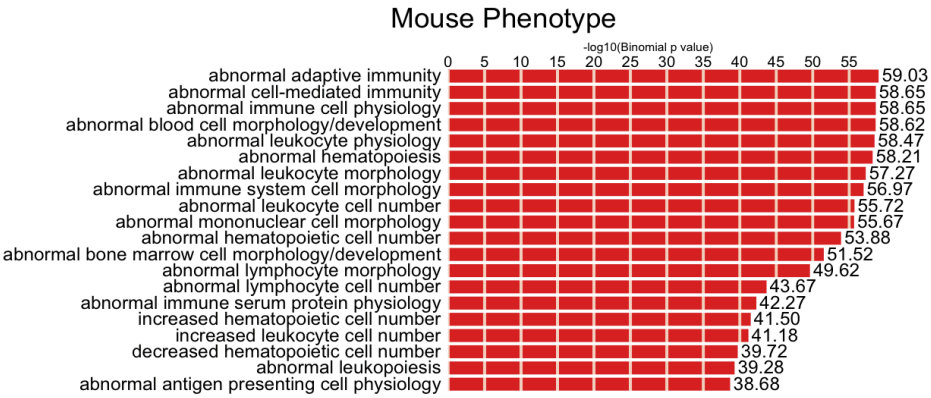
Name	MA0043.1
Alt. Name	HLF
Database	JASPAR non redun _vertebrate_MEME.txt
***p*-value**	2.00984e-05
***E*-value**	0.00261279
***q*-value**	0.00254512
Overlap	10
Offset	1
Orientation	Normal

Alignment



C





D

Rank	Motif	P-value	log P-pvalue	% of Targets	% of Background	STD(Bg STD)	Best Match/Details
1		1e-1960	-4.515e+03	57.37%	2.57%	30.8bp (63.4bp)	MF0006.1_bZIP_cEBP-like_subclass/Jaspar More Information | Similar Motifs Found
2		1e-369	-8.511e+02	26.13%	4.53%	51.0bp (57.4bp)	PB0058.1_Sfpi1_1/Jaspar More Information | Similar Motifs Found
3		1e-130	-2.997e+02	22.87%	8.54%	52.9bp (61.1bp)	RUNX(Runt)/HPC7-Runx1-ChIP-Seq/Homer More Information | Similar Motifs Found
4		1e-122	-2.821e+02	3.39%	0.10%	26.0bp (59.4bp)	MA0138.2_REST/Jaspar More Information | Similar Motifs Found
5		1e-88	-2.030e+02	2.79%	0.11%	31.8bp (61.6bp)	MA0138.1_REST/Jaspar More Information | Similar Motifs Found
6		1e-54	-1.257e+02	10.31%	3.87%	55.4bp (60.1bp)	PU.1-IRF(ETS:IRF)/Bcell-PU.1-ChIP-Seq(GSE21512)/Homer More Information | Similar Motifs Found
7		1e-37	-8.688e+01	6.76%	2.44%	53.3bp (60.6bp)	Jun-AP1(bZIP)/K562-cJun-ChIP-Seq/Homer More Information | Similar Motifs Found
8		1e-23	-5.472e+01	2.76%	0.74%	60.6bp (57.5bp)	CTCF(Zf)/CD4+-CTCF-ChIP-Seq/Homer More Information | Similar Motifs Found

Figure S6 (related to Figure 5). Genome-wide Hlf binding characteristics. (A) Pie chart displaying the distribution of the identified Hlf peaks in different genomic regions (graph obtained using CEAS (Shin et al., 2009), available on cistrome.org/ap/). (B) Sequence logo showing the deduced Hlf binding motif (HOMER de novo discovery) and the binding motif for human HLF in the JASPAR database (MA0043.1). (C) Gene ontology categories and mouse phenotypes associated with the putative Hlf targets determined using GREAT analysis (http://www.great.stanford.edu) (McLean et al., 2010) ranked in a decreasing likelihood order. (D) Sequence logos showing the 8 most overrepresented TF motifs in the Hlf peaks as obtained from HOMER *de novo* motif discovery.

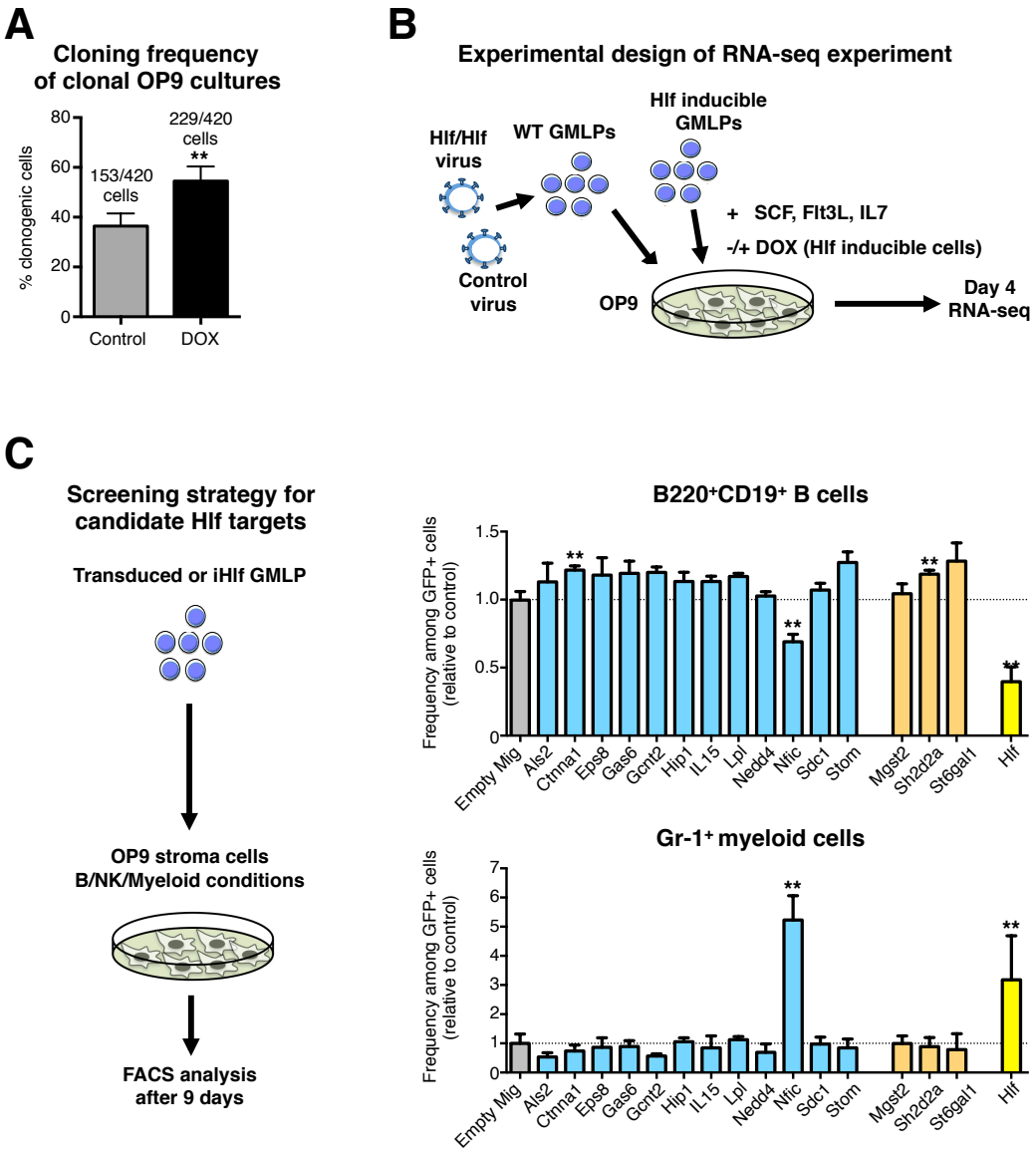


Figure S7 (related to Figure 6 and 7). Cloning frequency of Hlf-inducible GMLPs, the experimental outline for RNA-seq experiments and functional screening of candidate Hlf targets. (A) Cloning frequency (defined as percent of the total number of seeded wells containing at least 30 B-, NK or myeloid cells) of cultures initiated with GMLPs in the absence or presence of Hlf induction (420 initiated cultures/group from 5 independent experiments). (B) Schematic outline for the RNA sequencing experiments; WT GMLPs transduced with the Hlf/Hlf construct or a control virus (see Figure 1) and Hlf inducible GMLPs maintained in the absence or presence of DOX were cultured for 4 days on OP9 stroma under B cell permissive conditions, prior to processing for RNA sequencing (duplicate samples per group). (C) Bicistronic retroviral vectors for indicated genes and GFP were generated and used to transduce GMLPs. Cells were subsequently evaluated using the OP9 stromal co-culture system for their ability to produce myeloid and lymphoid cells following 9 days of culture (n = 3 replicates per vector, from one experiment). Error bars depict SEM.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Construction of viral vectors and production of high-titer viruses

