nature research

Corresponding author(s): Shahin Rafii MD; Peipei Guo, PhD; Duancheng Wen

Last updated by author(s): 08/16/2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

The peripheral blood readings were collected using Advia 120 hemological analyzer (Siemens); The quantification of hematopoietic stem cell number and progenitor cell numbers in the bone marrow, the quantification of the lineage composition of bone marrow, spleen are carried out using BD LSRII or BD Fortessa or BD FACS AriaII machine, using a software named BD Diva. The RNA-seq and ChIP-seq libraries are sequenced using Illumina Hiseq 4000 or Novaseq 6000 platforms. Quantitative real-time RT-PCR results are collected using ViiA-7, Thermo Fisher Scientific.

Data analysis

For flow cytometric data, data is analyzed in flowjo (version 10.4). Compensation is calculated using ultracompensation beads with single color controls, using the automatic compensation settings. In flowjo, the gating is drawn based on fluorescence minus one or unstained controls. The percentages of lineage cells among live cells are extracted and plotted.

Reads from RNA-seq were aligned to mouse genome version mm9 using TopHat. Fragments per kilobase of transcript per million fragments mapped (FPKM) were identified using Cufflinks with upper-quartile and GC-normalization. After log2 transformation of FPKM values, differentially expressed genes were then identified using limma packaged in R (3.5.3). A cutoff of pvalue < 0.05 and fold change smaller than abosolute(log2(1.8)) is used to identify downregulated genes. A cut off of pvalue < 0.05 and fold change bigger than log2(1.8) was used to identify upregulated genes. The gene ontology function was revealed using DAVID GO website, based on biological process and molecular functions.

Identification of significantly changed peaks using SICER_V1.1. Raw reads (fastq files) from ChIP-seq was aligned to mouse genome mm9 using bwa. Duplicate reads and reads aligning to more than one location were excluded. After QC and alignment, bam files were obtained. Bam files were then converted to bed files using bamtobed command (bedtools 2.27.1). SICER was used to identify the peaks for H3.3 (HA) by comparing to the input control sample using a FDR < 0.01, fragment size of 150bp, window size 200bp, gap size 600bp. For identification of significantly decreased or increased H3K9me3, H3K27me3, H3K27ac peak regions, bed files for BKO_H3K9me3, BKO_input, DKO_H3K9me3 and DKO_input were used. The parameters used are FDR < 0.01, fragment size of 150bp, window size 200bp, and gap size of 600bp. For H3K4me3 peak calling, gap size 200 bp is used. The peak region bed files were annotated to the nearest TSS using HOMER (v4.10.4) annotatepeaks.pl function. To visualize the changes of H3K9me3 at a given locus, the ChIP-seq bam files was

converted to tdf files (IGV genome browser version 2.4.10) and then visualized using IGV genome browser (Broad Institute). Normalization was selected for visualization. Annotations of RepeatMasker were loaded from the UCSC Table Browser by selecting the desired class of repeats.
To calculate the changes of histone PTM profiles between WT/BKO and DKO samples, deeptools (version 3.2.0) were used. Bam files were first converted to bw files using bamcoverage function with binsize 10bp and normalization method of RPGC (mm9 genome size is 2620345972). To compare the H3K9me3 marks at the significantly decreased or significantly increased H3K9me3 regions within WT and DKO HSPCs, computematrix scale region function was used. The normalized tag density at these peak regions was reported by "outFileNameMatrix" function, with each row representing the rows from original bed files, and the column represent the tag density within binned chromosome regions. The mean tag density of WT and DKO H3K9me3 are generated using plotProfile function. The heatmap and Kmeans clusters or Hclusters can be generated using plotHeatmap function.

Histone modification changes and motif analysis for promoter regions of differentially expressed genes
To examine how the histone modification changes influence the gene expressions, after obtaining the differentially expressed genes, the promoter regions of upregulated genes or downregulated genes was extracted to form a bed file in Rstudio. Bedtools "intersectBed" function was used to find the overlapping regions between H3K9me3 decreased peaks with the promoter regions of upregulated genes. To identify the transcription factors that influence of the differentially expressed genes, we carried out motif analysis for H3K27ac marked enhancer regions, or for the promoter regions of the differentially expressed genes. Motif discovery was carried out using HOMER, using size 400bp as option. The transcription factors identified are defined by HOMER. The reported transcription factors are selected based on the p-value.

