Supplemental methods

Hhex conditional knockouts

The following methods were briefly described in Smith et al[1]. Briefly, the murine *Hhex* gene was cloned from a 129 BAC by retrieval methods previously described[2]. A single *loxP* and a *loxP* plus *Frt*-flanked neo cassette were targeted to the resulting construct in two steps in EL350 cells through recombineering. First, to insert the single 5′ *loxP* site, a targeting cassette containing *Pgk-em7-neo* flanked by homology arms to 1st intron of *Hhex* was constructed in PL400. The homology arms are PCR amplified using the following primers:

5′-arm sense, 5′- CGC GAA GCT TGC AGA ACA TGA GTG TGA CCG-3′

5′-arm antisense, 5′- CGC GGA ATT CAT GAG AGC ACT TCC CAA GGC-3′;

3′-arm sense, 5′- CGC GCT CGA GAA AGA GCG CAC CCT GAG TCT -3′;

3′-arm antisense, 5′- CGC GGG ATC CCC AGA ACG CAA CCA TGT TCC -3′.

The homology arms were sequence verified, restriction digested, and cloned into PL400 via four-way ligation. The targeting cassette was released by *BamH*I/*Hin*dIII double digest and targeted through co-electroporation into heat shock-induced EL350 cells. The *Pgk-em7-neo* sequence was then removed by electroporation into arabinose-induced Cre-expressing EL350 cells, leaving behind a single *loxP* site. To insert the second *loxP* site after exon 4 of *Hhex*, a targeting cassette containing *frt-Pgk-Em7-neo-frt-loxP* flanked by homology arms to targeting site was constructed in PL451. Homology arms were amplified using the following primers:

5′-arm sense, 5′- CGC GAA GCT TCC TAA ACA TGA CAC CTA AAG -3′;

5′-arm antisense, 5′- CGC GGA ATT CCA CCC TGC TTG GTC CTC TTC-3′;

3′-arm sense, 5′- CGC GGG ATC CGA TTG GAG CTG CCA CTG AGT -3′;

3′-arm antisense, 5′- CGC GGC GGC CGC AGC AGC TGG AAC CTG ACA AC -3′.

The targeting cassette was released by *Not*I/*Hin*dIII double digest and targeted similarly as described above. The conditional targeting vector was then linearized by *Not*I digestion and electroporated into 129-derived CJ7 embryonic stem (ES) cells, using standard procedures. G418 (180 μg/ml) and ganciclovir (2 μM) double-resistant clones were analyzed by Southern blotting hybridization, using both 5′ and 3′ external probes. External probes were PCR amplified using the following primers:

5′ probe, sense, 5′- CCC CAC TAC ACC TGG CTA AC -3′

5′ probe, antisense, 5′- ACG TGG ATG GTA TCA AAG CC -3′

3′-probe, sense, 5′- GGG ATT TGT TGT TGC TGT GC-3′

3′-probe, antisense, 5′- CTG GAT GCT GGT GAC TCA GA -3′

Correctly targeted clones were then injected into C57BL/6 blastocysts using standard procedures, and resulting chimeras were mated with C57BL/6 females to obtain germline transmission of the targeted allele. The Neo cassette was removed by crossing to the *Flp* recombinase strain.

OP9 Assay

OP9-DL1 and OP9-GFP cells were maintained in culture as described[3]. Cells were cultured in α-MEM media with 20% fetal calf serum and 1% penicillin/streptomycin. Stem cells from the bone marrow of 6-8 week old mice were harvested and sorted using flow cytometry and $Lin^{-}Kit^{+}Sca\text{-}1^{+}$ (LSK) cells were collected. LSK cells were cultured in 24 well plates containing 75% confluent irradiated OP9-GFP cells with 6 ng/mL murine IL-7 and murine Flt3L (Peprotech Inc.). Cells were collected seven days after plating. Double negative thymocytes from the thymi of 8 week old mice were harvested and CD8 and CD4 was depleted by magnetic bead separation. DN cells were cultured in 24 well plates containing 75% confluent irradiated OP9-DL1 cells with 6 ng/mL murine IL-7 and murine Flt3L. Cells were collected, washed and plated on fresh OP9-DL1 cultures every 7 days.

RNA-seq and differential gene expression analyses

RNA-seq was performed as previously described[4]. Briefly, whole RNA was isolated by RNeasy kit (Qiagen) using established protocols. For RNA isolated from the LSK-OP9 co-culture, no amplification was performed. RNA was checked for quality on the Bioanalyzer. For RNA isolated from sorted LSK; $Flt3^{lo}$ and LSK; $Flt3^{hi}$ sorted cells, we had low input RNA (range 4.8ng-11.9ng) and was subjected to cDNA amplification using the SMARTer Ultra Low RNA Kit (Clontech, Catalog # 634936). The cDNA was sheared using the Covaris S2 instrument with the following parameters: duty 10%, Intensity 5, 200 Cycles/Burst Time 5 minutes, and mode frequency sweeping. Once the cDNA was sheared, the entire volume was used as input into the Illumina TruSeq ChIP Sample Prep Kit (Catalog #IP-202-1012) and the Illumina protocol was primarily used. However, instead of the Illumina PCR ChIP mastermix, the KAPA Hot Start PCR Kit (Kapa Biosystems) was used with only 15 cycles PCR to minimize duplication. The samples were then quality controlled, normalized, clustered, and sequenced according to Illumina's best practices. Multiple stage of quality control (QC) of sequencing data was carried out [5]. Raw data and alignment QC were performed using QC3 [6], expression analysis were carried out using MultiRankSeq [7]. All data passed QC. Raw data were aligned with TopHat 2 [8] and gene expression levels were quantified using Cufflinks [9]. RPKM (reads per kilobase per million reads) based approaches (Cuffdiff) were used to detect differentially expressed genes. False discovery rate (FDR < 0.05) was used to correct for multiple testing. To identify lincRNA, we performed gene annotation using lincRNA reference Gencode v19 released by the ENCODE project [10].