All viral constructs unless indicated were constructed using PCR-mediated Gibson Assembly cloning (New England Biolabs). In brief, the single-copy Hlf gene construct was generated by PCR amplification of the coding sequence of Hlf with the addition of a N-terminal HA-tag and assembly into a BamHI and NotI (New England Biolabs) linearized pHAGE2 vector (Mostoslavsky et al., 2005). The single-copy Nfil3 construct was generated by PCR amplification of the Nfil3 coding sequence with 5' addition of a NotI restriction enzyme site and 3' addition of a BamHI restriction enzyme site, A-tailing using Taq DNA polymerase (VWR), TA-cloning into the pCR2.1 plasmid (Invitrogen), restriction enzyme digestion of the Nfil3 fragment by BamHI and NotI (New England Biolabs) and ligation into a BamHI and NotI linearized pHAGE2 vector using T4 DNA ligase (New England Biolabs). For the Hlf based forced dimer constructs, the coding sequence of Hlf was amplified without a stop codon with an incorporated linker sequence (Neuhold and Wold, 1993) and assembled into pHAGE2 along with the coding sequences of Hlf, Tef, Dbp and Nfil3. To generate the vector used for the ChIP-seq experiments, the coding sequence of an Hlf/Hlf homodimer was first PCR amplified with the addition of an N-terminal 3xFLAG tag (GACTACAAAGACCATGACGGTGATTATAAAGATCATGATATCGATTACAAGGATGA CGATGACAAG) and assembled into BamHI and NotI linearized pHAGE2. Next, the 3xFLAG-Hlf/Hlf fragment along with an IRES element was PCR amplified and assembled into a NcoI linearized pMX-GFP retroviral vector (Cell Biolabs). The M33-Hlf lentiviral fusion construct was generated by fusing the repressive domain of M33 to the coding sequence of Hlf as described for Meis1 (Argiropoulos et al., 2010) with the addition of a N-terminal 3xFLAG tag, followed by cloning into a BamHI and NotI linearized pHAGE2 vector. The retroviral M33-Hlf fusion construct was generated by PCR amplification of 3xFLAG-M33-Hlf from the pHAGE2 vector and assembly into EcoRI and XhoI linearized MigR1 retroviral vector (Pear et al., 1998). The lentiviral M33-Hlf fusion construct was generated in an identical manner but assembled into the pHAGE2 vector. The constructs used for transdifferentiation experiments were generated by PCR amplification of single-copy Hlf and Cebpb and assembled into a EcoRI and XhoI linearized MigR1 (Pear et al., 1998). To generate the retroviral vectors containing the selected candidate Hlf targets, the coding sequences of each gene was PCR amplified and assembled into a EcoRI and XhoI linearized MigR1 (Pear et al., 1998). All Hlf based constructs were designed to retain a minimal Kozak sequence (ATCACG). The gene products for the Hlf, M33, Dbp, Nfil3 and Tef constructs were PCR amplified using the Q5 High-Fidelity DNA Polymerase (New England Biolabs) from a full-length cDNA library generated from unfractionated BM cells using the Qiagen RNeasy Micro Kit (Qiagen) and Superscript III (Invitrogen). The gene products for Als2, Ctnna1, Eps8, Gas6, Gcnt2, Hip1, Il15, Lpl, Mgst2, Nedd4, Nfic, Sdc1, Sh2d2a, St6gal1 and Stom were PCR amplified using the Q5 High-Fidelity DNA Polymerase (New England Biolabs) from full-length cDNA libraries generated from unfractionated BM cells, with the exception of Gas6 that was generated from testis cDNA, using the RNeasy Micro mRNA purification kit, with first- and second-strand synthesis carried out using Superscript II (Invitrogen). The resulting cDNA was next amplified using KAPA HiFi Hotstart Readymix (Kapa Biosystems Inc.). Lentiviruses were produced by transient transfection of 293T producer cells using Lipofectamine LTX (Invitrogen) with the lentiviral plasmids and

packaging constructs (HDM-Hgpm2, HDM-Tat1b, HDM-VSVG and RC-CMV-Rev1b). Lentiviral supernatants were harvested 48 and 72h post transfection and concentrated by ultracentrifugation. The retroviral supernatants were produced by Lipofectamine LTX transfection of Plat-E packaging cells (Cell Biolabs) with the retroviral vectors, followed by harvest of the supernatants 48h post transfection.

Generation of transgenic mice and in vivo Hlf transgenic strategy

To generate DOX inducible Hlf transgenic mice, the conserved coding sequence, along with the 5' and 3' UTR of the murine Hlf gene was subcloned from the pYX-Asc-Hlf plasmid (K.K. DNAFORM) into the pBS31 targeting vector (Beard et al., 2006). Next, the targeting construct was inserted in the Col1a1 locus via Flpase mediated recombination into the KH2 ES cell line (Beard et al., 2006). See also Figure S3B for a schematic depiction of the model. Specific integration of the targeting construct and the presence of the reverse tetracycline transactivator in the Rosa26 locus was verified by PCR. Following karyotyping and test of Hlf inducibility in vitro, engineered ES cells were injected into E3.5 C57Bl/6 blastocysts and implanted into pseudopregnant mice at the Transgenic Core Facility at Lund University. Following germline transmission, mice were backcrossed to CD45.1$^+$ C57BL/6 mice and bred to homozygozity for both the modified Col1a1 and Rosa26 loci. To achieve transgenic expression of Hlf in vivo, transgenic mice or mice transplanted with Hlf inducible GMLPs were administered DOX-containing chow (or normal food in control mice) ad libitum (2 g/kg, Ssniff Specialdiäten). All animal experiments were performed with consent from a local ethical committee.