Quantification of H3K9me3 changes at ERVs
The genome coordinates of different repfamilies of repeats are downloaded from UCSC genome browser. Deeptools compute matrix (reference point) function was used to calculate the tag density (H3K9me3) within each repeat region for WT or DKO HSCs. The matrix was exported using the outfileNameMatrix function; the boxplot of tag density at specified regions are plotted using R. T-test was used to test whether the resulting WT_H3K9me3 tags and DKO_H3K9me3 tags are significantly different from each other.

Quantification of mRNA expression changes of ERVs
Two methods were used to quantify the mRNA expression changes of ERVs. The first one is quantification using RNA-seq methods followed by analyzerepeat function of HOMER. In detail, RNA-seq raw reads was mapped the mouse genome mm9 with TopHat tools. The resulting bam files are converted to bed files. Make tag directory and analyzerepeat function are then used to generate the read count within each repeat region. The read counts are normalized to 10 million mapped reads. The resulting matrix loaded in R; differentially expressed ERVs are calculated using package limma. A cutoff of pvalue < 0.05 is used to find the significantly changed ERVs. The genome coordinate .bed or .gtf files for specific repfamily such as MMERVK10C-int and IAPEz-int are downloaded from UCSC table browser. To unbiasely identify the significantly changed ERVs across the genome, .gtf files for LTR class of repeats from chromosome 1 was downloaded from UCSC table browser. Significantly changed ERV name and locus are identified according to the above method. The nearby coding genes near the up or regulated ERVs are annotated using HOMER.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The ChIP-seq data including the fastq data and processed wig files have been uploaded into the GEO website. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE148750. Reviewers please use the token to access the data: cpixskietvsrhmt. All the code will be made available before the publication of this manuscript, including but not limited to the following, identification of the the differentially expressed genes, identification of significantly increased or decreased peaks for H3K27me3 and H3K9me3 enrichment; calculation of H3K9me3 at ERV site, calculation of ERV mRNA expression, motif analysis, etc.

The publicly available dataset is also used, UCSC genome browser. Bloodspot (http://servers.binf.ku.dk/bloodspot/).

The ChIP-seq data including the fastq data and processed wig files have been uploaded into the GEO website. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE148750. Reviewers please use the token to access the data: cpixskietvsrhmt. All the code will be made available before the publication of this manuscript, including but not limited to the following, identification of the differentially expressed genes, identification of significantly increased or decreased peaks for H3K27me3 and H3K9me3 enrichment; calculation of H3K9me3 at ERV site, calculation of ERV mRNA expression, motif analysis, etc. Please refer to Supplementary Table 10 metadata spreadsheet for detailed information.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical method was used to predetermine sample size for a given effect size. The differences in the phenotypic analysis of experimental parameters were statistically compared using a 2-tailed t test. P value of less than 0.05 was considered to be statistically significant; *P < 0.05, **P < 0.01, and ***P < 0.001. Experiments were repeated at least 2–3 times. The variance was similar between the groups being statistically compared. If higher variability was found, we increased the sample size to fully confirm statistical significance. Sample size was chosen based on similar studies using animal models to help dissect the role of hematopoietic stem cells.
Data exclusions	We included all tested animals for quantification to analyze statistical difference.
Replication	Experiments were repeated at least 2–3 times, both for in vivo and in-vitro HSPC quantifications. For in vivo experiments, we collected results at various time points, week 2, week 3 and week 8 post H3.3 deletion. The number of biological replicates are presented in the data plot as individual dots. Except for ChIP-seq for H3K4me3, H3K27me3 and H3.3 using a limited number of primary cells after in vitro culture of H3.3DKO HSCs and adding back H3.3 transgene, all the rest of experiments were sucessfully replicated each time using different biological replicates.
Randomization	Age and gender matched control or H3.3DKO mice are used to compare the phenotypes. Animals were injected with tamoxifen around 6 weeks of age. For in vitro studies, primary HSCs were isolated from control or knockout animals and were subjected to in vitro culture.
Blinding	We conduct experiment with the information regarding the number of the animals, without knowing their genotypes. But due to the severe phenotypes of enlarged spleen, RBC reduction, and anemia phenotype etc, it would become apparent which mice we are analyzing once we dissected the mouse organs. For in vitro cultures, the wells were labeled with the animal numbers. The operators were blinded to the genotype of each mouse during the course of experiment, including cell culture, cell counting, collecting cells for flow cytometry and sorting.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☐	☒ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

