

Reporting Summary

Nature Portfolio 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 Portfolio policies, see our [Editorial Policies](#) 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
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| ☐ | ☒ | 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
<i>Give P values as exact values whenever suitable.</i> |
| ☒ | ☐ | 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	No software was used for data collection; data were already collected, de-identified and provided to researchers for use.
Data analysis	R (v3.6.0), Prism (v9), and FlowJo (v10.9.0), Seurat (v3.0), DoubletFinder (v3), SingleCellNet (v0.1.0), CellRouter (available via the Framework for Unified Single-Cell Analysis (FUSCA) v1.3.1), cNMF (v1.2), PBA (no formal versions curated), StationaryOT (no formal versions curated), survminer (v0.4.9).

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 and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The raw sequencing data from this study is deposited within the National Institutes of Health database of Genotypes and Phenotypes (NIH dbGAP; accession

phs002750). Processed sequencing data is available from the National Center for Biotechnology Information's Gene Expression Omnibus (NCBI-GEO; accession GSE189161).

Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

Reporting on sex and gender	Our results include both sexes - there was no bias in sample selection. All are analyzed together.
Population characteristics	We used secondary healthy donor human HSPC specimens collected on banking protocols. There were no identifying or demographic information (except age and sex) collected.
Recruitment	These were secondary specimens collected on healthy donor banking protocols.
Ethics oversight	Informed consent for all studies was obtained in accordance with the declaration of Helsinki. Human studies were approved and regulated by either the Human Subjects Protection Committees of the Dana-Farber/Harvard Cancer Center, or the Institutional Research Board of Boston Children's Hospital, Brigham and Women's Hospital, or The University of Pennsylvania.

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

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](https://www.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	The main dataset surveying human hematopoiesis was assembled based on specimen availability and we described methods used for statistical analysis. For validating experiments, number of replicates are stated
Data exclusions	No data were excluded.
Replication	Replicates were typically performed three times or as described and attempts at replication were successful.
Randomization	No randomization was required. To control for donor variability, experiments comparing control and experimental conditions used the same donor, although multiple donors were tested independently.
Blinding	Blinding was used for experiments with morphologic analysis - colony formation assays and clonogenesis assays. Blinding was not done during flow cytometry as it was not relevant given that gating was kept consistent across sample and experiments. Blinding was not applicable to informatics analysis.

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 and archaeology
☐	☒ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

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

Antibodies

Antibodies used

anti-mouse CD45.1-FITC (BD 553775, lot 1284545 clone A20)
 anti-human CD45-APC (BD, 555485, lot 8338550, clone HI30)
 anti-human CD41a-FITC (BD 340929, lot 330674, clone HIP8)
 anti-human CD34-PECy7 (BD 348791, lot 3157098, clone 8G12)
 anti-human CD235a-PECy7 (BioLegend 349112, lot 0216883, clone HI264)
 anti-human CD19-PE (BioLegend 392506, lot B369430, clone 4G7)
 anti-human CD33-APCCy7 (BioLegend, 366614, lot, B252646, clone P67.6)
 anti-human CD69-APC (BD, 555533, lot 0010330, clone FN50)
 anti-human CD38-PE (BioLegend 303506, lot B329920, clone HIT2)
 anti-human lineage cocktail (CD3/14/16/19/20/56 BioLegend 348805 lot B240424)-Pacific Blue
 anti-human CD16-BV510 (BioLegend 302048, lot B265264, clone 3G8)
 anti-human CD14-APC (BD 561383, lot 2213036, clone MSE5)
 anti-human CD45-APCCy7 (BioLegend 304014, lot B402225, clone HI30)
 anti-human CD16-Pacific Blue (BioLegend 302024, clone 3G8)

anti-mouse CD34-FITC (eBioScience 11-0341-85, lot 2527366, clone RAM34)
 anti-mouse CD127-PE/Dazzle 594 (BioLegend 135031, lot B376378, clone A7R34)
 anti-mouse CD135-PE (BioLegend 135306, lot B385169, clone A2F10)
 anti-mouse Sca-1-PE/Cy7 (BioLegend 108114, lot B395526, clone D7)
 anti-mouse CD16/CD32-PerCP/Cy5.5 (eBioScience 45-0161-82, lot 2608880, clone 93)
 anti-mouse CD150-APC (BioLegend 115910, lot B317517, clone TC15-12F12.2)
 anti-mouse CD69-Alexa700 (BioLegend 104539, lot B360229, clone H1.2F3)
 anti-mouse CD117-APC/Cy7 (BioLegend 105826, lot B405703, clone 2B8)
 anti-mouse/human CD45R/B220-BV421 (BioLegend 103251, lot B361300, clone RA3-6B2)
 anti-mouse CD3-BV421 (BioLegend 100228, lot B396106, clone 17A2)
 anti-mouse Ly-6G/Ly-6C-BV421 (BioLegend 108434, lot B382772, clone RB6-8C5)
 anti-mouse TER-119-BV421 (BioLegend 116234, lot B384031, clone TER-119)
 anti-mouse CD48-BV711 (BioLegend 103439, lot B388165, clone HM48-1)