Immunophenotypic analysis and cell sorting

For isolation and analysis of HSCs, the indicated progenitor populations and mature cells by FACS were performed as previously described (Inlay et al., 2009; Pronk and Bryder, 2011; Pronk et al., 2007; Tung et al., 2004). In brief, single-cell suspensions from BM, spleen, thymus and peripheral blood were prepared and stained with fluorescently labeled antibodies and subjected to FACS analysis and/or sorting. In some cases, overall organ cellularity was also assessed using a Sysmex KX-21 N machine (Sysmex) and intersected with FACS data to determine the absolute number of the indicated cell fractions in the different organs. For cell sorting, cell suspensions were depleted of mature hematopoietic cells using biotinylated antibodies against B220 (RA3-6B2), Gr-1 (RB6-8C5), CD11b (M1/70), CD8a (53–6.7), Ter-119, CD4 (GK1.5) (all from Biolegend) and anti-biotin MACS beads (Miltenyi Biotec), and was followed by staining with fluorescently labeled antibodies. Cell sorting was performed on FACS Aria I,II and III cell sorters and analysis on an LSRII (Becton Dickinson). The immunophenotypes of the cells used for FACS analysis and sorting were: HSC; Lin$^-$cKit$^+$Sca1$^+$CD48$^-$CD150$^+$, GMLP; Lin$^-$cKit$^+$Sca1$^+$CD48$^+$CD150$^-$Flt3$^+$, pGM; Lin$^-$cKit$^+$Sca1$^-$CD105$^-$CD150$^-$CD16/32$^-$, GMP; Lin$^-$cKit$^+$Sca1$^-$CD105$^-$CD150$^-$CD16/32$^+$, ALP; Lin$^-$cKit^{low}Sca1^{low}Flt3$^+$IL7Ra$^+$Ly6D$^-$, BLP; Lin$^-$Kit^{low}Sca1^{low}Flt3$^+$IL7Ra$^+$Ly6D$^+$, pMegE; Lin$^-$cKit$^+$Sca1$^-$CD105$^-$CD150$^+$CD41$^-$, pCFU-E; Lin$^-$cKit$^+$Sca1$^-$CD105$^+$CD150$^+$CD41$^-$, CFU-E; Lin$^-$cKit$^+$Sca1$^-$CD105$^+$CD150$^-$CD41$^-$, MkP; Lin$^-$cKit$^+$Sca1$^-$CD150$^+$CD41$^+$, Fr. A; Lin$^-$B220$^+$CD19$^-$CD24$^-$IgM$^-$IgD$^-$, Fr. B-C; Lin$^-$B220$^+$CD19^{high}CD24^{low}IgM$^-$IgD$^-$, Fr. C'-D; Lin$^-$B220$^+$CD19^{low}CD24^{high}IgM$^-$IgD$^-$, Fr. E;

$Lin^-B220^+CD19^{low}CD24^{high}IgM^+IgD^-$, Fr. F; $Lin^-B220^+CD19^{high}CD24^{low}IgM^{-/+}IgD^{-/+}$, Follicular B; $Lin^-B220^+CD19^+IgM^{low}IgD^+CD21^+$, Marginal Zone B; $Lin^-B220^+CD19^+IgM^+IgD^{low}CD5^-CD43^-CD21^+$, Immature B; $Lin^-B220^+CD19^+IgM^+IgD^{low}CD5^-CD43^-CD21^-$, ETP/DN1; $Lin^-CD4^-CD8^-cKit^+CD25^-$; DN2; $Lin^-CD4^-CD8^-cKit^+CD25^+$; DN3; $Lin^-CD4^-CD8^-cKit^-CD25^+$; CD4+CD8+; $Lin^-CD4^+CD8^+cKit^-CD25^-$; CD4+; $Lin^-CD4^+CD8^-cKit^-CD25^-$; CD8+; $Lin^-CD4^-CD8^+cKit^-CD25^-$.

In vitro evaluation of NK-, B-, and T cell potential by OP9/OP9-DL1 coculture

To determine the differentiation potential of bulk sorted GMLPs from Hlf inducible mice (Figure 1E), cells were sorted into 48-well plates pre-plated with OP9 (Kodama et al., 1994) or OP9-DL1 stromal cells. The lineage potential of single GMLPs was assessed by culturing individually FACS-deposited GMLPs in OP9 coated 96-well plates. For assessments of the differentiation potential following the lentiviral transduction in Figure 1D, 3,000 WT GMLPs were transduced twice during 36h on retronectin-coated plates (Takara) and were next divided over three wells of a 6-well plate pre-plated with OP9 or OP9-DL1 stromal cells. For the experiments involving the M33-Hlf fusion construct shown in Figure 4B and 4C, 2,000 WT GMLPs were transduced overnight with a retroviral M33-Hlf vector on retronectin-coated plates (Takara) and were subsequently divided over three wells of a 6-well plate pre-plated with OP9 and three wells pre-plated with OP9-DL1 stromal cells. To investigate effects of candidate Hlf targets, 2,600 WT GMLPs were transduced on retronectin-coated plates overnight with the indicated viruses and were thereafter divided over three wells of a 6-well plate pre-plated with OP9 and three wells pre-plated with OP9-DL1 stromal cells. To evaluate the effect of Nfic overexpression, 32,000 GMLPs were sorted and prestimulated in OptiMEM and cytokines for 6h. Cells were next retrovirally transduced on retronectin-coated 96-well plates (Takara). After 24h, cells were split (between 100 and 5,000 cells) onto OP9 cells preplated 3h before. Frequencies of GFP+ cells were analyzed every 48h by FACS for 14 days. For B/NK cell permissive cultures, cells were grown on OP9 stroma cells and supplemented with 20 ng/mL interleukin 15 (IL-15), 40 ng/mL IL-2, 10 ng/mL Stem Cell Factor (SCF), 10 ng/mL fms-like tyrosine kinase 3 ligand (Flt3L), and 10 ng/mL IL-7. For T cell generation, cells were grown on OP9-DL1 stroma supplemented with 10 ng/mL Flt3L, and 10 ng/mL IL-7. The cocultures were maintained in the presence or absence of 1 ug/ml DOX (Sigma-Aldrich) and evaluated at the indicated time points by cell counting and FACS staining with CD19 (1D3), B220 (RA3-6B2), Gr-1 (RB6-8C5), NK1.1 (PK136) and PI for B- and NK cell OP9 cocultures, and CD19 (1D3), CD90.2/Thy1.2 (53-2.1), CD25 (7D4), c-Kit (2B8) and PI for T cell OP9-DL1 cocultures. The basal medium for maintaining the OP9/OP9-DL1 stromal cells, as well as the cocultures, was OptiMEM (Invitrogen), 10% fetal calf serum (Sigma-Aldrich), 50 µg/mL gentamicin (Invitrogen) and 50µM β-mercaptoethanol (Invitrogen). The OP9 and OP9-DL1 stroma cells were pre-plated 3h prior to the addition of hematopoietic cells at a seeding density of 2,000 cells/cm^2.

Evaluation of apoptosis

To investigate whether Hlf expression resulted in increased levels of apoptosis in B and T lymphoid progenitors, 100,000 thymic $CD4^+CD8^+$ and 100,000 bone marrow-derived Fraction B-C cells ($Lin^-B220^+CD19^{high}CD24^{low}IgM^-IgD^-$) were FACS sorted from Hlf inducible mice and cultured at 37ºC in wells of a 24-well plate preplated with 10,000 OP9-DL1 and 10,000 OP9 respectively in the absence or presence of DOX (1 ug/ml) for 48h. Next, the cells were harvested and the Fraction B-C cultures were first stained with a c-Kit antibody (2B8, eBioscience) for 30 min. Thereafter, both culture types were incubated with Annexin V conjugated to Cy5 and Propidium Iodide for 15 minutes before immediate FACS analysis. The OP9 cultures were maintained in basal OP9/OP9-DL1 medium supplemented with 10 ng/ml Flt3L, 10 ng/ml IL-7 and 50 ng/ml SCF, while the OP9-DL1 cultures were maintained in basal medium supplemented with 10 ng/ml Flt3L, 10 ng/ml IL-2 and 10 ng/ml IL-7.

Affymetrix gene expression analysis and qRT-PCR

The microarray data in Figure 1B and S1 can be found in the Gene Expression Omnibus (GEO) (HSC accession numbers GSE44923 and GSE27686) (Pre Meg/E, Pre CFU-E, CFU-E, MkP, pGM and CLP, accession number GSE8407) (GMLP, accession number GSE18734) (GMP, accession number GSE14833). For data preprocessing, probe level expression values were extracted using RMAExpress (Bolstad et al., 2003) and analyses were performed using the dChip software (Li and Hung Wong, 2001) by filtering out probes with a lower expression than 50 in all subsets to eliminate noise in expression, followed by fold-change calculations and hierarchical clustering. For the quantitative RT-PCR (qRT-PCR) experiments described in Figure S1 and S2, the indicated cellular fractions were FACS sorted directly into RLT lysis buffer and purified using the RNeasy Micro mRNA purification kit (Qiagen), followed by first-strand cDNA synthesis as previously described (Norddahl et al., 2011). qRT-PCR reactions were run with SYBR GreenER (Invitrogen). For the qRT-PCR experiments depicted in Figure S5B, RNA from the sorted populations was purified using the RNeasy Micro mRNA purification kit, but was followed by first-strand and second-strand synthesis using Superscript II (Invitrogen) and amplification using KAPA HiFi Hotstart Readymix (Kapa Biosystems Inc.) and used for qRT-PCR using EvaGreen (Bio-rad).