Name Company Clone Applications Catalog number Dilutions
CD45.1 Biolegend A20 Flow Cytometry 110707 50.00
CD45.2 Biolegend 104 Flow Cytometry 109843 50.00
Ter119 Biolegend Ter119 Flow Cytometry 116211 50.00
CD3 Biolegend 17A2 Flow Cytometry 100227 50.00
B220 Biolegend RA3-6B2 Flow Cytometry 103224 50.00
CD11b Biolegend M1/70 Flow Cytometry 101207 50.00
Gr1 Biolegend RB6-85C Flow Cytometry 108416 50.00
c-Kit Biolegend 2B8 Flow Cytometry 105834 50.00
Sca1 Biolegend D7 Flow Cytometry 108114 50.00
H3.3 Millipore N/A Western Blot 09-838 1000.00
Beta-actin Sigma AC-74 Western Blot A5316 10000.00
Gapdh Cell Signaling Technologies 2118 Western Blot 2118 10000.00
Ki67 BD Biosciences B56 Flow Cytometry 558616 15.00
CD16/32 Biolegend 93 Flow Cytometry 101335 50.00
CD150 Biolegend TC15-12F12.2 Flow Cytometry 115904 50.00
CD48 Biolegend HM48-1 Flow Cytometry 103426 50.00
CD34 BD Biosciences RAM34 Flow Cytometry 553733 50.00
CD16/32 Biolegend 93 Flow Cytometry 101328 50.00
H3K4me3 Cell Signaling Technologies 9727S Chip 9727S 1 µg/reaction
H3K9me3 Active Motif 39161 Chip 39161 1 µg/reaction
H3K27me3 Diagnode C15410069 Chip C15410069 1 µg/reaction
H3K27ac Abcam ab4729 Chip ab4729 1 µg/reaction

HA Bon Opus Biosciences (BA020061) 12CA5 Chip, western blot 1 ug/reaction
CD71 Biolegend RI7217 Flow cytometry 113808 50.00
CD19 eBiosciences eBio1D3 Flow cytometry 14-0193-82 50.00
CD44 Biolegend IM7 Flow cytometry 103022 50.00
CD25 Biolegend PC61 Flow cytometry 102006 50.00
Flt3 BD Biosciences A2F10.1 Flow Cytometry 560718 50.00
Il7r BD Biosciences A7R34 Flow Cytometry 566377 50.00
BUV395 BD Biosciences 564176 Flow Cytometry 564176 100.00
CD105 Biolegend MJ7/18 Flow Cytometry 120414 50.00
H3 Abcam ab1791 Western Blot ab1791 1:10000

Validation

For the flow cytometry antibodies, all of those antibodies have been used by our group in the previous publications or used by other groups studying mouse hematopoietic stem cell and lineage differentiation. The cell population and gating generating from these antibodies is comparable to what have already been published.

For the ChIP-seq antibody including H3K4me3, H3K27me3, H3K9me3 or H3K27ac antibodies, previous articles have used the same clone of antibodies in a native ChIP-seq experiment using small number of mouse embryos. Our results generated ChIP-seq peaks comparable to ChIP-seq using bulk hematopoietic cells, suggesting the robustness of our antibodies and native ChIP-seq approaches. For the HA antibodies, we have generated results to indicate the binding regions of H3.3. This patterns matches well with the previous published enrichment patterns of H3.3 in mouse embryonic cells. All the ChIP-seq results were compared with control samples with no IgG incubations, demonstrating the specificity of peaks associated with each antibody.