***Igh* gene rearrangement PCR**

Genomic DNA from sorted $B220^{+}$ cells from Hhex cKO and WT spleens were subjected to nested PCR analysis. Primers were designed to amplify four possible junctions between D-Q52 and JH regions114. The first round of PCR was completed using the following primers and PCR conditions: DH1 Q52-1 5'-CACAGACCTTTCTCCATAGTTGATAACTCAG- 3' and JH4-1 5'- AGGCTCTGAGATCCCTAGACAG-3'; denaturing at 95°C for 1 min, annealing at 60°C for 1 min, extension at 72°C for 2.5 min (28 cycles). The second round of PCR was completed using 4% of the first PCR with the following primers and

conditions: DH Q52-2 5'- GCCTCAGAGTCCCTGTGGTCTCTGACTGGT-3' and JH4-2 5'-GGGTCTAGACTCTCAGCCGGCTCCCTCAGGG-3'; denaturing at 95°C for 20 sec, annealing at 60°C for 1 min, extension at 72°C for 2 min (35 cycles). These were resolved on 0.8% agarose gels and visualized under UV.

Homing assay and engraftment analysis

To analyze homing, bone marrow from WT and Hhex cKO mice were harvested and used for transplantation. BM cells (80,000) were first plated in methylcellulose culture media for CFU-C assay to quantify the input number of HSPCs. Ten million donor cells were retro-orbitally injected into lethally irradiated $CD45.1^{+}$ host mice; 16 hours post-transplant bone marrow from recipient mice were collected and cultured in triplicate for CFU-C assay. Homing was calculated using the following equation:

homing (%)= (Number of output colonies/dish/ number of input colonies)x2x4x100/18x100.

To analyze engraftment, recipient CD45.1 mice were lethally irradiated and injected with donor bone marrow from either WT or *Hhex* cKO CD45.2 mice. 7 days post injection mice were sacrificed and flow cytometry was used to assess the presence of both CD45.1 and CD45.2 in bone marrow, spleen,

Supplemental References

1. Smith S, Tripathi R, Goodings C, et al. LIM Domain Only-2 (LMO2) Induces T-Cell Leukemia by Two Distinct Pathways. *PLoS ONE.* 2014;9(1):e85883.

2. Liu P, Jenkins NA, Copeland NG. A highly efficient recombineering-based method for generating conditional knockout mutations. *Genome Res.* 2003;13(3):476-484.

3. Schmitt TM, Zuniga-Pflucker JC. Induction of T cell development from hematopoietic progenitor cells by delta-like-1 in vitro. *Immunity.* 2002;17(6):749-756.

4. Cleveland SM, Smith S, Tripathi R, et al. Lmo2 induces hematopoietic stem cell-like features in T-cell progenitor cells prior to leukemia. *Stem Cells.* 2013;31(5):882-894.

5. Guo Y, Ye F, Sheng Q, Clark T, Samuels DC. Three-stage quality control strategies for DNA re-sequencing data. *Brief Bioinform.* 2013.

6. Guo Y, Zhao S, Sheng Q, et al. Multi-perspective quality control of Illumina exome sequencing data using QC3. *Genomics.* 2014.

7. Guo Y, Zhao S, Ye F, Sheng Q, Shyr Y. MultiRankSeq: Multiperspective Approach for RNAseq Differential Expression Analysis and Quality Control. *BioMed Research International.* 2014;2014:8.

8. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 2013;14(4):R36.

9. Trapnell C, Williams BA, Pertea G, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol.* 2010;28(5):511-515.

10. Consortium EP, Bernstein BE, Birney E, et al. An integrated encyclopedia of DNA elements in the human genome. *Nature.* 2012;489(7414):57-74.

11. Mingueneau M, Kreslavsky T, Gray D, et al. The transcriptional landscape of alphabeta T cell differentiation. *Nat Immunol.* 2013;14(6):619-632.

12. Fu W, Wojtkiewicz G, Weissleder R, Benoist C, Mathis D. Early window of diabetes determinism in NOD mice, dependent on the complement receptor CRIg, identified by noninvasive imaging. *Nature immunology.* 2012;13(4):361-368.

13. Dykstra B, Kent D, Bowie M, et al. Long-Term Propagation of Distinct Hematopoietic Differentiation Programs In Vivo. *Cell Stem Cell.* 2007;1(2):218-229.

14. Wilson A, Laurenti E, Oser G, et al. Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair. *Cell.* 2008;135(6):1118-1129.

15. Rossi DJ, Bryder D, Seita J, Nussenzweig A, Hoeijmakers J, Weissman IL. Deficiencies in DNA damage repair limit the function of haematopoietic stem cells with age. *Nature.* 2007;447(7145):725-729.