ChIP-seq

80,000 WT $Lin^-Sca\text{-}1^+c\text{-}Kit^+$ cells were infected on retronectin coated plates (Takara) with either a pMX-GFP control or a pMX-3xFLAG-Hlf/Hlf-IRES-GFP virus and maintained for 5 days in basal OP9 medium supplemented with 50 ng/ml SCF, 10 ng/ml IL-7, 10 ng/ml Flt3L and 5 ng/ml IL-3 (all from Peprotech). Next, 2 x 10^7 cells were cross-linked in 1% formaldehyde for 12 min at room temperature with constant stirring before quenching the crosslinking reaction by the addition of glycine to a concentration of 0.125M and incubation for 5 min at room temperature. Cells were washed and nuclei were prepared by incubation on ice in lysis buffer (10 mM Tris pH8.0, 10 mM NaCl, 0.2% NP40 containing protease inhibitors (protease inhibitor cocktail (Sigma) and PMSF) for

10 min. The nuclei were harvested by centrifugation at 600 x g for 5 minutes at 4° C and snap-frozen in a dry ice/isopropanol bath. The frozen nuclei were resuspended in 1 ml of nuclei lysis buffer (50mM Tris pH 8.0, 10mM EDTA, 1% SDS) containing protease inhibitors (as above) and incubated on ice for 10 minutes. An equal volume of IP dilution buffer (20mM Tris pH 8.0, 2mM EDTA, 150mM NaCl, 1% Triton X-100, 0.01% SDS) containing protease inhibitors was added and chromatin was sonicated in a Bioruptor (Diagenode) for 4 cycles (30s on, 30s off). The chromatin solution was centrifuged for 10 minutes at 3220 x g and the supernatant was transferred to a new tube. An additional 3 ml of IP buffer was added together with 50 µl of rabbit IgG (2 µg/µl) and incubated at 4° C for 1 hour. 200µl of Protein G sepharose beads (1:1 slurry in IP dilution buffer) were added to the chromatin solution and further incubated at 4° C for 2 hours. The beads/IgG were collected by centrifugation at 1800 x g for 2 minutes. The chromatin was transferred to 1.5 ml tubes, an input sample was removed and 7 µg anti-FLAG (F3165, Sigma-Aldrich) were added then incubated overnight at 4° C with rotation. 60µl of protein G agarose beads (1:1 slurry in IP dilution buffer) were added and incubated with the samples for 2 hours. The beads were harvested at 5400 x g for 2 minutes and washed twice with low salt buffer (20mM Tris pH 8.0, 2mM EDTA, 50mM NaCl, 1% Triton X-100, 0.1% SDS), then once with LiCl buffer (10mM Tris pH 8.0, 1mM EDTA, 0.25M LiCl, 1% NP40, 1% Sodium deoxycholate monohydrate) and twice with 1x TE pH 8.0. The complexes were eluted twice from the beads by adding 150µl elution buffer (100mM NaHCO3, 1% SDS). To reverse the cross-linking and deproteinate the samples, 0.3M NaCl, RnaseA and Proteinase K were added to all the IP samples and input, followed by overnight incubation at 65° C. DNA was purified using Qiagen PCR clean up columns. Sequencing libraries were prepared using the Illumina TruSeq ChIP Library preparation kit (IP-202-1012, Illumina) according to the manufacturers instructions. Libraries were quantified using the KAPA universal library quantification kit (KK4824, KAPA Biosystems) and the average size estimated using an Agilent Bioanalyzer with an Agilent DNA 1000 chip (Agilent Technologies). Libraries were sequenced on a Illumina HiSeq 2500. Raw sequence reads in fastq format were mapped to the mouse genome (Genome Reference Consortium Mouse Build 38 - mm10) using Bowtie2 (Langmead and Salzberg, 2012). Peak regions were called using MACS2 (Zhang et al., 2008) and density plots generated in bigWig format. These were then displayed together as custom tracks on the UCSC Genome Browser. Peaks were assigned to genes if a) the peak was in a promoter according to MPromDB (Gupta et al., 2011), or b) if the peak was not in a promoter but intragenic, or c) if the peak was neither in a promoter or intragenic but intergenic within 50 kb from the start or end of a gene. Motif discovery was performed using the HOMER program (Heinz et al., 2010). To compute histone acetylation density profiles around the Hlf peaks we used annotatePeaks.pl from Homer (Heinz et al., 2010). Given the bigWig profiles from Lara-Astiaso et al. 2014, we set up Homer to produce the histograms in regions ±1kbp around the center of the peaks of the Hlf ChIP-Seq experiment. We plotted this set of histograms from a given histone experiment together as a heatmap-plot, where each row represents a genomic region around an Hlf peak and the color intensity is proportional to the sequencing library density. Additionally, we plotted the cumulative density around the peaks' center on top. For convenience, we sorted the heatmaps according to Hlf peak height. The script to produce the histograms (bw2histogram.sh, Linux-bash) and to plot heatmaps (HistoneMap MATLAB) can be found at: https://github.com/mscastillo/ChIP-Seq. The naming convention used in Figure 5B is Factor_cell type/line_CL (cell line)/PC (primary cells) and the heatmap was

generated using tools and data available on http://codex.stemcells.cam.ac.uk (Sánchez-Castillo et al., 2015). The generated ChIP-seq data is available on GEO under accession number GSE69817.

RNA-seq

Duplicate cultures of WT GMLPs, transduced overnight on retronectin coated plates with pHAGE2-Hlf/Hlf-IRES-ZsGreen or a pHAGE2 control virus, and Hlf inducible GMLPs maintained in the presence or absence of DOX, were cultured on OP9 stroma in B/NK cell permissive conditions for 4 days. After 4 days of coculture, $ZsGreen^{+}/CD45^{+}$ cells were sorted directly into RLT lysis buffer (1,000 – 50,000 cells). Following initial qualitative assessments using a Bioanalyzer (Agilent Technologies), the samples were subjected to library preparation (SMARTer Ultra Low Input RNA Kit for Sequencing, Clontech) and RNA-seq analysis using an Illumina HiSeq 2500 platform by the Genome Access Technology Center (GTAC, Washington University School of Medicine, St. Louis, MO). The data was next processed according to the GTAC RNA-seq analysis pipeline: sequence reads were mapped to the mouse genome (Ensemble version R72) using Tophat v2.0.8 (Trapnell et al., 2009) with Bowtie2 v2.1.0 (Langmead and Salzberg, 2012). Reads within exons were counted using HTSeq (Anders et al., 2015), and differentially expressed genes were identified using edgeR (Robinson et al., 2010) using the protocol described in the edgeR vignette, and called differential if the FDR was less than 0.05. The RNA-seq data is available in GEO under the accession number GSE69858.

Luciferase assays

To generate the vectors used for the luciferase assay, a 1kb region including the Nfic binding peak and located approximately 6.5kb downstream of the Nfic TSS (GRCm38.p5; Chr10:81413190-81414253), was cloned in the pGL2 vector followed by a SV40 promoter sequence upstream of the firefly luciferase reporter gene (= +6.5kb). The same procedure was used for a shorter 204bp region (GRCm38.p5; Chr10:81413934-81414137) which contains the three candidate HLF binding sites (= Min). As an additional control, an empty pGL2 vector was used. 10,000 Hlf inducible ES cells per sample were seeded onto irradiated MEFs in 48-well plates in complete media without antibiotics (DMEM, FCS 15%, NEAA 1x, Sodium pyruvate 1mM, b-mercaptoethanol 0,1mM and Leukemia Inhibitory Factor 10^{6} units/ml) in the absence or presence of doxycycline (1ug/ml). Each well was co-transfected using Lipofectamine LTX reagent (Invitrogen) with 180ng of each corresponding vector and 20ng of the pRL-TK vector containing the Renilla luciferase reporter gene. 24h after transfection, the luciferase reporter assay was performed using the Dual-Luciferase Reporter Assay System kit (Promega) according to the manufacturer's instructions.