The western blot antibody for H3.3 from Millipore is used previous by other research groups. We have used our H3.3 knockout cells as a negative control and the H3.3 antibody does not recognize H3.3 null cells, but shows a band with appropriate size in wild type cells.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals

H3.3A floxed allele was generated by Dr. Laura Banaszynski (Banaszynski et al, 2013) H3.3B EYFP/mcherry deletion allele and mice were generated by Dr. Duancheng Wen. The parental ESC line (129B-13) to generate both H3.3A and H3.3B knock out mice was derived from F1 hybrid embryos crossed with 129SV and C57BL/6j by Dr. Duancheng Wen. Pluripotency of the ES cell line was assessed by both chimera generation and tetraploid complementation. by Dr. Duancheng Wen. The CD45.1 NSG mice used for bone marrow transplantation were purchased from Jax lab (strain 005557). Rosa26creERT2 mice were purchased from the Jax lab (https://www.jax.org/strain/008463). Rosa26creERT2; H3.3Afl/fl mice, H3.3B-/- breeder and offsprings were maintained at B6/129 background and backcrossed to C57BL/6j for four to five times.

We use sex and age matched adult mice for analysis of their hematopoietic phenotypes, from 6-16 weeks old at the time of tamoxifen injection.

Wild animals

No wild animals were used

Field-collected samples

No field-collected samples were used

Ethics oversight

All breeding of animals and animal experiments were performed under the approval of Weill Cornell Medicine Institutional Animal Care and Use Committee, New York, NY. To compare the phenotypes between different mouse genotypes, sex- and weight-matched control or H3.3 double knockout mice were used. Mice at 20–28 g body weight were used.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

ChIP-seq

Data deposition

☒ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links
May remain private before publication.

Our raw data are currently being uploaded and waiting the assignment of the GEO number.

Files in database submission

in vitro_DKO_H3K4me3.wig
in vitro_DKO_H3K9me3.wig
in vitro_DKO_H3K27ac.wig
in vitro_DKO_H3K27me3.wig
in vitro_H3.3.wig
in vitro_WT_H3K4me3.wig
in vitro_WT_H3K9me3.wig
in vitro_WT_H3K27ac.wig

in vitro_WT_H3K27me3.wig
in vivo_d14_BKO_H3K4me3.wig
in vivo_d14_BKO_H3K9me3.wig
in vivo_d14_BKO_H3K27ac.wig
in vivo_d14_BKO_H3K27me3.wig
in vivo_d14_BKO_LK_H3K9me3.wig
in vivo_d14_DKO_H3K4me3.wig
in vivo_d14_DKO_H3K9me3.wig
in vivo_d14_DKO_H3K27ac.wig
in vivo_d14_DKO_H3K27me3.wig
in vivo_d14_DKO_LK_H3K9me3.wig
in vivo_d17_BKO_H3K9me3.wig
in vivo_d17_BKO_H3K27me3.wig
in vivo_d17_DKO_H3K9me3.wig
in vivo_d17_DKO_H3K27me3.wig
in vivo_d56_BKO_H3K4me3.wig
in vivo_d56_BKO_H3K9me3.wig
in vivo_d56_BKO_H3K27ac.wig
in vivo_d56_BKO_H3K27me3.wig
in vivo_d56_DKO_H3K4me3.wig
in vivo_d56_DKO_H3K9me3.wig
in vivo_d56_DKO_H3K27ac.wig
in vivo_d56_DKO_H3K27me3.wig
in vitro_DKO_H3K4me3.fastq.gz
in vitro_DKO_H3K9me3.fastq.gz
in vitro_DKO_H3K27ac.fastq.gz
in vitro_DKO_H3K27me3.fastq.gz
in vitro_DKO_input.fastq.gz
in vitro_WT_H3K4me3.fastq.gz
in vitro_WT_H3K9me3.fastq.gz
in vitro_WT_H3K27ac.fastq.gz
in vitro_WT_H3K27me3.fastq.gz
in vitro_WT_input.fastq.gz
in vivo d14_BKO_input.fastq.gz
in vivo d14_DKO_input .fastq.gz
in vivo d56_BKO_input_R1_001.fastq.gz
in vivo d56_BKO_input_R2_001.fastq.gz
in vivo d56_DKO_input_R2_001.fastq.gz
in vivo d56_DKO_input.fastq.gz
in vivo_d14_BKO_H3K4me3.fastq.gz
in vivo_d14_BKO_H3K9me3.fastq.gz
in vivo_d14_BKO_H3K27ac.fastq.gz
in vivo_d14_BKO_H3K27me3.fastq.gz
in vivo_d14_DKO_H3K4me3.fastq.gz
in vivo_d14_DKO_H3K9me3.fastq.gz
in vivo_d14_DKO_H3K27ac.fastq.gz
in vivo_d14_DKO_H3K27me3.fastq.gz
in vivo_d56_BKO_H3K4me3_R1_001.fastq.gz
in vivo_d56_BKO_H3K4me3_R2_001.fastq.gz
in vivo_d56_BKO_H3K9me3_R1_001.fastq.gz
in vivo_d56_BKO_H3K9me3_R2_001.fastq.gz
in vivo_d56_BKO_H3K27ac_R1_001.fastq.gz
in vivo_d56_BKO_H3K27ac_R2_001.fastq.gz
in vivo_d56_BKO_H3K27me3_R1_001.fastq.gz
in vivo_d56_BKO_H3K27me3_R2_001.fastq.gz
in vivo_d56_DKO_H3K4me3_R1_001.fastq.gz
in vivo_d56_DKO_H3K4me3_R2_001.fastq.gz
in vivo_d56_DKO_H3K9me3_R1_001.fastq.gz
in vivo_d56_DKO_H3K9me3_R2_001.fastq.gz
in vivo_d56_DKO_H3K27ac_R1_001.fastq.gz
in vivo_d56_DKO_H3K27ac_R2_001.fastq.gz
in vivo_d56_DKO_H3K27me3_R1_001.fastq.gz
in vivo_d56_DKO_H3K27me3_R2_001.fastq.gz

