nature research

Corresponding author(s): Sean J. Morrison

☐ Initial submission ☐ Revised version ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

Describe how sample size was determined.

The sample size used in each experiment was not predetermined or formally justified for statistical power. For samples with low variation, for example metabolite levels in HSCs, generally 4-6 samples per condition were analysed in each experiment. For samples with large variation, for example transplantation experiments, generally 10-20 samples per condition were analysed in each experiment.

2. Data exclusions

Describe any data exclusions.

In Figure 5f, data from 1 human patient was excluded due to clinical evidence of myelodysplastic syndrome. In Figures 3e and ED Figure 8g, some recipient mice were excluded because they showed no reconstitution, suggesting no HSCs were present among the donor cells, therefore those samples provided no information on HSC activity. In ED Figure 6d, time point 4 was excluded from the statistical analysis because it involved comparisons between genotypes before Tet2 conditional deletion.

3. Replication

Describe whether the experimental findings were reliably reproduced.

The experimental findings were reproduced in multiple independent experiments. The number of independent experiments and biological replicates in each data panel is indicated in the figure legends. Data shown in the figures represent the aggregate of all independent experiments in almost all cases. Data shown in a minority of panels are from a representative experiment (e.g. western blots) and in those cases the number of independent experiments that reproduced the finding is also indicated in the figure legends.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

No formal randomization techniques were used, however, samples were allocated randomly into experimental groups and processed in an arbitrary order.

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

The investigators were not blinded to group allocation during data collection and analysis

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ A statement indicating how many times each experiment was replicated
☐	☒ The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒ A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☒	☐ The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
☐	☒ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒ Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.

GraphPad Prism v7.0 and R 3.2.1 were used for all statistical tests. RNA-seq analysis was performed using FastQC 0.11.2 (raw reads quality checking), Trimmomatic 0.32 (raw reads quality filtering), TopHat 2.0.12 with Bowtie2 2.2.3 (reads mapping), SAMtools 0.1.19 (mapped reads quality filtering), HTSeq 0.6.1 (mapped read counts quantification), DESeq2 1.8.2 and RUVSeq 1.2.0 with R 3.2.1 (differential expression), and Cluster 3.0 (clustering and heatmap generation). Multiquant, MetaboAnalyst and Simca were used to analyze metabolomics data.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

There are no restrictions on the availability of materials. No unique materials were used in this study. Antibodies and chemical reagents were purchased from for-profit companies as detailed in the methods.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

Please see attached list of antibodies, clone names, suppliers, catalogue numbers and lot numbers. Please see statement and citations on suppliers' websites regarding antibody specificity. Antibodies used to purify haematopoietic stem cells have been thoroughly validated in previous publications by this laboratory (e.g. Cell 121:1109, Blood 107:924, Nature 526:126).

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

No cell lines were used

b. Describe the method of cell line authentication used.

No cell lines were used

c. Report whether the cell lines were tested for mycoplasma contamination.

No cell lines were used

d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.

No cell lines were used

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.

C57BL/6 mice were used. Details and references for each mutant strain are described in the methods section. Both male and female mice were used for most experiments unless otherwise specified. 2-6 month old animals were used for most experiments as specified in the text.

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

Bone marrow aspirates were collected from male and female patients, aged 34-85, who were being assessed for lymphoma. Experiments in Figure 5f-g were performed using aspirates in which there was no evidence of lymphoma or myelodysplastic syndrome was observed. Age and sex of the patients were shared with us. We did not obtain additional demographic data on the patients who donated the samples because per our IRB approval, the samples were de-identified before being shared with us.

nature research

Corresponding author(s): Sean J. Morrison

☐ Initial submission ☐ Revised version ☒ Final submission

Flow Cytometry Reporting Summary

Form fields will expand as needed. Please do not leave fields blank.

Data presentation

For all flow cytometry data, confirm that:

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

☒ 2. 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).

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

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

Methodological details

5. Describe the sample preparation.	Bone marrow cells were obtained by flushing femurs and tibias with a 25G needle or by crushing femurs, tibias, vertebrae, and pelvic bones with a mortar and pestle, in Ca2+ and Mg2+ free Hank's buffered salt solution (HBSS; Gibco), supplemented with 2% heat-inactivated bovine serum (HIBS, Gibco). Spleens and thymuses were dissociated by crushing followed by gentle trituration. All cell suspensions were filtered through a 40 µm cell strainer. For flow cytometric analysis and isolation, cells were incubated with combinations of fluorescent antibodies to cell-surface markers. Antibody staining was generally performed at 4°C for 30 minutes. For isolation of all c-kit+ cell populations, cells were pre-enriched before flow cytometry using paramagnetic microbeads and an autoMACS magnetic separator (Miltenyi Biotec).
6. Identify the instrument used for data collection.	BD FACS Aria II or FACS Aria Fusion (for cell sorting or analysis), BD Canto (for analysis).
7. Describe the software used to collect and analyze the flow cytometry data.	BD FACS DiVa
8. Describe the abundance of the relevant cell populations within post-sort fractions.	The abundance of the relevant cell populations within post-sort fractions was almost 100% in pilot experiments. The abundance could not be tested after every experiment because cells were usually sorted directly in lysis buffer.
9. Describe the gating strategy used.	Please see the attached supplementary figure showing the gating strategy for haematopoietic stem and progenitor cells

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