RESOURCE TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
anti-FLAG	Sigma-Aldrich	Cat#F3165
B220	Biolegend	Clone: RA3-6B2
c-Kit	eBioscience	Clone: 2B8
CD105	Biolegend	Clone: MJ7/18
CD11b	Biolegend	Clone: M1/70
CD150	Biolegend	Clone: TC15-12F12.2
CD16/32	eBioscience	Clone: 93
CD19	Biolegend	Clone: 1D3
CD21	Biolegend	Clone: 7E9
CD24	Biolegend	Clone: M1/69
CD25	BD Pharmingen	Clone: 7D4
CD25	eBioscience	Clone: PC61.5
CD3ε	Biolegend	Clone: 17A2
CD4	Biolegend	Clone: GK1.5
CD41	BD Pharmingen	Clone: MWReg30
CD41	Biolegend	Clone: MWReg30
CD45.1	Biolegend	Clone: A20
CD45.2	Biolegend	Clone: 104
CD48	Biolegend	Clone: HM48-1
CD5	Biolegend	Clone: 53-7.3
CD8a	Biolegend	Clone: 53-7.6
CD8a	Biolegend	Clone: 53-6.7
Flt3	eBioscience	Clone: A2F10
Gr-1	Biolegend	Clone: RB6-8C5
IgD	Biolegend	Clone: 11-26c.2a
IgM	Biolegend	Clone: RMM-1

IL-7Rα	eBioscience	Clone: A7R34
Ly-6D	Sigma-Aldrich	Clone: RGRSL114.8.1
NK1.1	Biolegend	Clone: PK136
Sca1	Biolegend	Clone: D7
Ter-119	Biolegend	Clone: TER-119
Chemicals, Peptides, and Recombinant Proteins		
Annexin V – Cy5	BD Pharmingen	Cat#559933
BamHI	New England Biolabs	Cat#R0136S
Doxycycline food pellets (2g/kg)	Ssniff Specialdiäten	Cat#A112D72003
Doxycycline hyclate	Sigma-Aldrich	Cat#D9891-5G
EcoRI	New England Biolabs	Cat#R0101S
Formaldehyde	Sigma-Aldrich	Cat#F8775
Flt3L	PeproTech	Cat#300-19
IL-2	PeproTech	Cat#212-12
IL-3	PeproTech	Cat#213-13
IL-7	PeproTech	Cat#200-07
IL-15	PeproTech	Cat#210-15
M-CSF	PeproTech	Cat#315-02
NcoI	New England Biolabs	Cat#R0193S
NotI	New England Biolabs	Cat#R0189S
Propidium Iodide	Invitrogen	Cat#P3566
Retronectin	Takara	Cat#T100B
SCF	PeproTech	Cat#250-03
XhoI	New England Biolabs	Cat#R0146S
Critical Commercial Assays		
Anti-biotin Microbeads	Miltenyi Biotec	Cat#130-090-485
Dual-Luciferase Reporter Assay System kit	Promega	Cat#E1910

Gibson Assembly Cloning Kit	New England Biolabs	Cat#E5510S
Q5 High-Fidelity DNA Polymerase	New England Biolabs	Cat#M0491S
KAPA Hifi Hotstart Ready Mix	KAPA Biosystems	Cat#KK2601
KAPA Universal Library Quantification Kit	KAPA Biosystems	Cat#KK4824
Lipofectamine LTX	Invitrogen	Cat#15338100
RNeasy Micro Kit	Qiagen	Cat#74004
SMARTer Ultra Low Input RNA Kit for Sequencing	Clontech	Cat#634888
Ssofast EvaGreen Supermix with Low ROX	Biorad	Cat#1725212
Superscript II Reverse Transcriptase	Invitrogen	Cat#18064014
Superscript III Reverse Transcriptase	Invitrogen	Cat#18080093
SYBR GreenER qPCR SuperMix	Invitrogen	Cat#11762500
T4 DNA ligase	New England Biolabs	Cat#M0202S
Taq DNA polymerase	VWR	Cat#N224-500U
TOPO TA Cloning Kit	Invitrogen	Cat#K4500J10
TruSeq ChIP Library Preparation Kit	Illumina	Cat#IP-202-1012
Deposited Data		
ChIP-seq raw data	NCBI GEO	Accession number: GSE69817
ChIP-seq UCSC track to visualize our data along with that from Lara-Astiaso et al., 2014	NCBI GEO	(http://genome-euro.ucsc.edu/cgi-bin/hgTracks?hgS_doOtherUser=submit&hgS_otherUserName=promufa&hgS_otherUserSessionName=Lund_160719
RNA-seq raw data	NCBI GEO	Accession number: GSE69858
Experimental Models: Cell Lines		
Dox-inducible Hlf KH2 ES cells	This paper	N/A
KH2 ES cells	(Beard et al., 2006)	N/A

OP9 stromal cells	(Kodama et al., 1994)	N/A
OP9-DL1 stromal cells	(Schmitt and Zúñiga-Pflücker, 2002)	N/A
Plat-E cells	Cell Biolabs	Cat#RV-101
Lenti-X 293T cells	Clontech	Cat#632180
Experimental Models: Organisms/Strains		
Mouse: Dox-inducible Hlf mice	This paper	N/A
Oligonucleotides		
KH2 ES cells genotyping primers: Col1a1 fw: tccctcacttctcatccagatatt	Integrated DNA Technologies	(Jaako et al., 2011)
KH2 ES cells genotyping primers: Col1a1 WT rev: agtcttggatactccgtgaccata	Integrated DNA Technologies	(Jaako et al., 2011)
KH2 ES cells genotyping primers: Col1a1 Hlf rev: ggacaggataagtatgacatcatcaa	Integrated DNA Technologies	(Jaako et al., 2011)
KH2 ES cells genotyping primers: Rosa26 fw: aaagtcgctctgagttgttat	Integrated DNA Technologies	(Jaako et al., 2011)
KH2 ES cells genotyping primers: Rosa26 WT rev: ggagcgggagaaatggatatg	Integrated DNA Technologies	(Jaako et al., 2011)
KH2 ES cells genotyping primers: Rosa26 M2rtTA rev: gcgaagagtttgtcctcaacc	Integrated DNA Technologies	(Jaako et al., 2011)
pHAGE2 Hlf/Hlf assembly Fragment 1 Fw: caggtgtcgtgaagcatcacgATGGAGAAAATGTCCCGACAGCTCCCC	Integrated DNA Technologies	Own design
pHAGE2 Hlf/Hlf assembly Fragment 1 Rev: ccagtcccCAGGGGCCCGTGCCTGGC	Integrated DNA Technologies	Own design

pHAGE2 Hlf/Hlf assembly Fragment 2 Fw: ggcccctggggactggtggaggctcaggtggaggctcaggtggaggctcaggtggaggctcaggtggaggcactATGGAGAAAATGTCCCGACAGC	Integrated DNA Technologies	Own design
pHAGE2 Hlf/Hlf assembly Fragment 2 Rev: taggggggggggaggTTACAGGGGCCCGTGCCT	Integrated DNA Technologies	Own design
pHAGE2 Hlf/Dbp assembly Fragment 1 Fw: caggtgtcgtgaagcatcacgATGGAGAAAATGTCCCGACAGCTCCCC	Integrated DNA Technologies	Own design
pHAGE2 Hlf/Dbp assembly Fragment 1 Rev: ccagtcccCAGGGGCCCGTGCCTGGC	Integrated DNA Technologies	Own design
pHAGE2 Hlf/Dbp assembly Fragment 2 Fw: ggcccctggggactggtggaggctcaggtggaggctcaggtggaggctcaggtggaggctcaggtggaggcactATGGCGCGGCCTCTGAGC	Integrated DNA Technologies	Own design
pHAGE2 Hlf/Dbp assembly Fragment 2 Rev: taggggggggggaggTCACAGTGTCCCATGCTGGG	Integrated DNA Technologies	Own design
pHAGE2 Hlf/Tef assembly Fragment 1 Fw: caggtgtcgtgaagcatcacgATGGAGAAAATGTCCCGACAGCTCCCC	Integrated DNA Technologies	Own design