Genome browser session
(e.g. UCSC)

N/A

Methodology

Replicates

For H3.3 ChIP-seq in LKS cells, 1 biological replicate is performed. One replicate is on 2 millions of E4-HUVEC cocultured and purified LKS cells.

For H3K4me3 enrichment within LKS from control or H3.3DKO cells, 3 biological replicated are performed, including LKS cells at day 4-post H3.3 deletion in vitro, day 14, day 56 post H3.3 deletion in vivo.

	For H3K27me3 enrichment within LKS cells from control or H3.3DKO cells, 3 biological replicates are performed, including LKS cells at day 4 post H3.3 in deletion in vitro, day 14 and day 56 post H3.3 deletion in vivo. For H3K9me3 enrichment within LKS cells from control or H3.3DKO cells, 4 biological replicates are performed, including LKS cells at day 4-post H3.3 in deletion in vitro, day 14, day17 or day 56 post H3.3 deletion in vivo. For H3K27ac enrichment within LKS for control or H3.3DKO cells, 3 biological replicates are performed, including LKS cells at day 4-post H3.3 in deletion in vitro, day 14, day 56 post H3.3 deletion in vivo. For or H3K9me3 of cKit+Lin-Sca1- cells for BKO or H3.3DKO cells, 1 biological replicates are performed at day 17 post H3.3 deletion in vivo. For H3K27me3 on Hirafl/fl or HiraKO LKS cells at day 14 post Hira deletion, 1 biological replicate are performed.
Sequencing depth	All ChIP-seq libraries are sequenced using Illumina Hiseq 4000 platform, single read, 50 cycles, except for the ChIP-seq for day 56 post H3.3 deletion H3K27me3 and H3K9me3, which are sequenced using Novaseq 6000, pair-end, 50 cyles. The two platforms generated comparable peaks, as shown in Figure 3 and Figure 6.
Antibodies	H3K4me3 Cell Signaling Technologies 9727S ChIP H3K9me3 Active Motif 39161 ChIP H3K27me3 Diagnode C15410069 ChIP H3K27ac Abcam ab4729 ChIP HA Bon Opus Biosciences (BA020061) 12CA5 Chip, western blot
Peak calling parameters	Raw reads (fastq files) from ChIP-seq was aligned to mouse genome mm9 using bwa. Duplicate reads and reads aligning to more than one location were excluded. After QC and alignment, bam files were obtained. Bam files were then converted to bed files using bamtobed command (bedtools 2.27.1). SICER was used to identify the peaks for H3.3 (HA) by comparing to the input control sample using a FDR < 0.01, fragment size of 150bp, window size 200bp, gap size 600bp. For identification of significantly decreased or increased H3K9me3, H3K27me3, H3K27ac peak regions, bed files for BKO_H3K9me3, BKO_input, DKO_H3K9me3 and DKO_input were used. The parameters used are FDR < 0.01, fragment size of 150bp, window size 200bp, and gap size of 600bp. For H3K4me3 peak calling, gap size 200 bp is used.
Data quality	After the ChIP-seq results and peak calling, we routinely recover about 20K-40K peak regions, using the parameter FDP < 0.01. The H3K4me3, H3K27me3, H3K9me3, H3K27ac and H3.3 peaks are comparable to what is previously reported for their genomewide localizations.