Validation

anti-mouse CD45.1-FITC (BD 553775, lot 1284545 clone A20) - validated on unengrafted negative control mice to confirm specificity
 anti-human CD45-APC (BD, 555485, lot 8338550, clone HI30) - validated on human peripheral blood mononuclear cells
 anti-human CD41a-FITC (BD 340929, lot 330674, clone HIP8) - validated on human peripheral blood mononuclear cells
 anti-human CD34-PECy7 (BD 348791, lot 3157098, clone 8G12) - validated on CD34-enriched human hematopoietic stem and progenitor cells
 anti-human CD235a-PECy7 (BioLegend 349112, lot 0216883, clone HI264) - validated on human peripheral blood mononuclear cells
 anti-human CD19-PE (BioLegend 392506, lot B369430, clone 4G7) - validated on human peripheral blood mononuclear cells
 anti-human CD33-APCCy7 (BioLegend, 366614, lot, B252646, clone P67.6) - validated on human peripheral blood mononuclear cells
 anti-human CD69-APC (BD, 555533, lot 0010330, clone FN50) - validated on activated T cells (PHA stimulated) from human peripheral blood
 anti-human CD38-PE (BioLegend 303506, lot B329920, clone HIT2) - validated on purified human hematopoietic stem and progenitor cells
 anti-human lineage cocktail (CD3/14/16/19/20/56 BioLegend 348805 lot B240424)-Pacific Blue - validated on human peripheral blood mononuclear cells
 anti-human CD16-BV510 (BioLegend 302048, lot B265264, clone 3G8) - validated on human peripheral blood mononuclear cells
 anti-human CD14-APC (BD 561383, lot 2213036, clone MSE5) - validated on human peripheral blood mononuclear cells and enriched monocytes/macrophages
 anti-human CD45-APCCy7 (BioLegend 304014, lot B402225, clone HI30) - validated on human peripheral blood mononuclear cells
 anti-human CD16-Pacific Blue (BioLegend 302024, clone 3G8) - validated on human peripheral blood mononuclear cells

anti-mouse CD34-FITC (eBioScience 11-0341-85, lot 2527366, clone RAM34) - validated in fluorescence minus one (FMO) studies on mouse bone marrow.
 anti-mouse CD127-PE/Dazzle 594 (BioLegend 135031, lot B376378, clone A7R34) - validated in FMO studies on mouse bone marrow.
 anti-mouse CD135-PE (BioLegend 135306, lot B385169, clone A2F10) - validated in FMO studies on mouse bone marrow.
 anti-mouse Sca-1-PE/Cy7 (BioLegend 108114, lot B395526, clone D7) - validated in FMO studies on mouse bone marrow.
 anti-mouse CD16/CD32-PerCP/Cy5.5 (eBioScience 45-0161-82, lot 2608880, clone 93) - validated in FMO studies on mouse bone marrow.
 anti-mouse CD150-APC (BioLegend 115910, lot B317517, clone TC15-12F12.2) - validated in FMO studies on mouse bone marrow.
 anti-mouse CD69-Alexa700 (BioLegend 104539, lot B360229, clone H1.2F3) - validated in FMO studies on mouse bone marrow.
 anti-mouse CD117-APC/Cy7 (BioLegend 105826, lot B405703, clone 2B8) - validated in FMO studies on mouse bone marrow.
 anti-mouse/human CD45R/B220-BV421 (BioLegend 103251, lot B361300, clone RA3-6B2) - validated in FMO studies on mouse bone marrow.
 anti-mouse CD3-BV421 (BioLegend 100228, lot B396106, clone 17A2) - validated in FMO studies on mouse bone marrow.
 anti-mouse Ly-6G/Ly-6C-BV421 (BioLegend 108434, lot B382772, clone RB6-8C5) - validated in FMO studies on mouse bone marrow.
 anti-mouse TER-119-BV421 (BioLegend 116234, lot B384031, clone TER-119) - validated in FMO studies on mouse bone marrow.
 anti-mouse CD48-BV711 (BioLegend 103439, lot B388165, clone HM48-1) - validated in FMO studies on mouse bone marrow.

Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

Cell line source(s)

MS5 cell lines were provided by Dr. John Dick (University of Toronto).

Authentication	The identity of this cell line is obvious by morphology and did not undergo molecular authentication.
Mycoplasma contamination	All cell lines in culture were tested every 14 days and prior to cryopreservation for Mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	N/A

Animals and other research organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research, and [Sex and Gender in Research](#)

Laboratory animals	NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ (NSG) mice were obtained from Jackson Laboratory (Stock 005557) and used for xenotransplantation experiments at 8 weeks of age. We used female mice as transplant recipients.
Wild animals	This study did not involve wild animals.
Reporting on sex	These animals were used only has human transplant recipients. All female mice were used as is standard in such assays.
Field-collected samples	This study did not involve field-collected samples
Ethics oversight	The Institutional Animal Care and Use Committee (IACUC) at Boston Children's Hospital approved the studies in this manuscript.

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

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	Fresh cells were immediately processed for staining by flow cytometry. In some experiments, cells were cultured in XVIVO medium (Lonza) with cytokines (SCF, FLT3L, TPO, IL-3, and IL-6) when indicated.
Instrument	Becton Dickinson (BD) LSRII Becton Dickinson (BD) LSR Fortessa Becton Dickinson (BD) FACS Aria
Software	FlowJo software v10.9.0
Cell population abundance	Post sort purity checking is shown in Extended Data Figure 8B. This was routinely done for experiments differentiating in vitro populations.
Gating strategy	Detailed gating strategy for HSPC sorting is shown in Extended Data Figure 2A - this is based on FSC-A/SSC-A to exclude debris, HSC-H/FSC-A and SSC-H/SSC-W to exclude doublets, and Sytox Blue/Pacific Blue negative to exclude viable cells.

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