pHAGE2 Hlf/Tef assembly Fragment 1 Rev: ccagtcccCAGGGGCCCGTGCC TGGC	Integrated DNA Technologies	Own design
pHAGE2 Hlf/Tef assembly Fragment 2 Fw: ggcccctggggactggtggaggctcaggt ggaggctcaggtggaggctcaggtggag gctcaggtggaggcactATGTCCGA CGCGGGCGGC	Integrated DNA Technologies	Own design
pHAGE2 Hlf/Tef assembly Fragment 2 Rev: taggggggggggaggTTACAAGGG CCCGTACTTGGTCTCG	Integrated DNA Technologies	Own design
pHAGE2 Hlf/Nfil3 assembly Fragment 1 Fw: caggtgtcgtgaagcatcacgATGGAG AAAATGTCCCGACAGCTCCC C	Integrated DNA Technologies	Own design
pHAGE2 Hlf/Nfil3 assembly Fragment 1 Rev: ccagtcccCAGGGGCCCGTGCC TGGC	Integrated DNA Technologies	Own design
pHAGE2 Hlf/Nfil3 assembly Fragment 2 Fw: ggcccctggggactggtggaggctcaggt ggaggctcaggtggaggctcaggtggag gctcaggtggaggcactATGCAGCT GAGAAAAATGC	Integrated DNA Technologies	Own design
pHAGE2 Hlf/Nfil3 assembly Fragment 2 Rev: taggggggggggaggTTACCTGGA GTCCGAAGC	Integrated DNA Technologies	Own design

pHAGE2 HA-Hlf assembly Fw: caggtgtcgtgaagcatcacgatgtaccca tacgatgttccagattacgctGAGAAAA TGTCCCGACAGCTCC	Integrated DNA Technologies	Own design
pHAGE2 HA-Hlf assembly Rev: taggggggggggaggTTACAGGGG CCCGTGCCT	Integrated DNA Technologies	Own design
pHAGE2 Nfil3 cloning Fw: tagtgcggccgcATGCAGCTGAGA AAAATGCAG	Integrated DNA Technologies	Own design
pHAGE2 Nfil3 cloning Rev: tagtggatccTTACCTGGAGTCCG AAGCC	Integrated DNA Technologies	Own design
pHAGE2 3x FLAG-M33-Hlf assembly fragment 1 Fw: caggtgtcgtgaagcatcacgatgGCAG AGCAGCAGAGAGGG	Integrated DNA Technologies	Own design
pHAGE2 3x FLAG-M33-Hlf assembly fragment 1 Rev: cattttctcATAATGCCTCAAGTTG AAGAAGCC	Integrated DNA Technologies	Own design
pHAGE2 3x FLAG-M33-Hlf assembly fragment 2 Fw: aggcattatGAGAAAATGTCCCGA CAGCTCC	Integrated DNA Technologies	Own design
pHAGE2 3x FLAG-M33-Hlf assembly fragment 2 Rev: taggggggggggaggTTACAGGGG CCCGTGCCT	Integrated DNA Technologies	Own design

pHAGE2 3x FLAG-Hlf/Hlf fragment 1 assembly Fw: caggtgtcgtgaagcatcacgatggactac aaagaccatgacggtgattataaagatcat gatatcgattacaaggatgacgatgacaa gGAGAAAATGTCCCGACAGCT CCCCTTG	Integrated DNA Technologies	Own design
pHAGE2 3x FLAG-Hlf/Hlf assembly fragment 1 Rev: ccagtcccCAGGGGCCCGTGCC TGGC	Integrated DNA Technologies	Own design
pHAGE2 3x FLAG-Hlf/Hlf fragment 2 assembly Fw: ggcccctggggactggtggaggctcaggt ggaggctcaggtggaggctcaggtggag gctcaggtggaggcactATGGAGAA AATGTCCCGACAGC	Integrated DNA Technologies	Own design
pHAGE2 3x FLAG-Hlf/Hlf assembly fragment 2 Rev: taggggggggggaggTTACAGGGG CCCGTGCCT	Integrated DNA Technologies	Own design
MigR1 Hlf assembly Fw: gccggaattagatctcatcacgATGGA GAAAATGTCCCGACAGC	Integrated DNA Technologies	Own design
MigR1 Hlf assembly Rev: agggggggggggcggTTACAGGG GCCCGTGCCT	Integrated DNA Technologies	Own design
pMX 3x FLAG-Hlf/Hlf assembly Fw: tccaccggtcgccaccatggATCACGA TGGACTACAAAG	Integrated DNA Technologies	Own design
pMX 3x FLAG-Hlf/Hlf assembly Fw: tccaccggtcgccaccatggATCACGA TGGACTACAAAG	Integrated DNA Technologies	Own design

pMX 3x FLAG-Hlf/Hlf assembly Rev: ctcgcccttgctcaccatggTGTGGCC ATATTATCATCG	Integrated DNA Technologies	Own design
MigR1 3x FLAG-M33-Hlf assembly Fw: gccggaattagatctcATCACGATGG ACTACAAAGACCATGAC	Integrated DNA Technologies	Own design
MigR1 3x FLAG-M33-Hlf assembly Rev: agggggggggggcggTTACAGGG GCCCGTGCCT	Integrated DNA Technologies	Own design
MigR1 Cebpb assembly Fw: gccggaattagatctcATGCACCGC CTGCTGGCC	Integrated DNA Technologies	Own design
MigR1 Cebpb assembly Rev: agggggggggggcggCTAGCAGTG GCCCGCCGAG	Integrated DNA Technologies	Own design
MigR1 Als2 assembly Fw: tctctaggcgccggaattagatctccaccA TGGACTCAAAGAAGAAAAGC	Integrated DNA Technologies	Own design
MigR1 Als2 assembly Rev: cagtaacgttagggggggggggcggCT AGTTAAGCTTCTCCCG	Integrated DNA Technologies	Own design
MigR1 Ctnna1 assembly Fw: aggcgccggaattagatctccaccATGA CTGCCGTCCACGCA	Integrated DNA Technologies	Own design
MigR1 Ctnna1 assembly Rev: acgttagggggggggggcggTCAGAT GCTGTCCATGGCTTTG	Integrated DNA Technologies	Own design

MigR1 Eps8 assembly Fw: aggcgccggaattagatctccaccATGA ATGGTCATATGTCTAACCGC	Integrated DNA Technologies	Own design
MigR1 Eps8 assembly Rev: acgttagggggggggggcggTCAGTG GCTGCTCCCTTC	Integrated DNA Technologies	Own design
MigR1 Gas6 assembly Fw: aggcgccggaattagatctccaccATGC CGCCACCGCCCGGG	Integrated DNA Technologies	Own design
MigR1 Gas6 assembly Rev: acgttagggggggggggcggCTAGG GGGTGGCATGCTCCACAGG	Integrated DNA Technologies	Own design
MigR1 Gcnt2 assembly Fw: aggcgccggaattagatctccaccATGG GCTCTTGGAAGTAC	Integrated DNA Technologies	Own design
MigR1 Gcnt2 assembly Rev: acgttagggggggggggcggTCAGAA ATACCAGCTCGG	Integrated DNA Technologies	Own design
MigR1 Hip1 assembly Fw: aggcgccggaattagatctccaccATGG ACCGAATGGCCAGC	Integrated DNA Technologies	Own design
MigR1 Hip1 assembly Rev: acgttagggggggggggcggCTACTC TTTGTCCGGTATTGCTTC	Integrated DNA Technologies	Own design
MigR1 Il15 assembly Fw: aggcgccggaattagatctccaccATGA AAATTTTGAAACCATATATGA G	Integrated DNA Technologies	Own design