Software	bwa, bedtools, SICER V1.1.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	At the time points after tamoxifen inejctions, mice were euthanized. Femurs and tibias were dissected. Following muscle cleanup, 2 femurs and 2 tibias per mouse were ground using mortars and pestles in 5 ml MACS buffer (PBS+2mM EDTA+0.5%BSA +1% antibiotics). The process is repeated two more times until all the bone chunks are broken and appear white. The cells were then filtered through a 40µm cell strainer. The total number of filtered cells was counted as the total hematopoietic cells (excluding mature RBC) in the BM. Lineage depletion was carried out for the hematopoietic cells using direct lineage depletion kit (Miltenyi, 130-110-470). 50µl of lineage antibody cocktail was added to the cells obtained from 2 femur and 2 tibius. The numbers of lineage negative cells (Lin-) were recorded. Lin- cells were blocked with Rat IgGs (not CD16/32!!!) for 5 min on ice, and then stained with HSC and progenitor cell markers c-Kit, Sca1, CD150, CD48, CD34, CD16/32 and CD105 to distinguish the LT-HSC and GMPs, CMPs, and MEPs, preGM, CFU-E and pre-megaE, etc. The total number of HSC per 2 femurs and 2 tibias was calculated using the total number of Lin- cells times the frequencies of CD150+CD48-c-Kit+Sca1+ gating. Likewise, the total number of GMPs or any other progenitor cell type was calculated using the total number of Lin-cKit+ cells and the frequencies of GMP cells. To obtain the total number of spleen cells, the weight of whole spleen was recorded. A small fraction of spleen was cut out and weighted. By counting the total cell number of the sampled spleen, we can infer the total number of mononuclear cells within spleen. For quantification of BM and spleen lineage cell types, the cells were blocked with mouse FcR block for 5 min on ice. Antibody cocktails 1 (Ter119, CD11b, Gr1, B220) was stained. The percentage of myeloid cells and B cells were obtained. The total numbers of lineage cell types are quantified by multiplying the total cell number within spleen and percentage of cell type among total live cells.

Instrument	The quantification of hematopoietic stem cell number and progenitor cell numbers in the bone marrow, the quantification of the lineage composition of bone marrow, spleen are carried out using BD LSRII or BD Fortessa or BD FACS AriaII machine, using a software named BD Diva.
Software	BD Diva is used to record the data. Subsequent gating and data analysis is carried out in FlowJo, version 10.4.
Cell population abundance	For quantification of LT-HSCs, we record at least 1 million events for the total Lin- cells, to allow for sufficient cell number for LT-HSCs; For quantification of lineage composition in spleen and bone marrow, more than 0.1 million of total cells are recorded, to allow for sufficient cell number within each populations.
Gating strategy	For the gating of CD150+CD48- cells on top of cKit+Lin-Sca1+ cells, we used CD150 (fluorescence minus one) FMO and CD48 FMO in the very first few batches of experiments. Later, we gated on CD150+CD48-LKS cells empirically, using the same gating for control and H3.3DKO cells. For the gating of CD16/32+ cells within the cKit+Lin- cell populations, CD16/32 FMO was used. For the gating of CD11b+Gr1+, or B220+ cells within spleen or bone marrow, unstained spleen or bone marrow samples are used as negative control.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.