nature research

Corresponding author(s): Qing Li

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.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. 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
☐	☒	An indication of 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 statistics 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
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	all FACS data were collected with FACSdiva V8.0
Data analysis	all FACS data were analyzed with Flowjo X on Windows and Flowjo 9.7 on Mac. All graphs were generated with Prism 7 Graphpad.

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 upon request. 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

Gene expression data reanalyzed here (DNAJB8 levels) are available from GEO under accession code GSE45194. Source data have been provided as Supplementary Table 2. All other data supporting the findings of this study are available from the corresponding author on reasonable request."

Field-specific reporting

Please select 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/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

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

Sample size	The sample size used in each experiment was not predetermined or formally justified for statistical power. For samples with low variation such as HSC frequency, generally 3-6 individual pairs were analyzed in 3-6 independent experiments. For samples with high variation such as transplantation assays, generally 5-15 recipients were analyzed.
Data exclusions	Data was excluded if FACS staining is deemed not satisfactory.
Replication	The experimental findings were reproduced in multiple independent experiments. Each biological replicates was analyzed in separate independent experiments. Data shown in the figures represent the aggregate of all independent experiments, with mean as the center and standard deviation to define error bars.
Randomization	No formal randomization techniques were used.
Blinding	The investigators were not blinded to group allocation during data collection and analysis

Reporting for specific materials, systems and methods

Materials & experimental systems

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

Methods

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

Antibodies

Antibodies used	Please see Supplementary Table 1 for a list of antibodies, clone names, suppliers, catalogue numbers. Please see statement and citations on suppliers' websites regarding antibody specificity.
Validation	Antibodies used to purify haematopoietic stem cells have been thoroughly validated in previous publications by this laboratory (e.g. Nature 504(7478): 143-147). Each new lot of antibodies are tested with samples with know staining pattern.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	The AML cell lines OCI-AML2 (Cat# ACC99) and OCI-AML3 (Cat# ACC582) were purchased directly from DSMZ (Germany). 293T (Cat# CRL-3216) was purchased from ATCC.
Authentication	Used directly from source; morphologic examination; validated overexpression data with western blot
Mycoplasma contamination	No, Mycoplasma was monitored by MycoAlert™ Mycoplasma Detection Kit from Lonza
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines used.

Animals and other organisms

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

Laboratory animals	C57BL/6 mice were used. Loxp-STOP-Loxp (LSL)-NrasG12D/+ 42, Tet2+/- 43,Stat5ab+/-31, Socs2+/- 32, Col1α1-H2B-GFP; Rosa26-M2-rtTA28, Mx1-Cre (Jackson, 003556), and the ERAI (Riken biological resource RBRC01099) mice were maintained in C57BL6 background. The B6;129S4-Ern1<tm2.1Tiw> (Ire1α knockout, RBRC05515) strain was obtained from Riken biological resource and backcrossed to a C57BL/Ka-CD45.2: Thy-1.1 background for at least six generations. Recipients in reconstitution assays were adult C57BL/Ka-CD45.1: Thy-1.2 mice, at least 8 weeks of age at the time of irradiation. All analyses were performed on samples from 8-14 week-old age- and sex-matched mice. We performed analysis separately for male and female mice and did not observe any differences between genders. Results represent analyses using both male and female mice. Equal number of male and female mice were used for analyses as much as possible.
Wild animals	none was used
Field-collected samples	none was used

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	Bone marrow cells were flushed from the long bones (tibias and femurs) with FACS Buffer (Hank's buffered salt solution without calcium or magnesium, supplemented with 2% heat-inactivated calf serum) from the mice. Cells were triturated and filtered through a nylon screen (60 mm; Sefar America) to obtain a single-cell suspension. Haematopoietic cells including CD150+CD48-Lineage-Sca1+c-kit+ SLAM HSCs and CD150-CD48-Lineage-Sca1+c-kit+ SLAM MPPs were isolated as previously described (e.g. Nature 504(7478): 143-147). For isolation of HSCs, whole bone marrow cells were incubated with antibodies to lineage (Lin) markers including B220 (6B2), CD2(RM2.5),CD3 (KT31.1), CD5 (53-7.3), CD8 (53-6.7), Gr-1 (8C5), CD41 (MWReg30) and Ter119 (Ter-119), anti-CD150 antibody (TC15-12F12.2), anti-CD48 antibody (HM48-1),anti-c-kit antibody (2B8) conjugated to APC, and anti-Sca1 antibody (D7) (all antibodies were purchased from Biolegend). After washing, cells were incubated with anti-APC conjugated to paramagnetic microbeads (Miltenyi Biotec). The microbead bound (c-kit+) cells were then enriched using MS columns (Miltenyi Biotec). Non-viable cells were excluded from sorts and analyses using the viability dye 4',6-diamidino-2-phenylindole (DAPI,Sigma) (1μg/ml). Cells were analyzed with BD Fortessa or sorted by BD Aria flow cytometry. The other haematopoietic cell populations were defined as the following: LSK (lineage-Sca1+cKit+), CMP (common myeloid progenitor; CD16/32-CD34+Lin-Sca1-c-kit+), GMP (Granulocyte-macrophage progenitor; CD16/32+CD34+Lin-Sca1-c-kit+), and MEP (megakaryocyte-erythroid progenitor; CD16/32-CD34-Lin-Sca1-c-kit+). Additional information of antibodies is provided in Supplementary Table 1.
Instrument	FACS analysis was done with BD Fortessa. FACS sorting was done with BD Aria III.
Software	all FACS data were collected by FACSdiva V8.0, and analyzed with Flowjo X on Windows or Flowjo 9.7 on Mac.
Cell population abundance	the frequency of functional long term HSC is ~50% within CD150+CD48-Lineage-Sca1+cKit+ fraction based on previous studies with single cell transplantation. (e.g. Nature 504(7478): 143-147.). The frequency of SLAM HSC, based on phenotypic characterization, is roughly <0.01% of the bone marrow cells.
Gating strategy	Please see Supplementary Figure 1a for details of gating strategy.

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