MigR1 Il15 assembly Rev: acgttagggggggggggcggTCAGG ACGTGTTGATGAAC	Integrated DNA Technologies	Own design
MigR1 Lpl assembly Fw: aggcgccggaattagatctccaccATGG AGAGCAAAGCCCTG	Integrated DNA Technologies	Own design
MigR1 Lpl assembly Rev: acgttagggggggggggcggTCAGC CAGACTTCTTCAG	Integrated DNA Technologies	Own design
MigR1 Nedd4 assembly Fw: tctctaggcgccggaattagatctccaccA TGAGCTCGGACATGGCA	Integrated DNA Technologies	Own design
MigR1 Nedd4 assembly Rev: cagtaacgttagggggggggggcggCT AATCAACGCCATCAAAGC	Integrated DNA Technologies	Own design
MigR1 Nfic assembly Fw: aggcgccggaattagatctccaccATGT ATTCCTCCCCGCTC	Integrated DNA Technologies	Own design
MigR1 Nfic assembly Rev: acgttagggggggggggcggCTAATC CCACAAAGGGAC	Integrated DNA Technologies	Own design
MigR1 Sdc1 assembly Fw: aggcgccggaattagatctccaccATGA GACGCGCGGCGCTC	Integrated DNA Technologies	Own design
MigR1 Sdc1 assembly Rev: acgttagggggggggggcggTCAGG CGTAGAACTCCTCCTGC	Integrated DNA Technologies	Own design

MigR1 Stom assembly Fw: aggcgccggaattagatctccaccATGT CTGACAAACGGCAG	Integrated DNA Technologies	Own design
MigR1 Stom assembly Rev: acgttagggggggggggcggTCAGTG ATTAGAACCCATG	Integrated DNA Technologies	Own design
MigR1 Mgst2 assembly Fw: aggcgccggaattagatctccaccATGG CCGGGGATTCAAGC	Integrated DNA Technologies	Own design
MigR1 Mgst2 assembly Rev: acgttagggggggggggcggTTAGAA GGGCTTCCTCAGTTTC	Integrated DNA Technologies	Own design
MigR1 Sh2d2a assembly Fw: aggcgccggaattagatctccaccATGG AGTTCTGCTTGGCCC	Integrated DNA Technologies	Own design
MigR1 Sh2d2a assembly Rev: acgttagggggggggggcggTCAGG AGGGGCTCCCTCT	Integrated DNA Technologies	Own design
MigR1 St6gal1 assembly Fw: aggcgccggaattagatctccaccATGA TTCATACCAACTTGAAGAG	Integrated DNA Technologies	Own design
MigR1 St6gal1 assembly Rev: acgttagggggggggggcggTCAACA GCGATTGTTCCG	Integrated DNA Technologies	Own design
qRT-PCR primer: Actb fw: CCACAGCTGAGAGGCAAATC	Integrated DNA Technologies	Own design
qRT-PCR primer: Actb rev: CTTCTCCAGGGAGGAAGAGG	Integrated DNA Technologies	Own design

qRT-PCR primer: Ccnb1 fw: AAGGTGCCTGTGTGTGAACC	Integrated DNA Technologies	Own design
qRT-PCR primer: Ccnb1 rev: GTCAGCCCCATCATCTGCG	Integrated DNA Technologies	Own design
qRT-PCR primer: Cfp fw: GAATGTGGCTCCTGGAACTC	Integrated DNA Technologies	Own design
qRT-PCR primer: Cfp rev: TTTGGAGCATGTGACAGAGC	Integrated DNA Technologies	Own design
qRT-PCR primer: Dntt fw: GCCATCCGTGTAGATCTGGT	Integrated DNA Technologies	Own design
qRT-PCR primer: Dntt rev: GCCGCAAGTCTCTCTCAAAC	Integrated DNA Technologies	Own design
qRT-PCR primer: Ebf1 fw: AGCTGCCAACTCACCCTATG	Integrated DNA Technologies	Own design
qRT-PCR primer: Ebf1 rev: CACTGCTGAGACCATGTTGG	Integrated DNA Technologies	Own design
qRT-PCR primer: Erg fw: CAGTAGCCGCCTTGCTAATC	Integrated DNA Technologies	Own design
qRT-PCR primer: Erg rev: TGATGCAGTTGGAGTTGGAG	Integrated DNA Technologies	Own design
qRT-PCR primer: Flt3 fw: TTCCTGCCTCTGGGTCTTTA	Integrated DNA Technologies	Own design
qRT-PCR primer: Flt3 rev: CTGGGTCTCTGTCACGTTCA	Integrated DNA Technologies	Own design
qRT-PCR primer: Hlf fw: GACCCACCTTATGGGACAAA	Integrated DNA Technologies	Own design
qRT-PCR primer: Hlf rev: GGATGCCATTCTCTGACAGG	Integrated DNA Technologies	Own design
qRT-PCR primer: Icosl fw: AGTTCACATGCCGGGTATT	Integrated DNA Technologies	Own design
qRT-PCR primer: Icosl rev: TCAGAGGTGCTGATGACAGG	Integrated DNA Technologies	Own design
qRT-PCR primer: Ikzf3 fw: TACAACCGACTGTGGAGCTG	Integrated DNA Technologies	Own design

qRT-PCR primer: Ikzf3 rev: GAGGTTTGGGCAAGCTGTAG	Integrated DNA Technologies	Own design
qRT-PCR primer: Lambda5 fw: CCATCTAAGCCCCAGTTTTG	Integrated DNA Technologies	Own design
qRT-PCR primer: Lambda5 rev: GGAAGGCAGGAACAGAGTGA	Integrated DNA Technologies	Own design
qRT-PCR primer: Lef1 fw: ACGACAAGGCCAGAGAACAC	Integrated DNA Technologies	Own design
qRT-PCR primer: Lef1 rev: TGTACGGGTCGCTGTTCATA	Integrated DNA Technologies	Own design
qRT-PCR primer: Lgr5 fw: CCCAATGCGTTTTCTACGTT	Integrated DNA Technologies	Own design
qRT-PCR primer: Lgr5 rev: AGGCTCGGTTCCCTGTTAAT	Integrated DNA Technologies	Own design
qRT-PCR primer: Pax5 fw: GGGAGACCTGTTCACACAGC	Integrated DNA Technologies	Own design
qRT-PCR primer: Pax5 rev: CCATGGCTGAATACTCTGTGG	Integrated DNA Technologies	Own design
qRT-PCR primer: Rag1 fw: CTCTCAGGGAGCTCATGGAC	Integrated DNA Technologies	Own design
qRT-PCR primer: Rag1 rev: CGAAACGCTGTGAGTTGAAA	Integrated DNA Technologies	Own design
qRT-PCR primer: Vpreb3 fw: CCTGCCTCTGCTCCTGATAG	Integrated DNA Technologies	Own design
qRT-PCR primer: Vpreb3 rev: CAGCTGAGATGAGCGTCTTG	Integrated DNA Technologies	Own design
Recombinant DNA		
MigR1	(Pear et al., 1998)	N/A
pBS31 targeting vector	(Beard et al., 2006)	N/A
pHAGE2	(Mostoslavsky et al., 2005)	N/A
pMX-GFP	Cell Biolabs	Cat#RTV-050

Forced dimer linker sequence (GGGACTGGTGGAGGCTCAGGTG GAGGCTCAGGTGGAGGCTCAGG TGGAGGCTCAGGTGGAGGCACT)	(Neuhold and Wold, 1993)	N/A
pYX-Asc-Hlf	K.K. DNAFORM	Clone: MGC:76396
Lentiviral construct Hlf/Hlf	This paper	pHAGE2 Hlf/Hlf
Lentiviral construct Hlf/Dbp	This paper	pHAGE2 Hlf/Dbp
Lentiviral construct Hlf/Tef	This paper	pHAGE2 Hlf/Tef
Lentiviral construct Hlf/Nfil3	This paper	pHAGE2 Hlf/Nfil3
Lentiviral construct Hlf	This paper	pHAGE2 HA-Hlf
Lentiviral construct Nfil3	This paper	pHAGE2 Nfil3
Lentiviral construct M33-Hlf	This paper	pHAGE2 M33-Hlf
Lentiviral construct 3x FLAG-Hlf/Hlf	This paper	pHAGE2 3x FLAG-Hlf/Hlf
Retroviral construct Hlf	This paper	MigR1 Hlf
Retroviral construct 3x FLAG-Hlf/Hlf	This paper	pMX 3x FLAG-Hlf/Hlf
Retroviral construct M33-Hlf	This paper	MigR1 3x FLAG-M33-Hlf
Retroviral construct Cebpb	This paper	MigR1 Cebpb
Retroviral construct Als2	This paper	MigR1 Als2
Retroviral construct Ctnna1	This paper	MigR1 Ctnna1
Retroviral construct Eps8	This paper	MigR1 Eps8
Retroviral construct Gas6	This paper	MigR1 Gas6
Retroviral construct Gcnt2	This paper	MigR1 Gcnt2
Retroviral construct Hip1	This paper	MigR1 Hip1
Retroviral construct Il15	This paper	MigR1 Il15
Retroviral construct Lpl	This paper	MigR1 Lpl
Retroviral construct Nedd4	This paper	MigR1 Nedd4
Retroviral construct Nfic	This paper	MigR1 Nfic
Retroviral construct Sdc1	This paper	MigR1 Sdc1
Retroviral construct Stom	This paper	MigR1 Stom
Retroviral construct Mgst2	This paper	MigR1 Mgst2

Retroviral construct Sh2d2a	This paper	MigR1 Sh2d2a
Retroviral construct St6gal1	This paper	MigR1 St6gal1
pGL2 Nfic +6.5kb	This paper	pGL2 Nfic +6.5kb
pGL2 Nfic Min	This paper	pGL2 Nfic Min
Software and Algorithms		
Bowtie2	(Langmead and Salzberg, 2012)	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
dChip	(Li and Hung Wong, 2001)	https://sites.google.com/site/dchipsoft/
edgeR	(Robinson et al., 2010)	
GEN-E		http://www.broadinstitute.org/cancer/software/GENE-E/index.html
Genome Reference Consortium Mouse Build 38 - mm10	Genome Reference Consortium	http://www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/mouse/
HistoneMap MATLAB		https://github.com/mscastillo/ChIP-Seq
HOMER	(Heinz et al., 2010)	http://homer.ucsd.edu/homer/
HTSeq	(Anders et al., 2015)	
MACS2	(Zhang et al., 2008)	https://github.com/taoliu/MACS/
MPromDB	(Gupta et al., 2011)	http://bioinformatics.wistar.upenn.edu/MPromDb/
Venny	(Oliveros, 2007-2015)	http://bioinfogp.cnb.csic.es/tools/venny/

SUPPLEMENTAL REFERENCES

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166-169.

Argiropoulos, B., Yung, E., Xiang, P., Lo, C.Y., Kuchenbauer, F., Palmqvist, L., Reindl, C., Heuser, M., Sekulovic, S., Rosten, P.*, et al.* (2010). Linkage of the potent leukemogenic activity of Meis1 to cell-cycle entry and transcriptional regulation of cyclin D3. Blood *115*, 4071-4082.

Beard, C., Hochedlinger, K., Plath, K., Wutz, A., and Jaenisch, R. (2006). Efficient method to generate single-copy transgenic mice by site-specific integration in embryonic stem cells. Genesis *44*, 23-28.

Bolstad, B.M., Irizarry, R.A., Astrand, M., and Speed, T.P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics *19*, 185-193.

Gupta, R., Bhattacharyya, A., Agosto-Perez, F.J., Wickramasinghe, P., and Davuluri, R.V. (2011). MPromDb update 2010: an integrated resource for annotation and visualization of mammalian gene promoters and ChIP-seq experimental data. Nucleic Acids Res *39*, D92-97.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576-589.

Inlay, M.A., Bhattacharya, D., Sahoo, D., Serwold, T., Seita, J., Karsunky, H., Plevritis, S.K., Dill, D.L., and Weissman, I.L. (2009). Ly6d marks the earliest stage of B-cell specification and identifies the branchpoint between B-cell and T-cell development. Genes Dev *23*, 2376-2381.

Jaako, P., Flygare, J., Olsson, K., Quere, R., Ehinger, M., Henson, A., Ellis, S., Schambach, A., Baum, C., Richter, J.*, et al.* (2011). Mice with ribosomal protein S19 deficiency develop bone marrow failure and symptoms like patients with Diamond-Blackfan anemia. Blood *118*, 6087-6096.

Kodama, H., Nose, M., Niida, S., and Nishikawa, S. (1994). Involvement of the c-kit receptor in the adhesion of hematopoietic stem cells to stromal cells. Exp Hematol *22*, 979-984.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods *9*, 357-359.

Li, C., and Hung Wong, W. (2001). Model-based analysis of oligonucleotide arrays: model validation, design issues and standard error application. Genome Biol *2*, RESEARCH0032.

Mostoslavsky, G., Kotton, D.N., Fabian, A.J., Gray, J.T., Lee, J.S., and Mulligan, R.C. (2005). Efficiency of transduction of highly purified murine hematopoietic stem cells by lentiviral and oncoretroviral vectors under conditions of minimal in vitro manipulation. Mol Ther *11*, 932-940.

Neuhold, L.A., and Wold, B. (1993). HLH forced dimers: tethering MyoD to E47 generates a dominant positive myogenic factor insulated from negative regulation by Id. Cell *74*, 1033-1042.

Norddahl, G.L., Pronk, C.J., Wahlestedt, M., Sten, G., Nygren, J.M., Ugale, A., Sigvardsson, M., and Bryder, D. (2011). Accumulating mitochondrial DNA mutations drive premature hematopoietic aging phenotypes distinct from physiological stem cell aging. Cell stem cell *8*, 499-510.

Oliveros, J.C. (2007-2015). *Venny. An interactive tool for comparing lists with Venn's diagrams.* http://bioinfogpcnbcsices/tools/venny/indexhtml.

Pear, W.S., Miller, J.P., Xu, L., Pui, J.C., Soffer, B., Quackenbush, R.C., Pendergast, A.M., Bronson, R., Aster, J.C., Scott, M.L.*, et al.* (1998). Efficient and rapid induction of a chronic

myelogenous leukemia-like myeloproliferative disease in mice receiving P210 bcr/abl-transduced bone marrow. Blood *92*, 3780-3792.
Pronk, C.J., and Bryder, D. (2011). Flow cytometry-based identification of immature myeloerythroid development. Methods in molecular biology *699*, 275-293.
Pronk, C.J., Rossi, D.J., Mansson, R., Attema, J.L., Norddahl, G.L., Chan, C.K., Sigvardsson, M., Weissman, I.L., and Bryder, D. (2007). Elucidation of the phenotypic, functional, and molecular topography of a myeloerythroid progenitor cell hierarchy. Cell stem cell *1*, 428-442.
Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139-140.
Schmitt, T.M., and Zúñiga-Pflücker, J.C. (2002). Induction of T cell development from hematopoietic progenitor cells by delta-like-1 in vitro. Immunity *17*, 749-756.
Sánchez-Castillo, M., Ruau, D., Wilkinson, A.C., Ng, F.S., Hannah, R., Diamanti, E., Lombard, P., Wilson, N.K., and Gottgens, B. (2015). CODEX: a next-generation sequencing experiment database for the haematopoietic and embryonic stem cell communities. Nucleic Acids Res *43*, D1117-1123.
Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105-1111.
Tung, J.W., Parks, D.R., Moore, W.A., and Herzenberg, L.A. (2004). Identification of B-cell subsets: an exposition of 11-color (Hi-D) FACS methods. Methods Mol Biol *271*, 37-58.
Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W.*, et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.