nature portfolio

Corresponding author(s): Dengli Hong

Last updated by author(s): Sep 12, 2024

# 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<br>*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<br>*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 | BD FACSDiva(V8.0.3), Summit (V6.3.1.16945), Las X (V4.7), ZEN (V2.3) |
| Data analysis | ImageJ(V1.52p), OpenComet(V1.3.1), Imaris (V9.0.1), FlowJo(V10), Graphpad Pism (V9.5.0), Trim_Galore(V0.6.7), IGV(V2.7.0), Bowtie2(V2.3.5.1), Burrow-Wheeler Aligner(V0.7.17), ATAC-seq QC (V1.14.4), DESeq2 (V1.26.0), HOMER(V4.11), hisat2 (V2.2.1), Picard(V2.25.5), deepTools(V3.5.1), MACS2(V2.2.6), R package ChIPseeker(V1.22.1), clusterProfiler(V3.14.3), circlise(V0.4.8), complex Heatmap package(V2.2.0), STAR(V2.7), HTSeq-count (V0.13.5), R (V3.6.3), HALO(V3.6.4134). |

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 ATAC sequencing, RNA sequencing, R-loop Cut & Tag sequencing, b-ZIPs Cut & Tag sequencing and whole genome sequencing data reported in this paper have

been deposited at the Genome Sequence Archive in BIG Data Center, Beijing Institute of Genomics (BIG), Chinese Academy of Sciences, under the study accession number PRJCA026810, accessible link: http://ngdc.cncb.ac.cn/bioproject/browse/ PRJCA026810. The reference genome used is mm10, accessible link: https://hgdownload.soe.ucsc.edu/goldenPath/mm10/bigZips/mm10.fa.gz. The annotation dataset used in RNA sequencing analysis is gencode.vM25.annotation.gtf, accessible link: http://ftp.ebi.ac.uk/pub/databases/gencode/Gencode_mouse/release_M25/gencode.vM25.annotation.gtf.gz.

## Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

| | |
|---|---|
| Reporting on sex and gender | Our study was not related to sex and gender, so the sex and gender were not considered. Eleven foetal liver samples , 3 cord blood samples, 14 bone marrow plasma samples from begnin infants, 10 bone marrow plasma samples from leukaemic infants were used in this study. |
| Reporting on race, ethnicity, or other socially relevant groupings | Socailly relevant categorization variables were not used in my manuscript. |
| Population characteristics | The foetal liver samples were collected from foetus (7-week-old, n=1; 8-week-old, n=3; 9-week-old, n=1; 10-week-old, n=1; 11-week-old, n=1; 13-week-old, n=1; 15-week-old, n=1; 16-week-old, n=1; 19-week-old, n=1). Cord blood samples were collected from placenta of new born babies. The bone marrow plasma samples were collected from 0-12-month-old patients (benign group: 2-month-old, n=1; 3-month-old, n=2; 4-month-old, n=2; 5-month-old, n=3; 6-month-old, n=3; 8-month-old, n=1; 9-month-old, n=1, 12-mpmth-old, n=1; leukaemic group: 2-month-old, n=2; 3-month-old, n=1; 4-month-old, n=2; 6-month-old, n=1; 8-month-old, n=1; 9-month-old, n=2; 11-month-old, n=1). Available samples were used without any selection. |
| Recruitment | The foetal liver samples were collected randomly from fetus by induced abortion at different gastational age. The cord blood samples were collected randomly from the plancenta of new born babies. The bone marrow plasma samples were collected from 0-12 month old infants and then seperated into begnin and leukaemia group accroding to their diagnostic reports. |
| Ethics oversight | Medical Ethics Committee of Tongji Hospital, Tongji University School of Medicine, Shanghai, China (k-w-2010-010). Medical Ethics Committee of Shanghai Children's Medical Center, Shanghai Jiao Tong University School of Medicine, Shanghai, China (SCMCIRB-K2024163-1). |

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

# Life sciences study design

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

| | |
|---|---|
| Sample size | No sample size calculation was performed. Sample size was chosen based on the statistical requirements. The exact sample sizes are included in the figure legends. |
| Data exclusions | No data were excluded. |
| Replication | A detailed experimental protocol is provided for replcation by others.The comet assay, rh2a.x staining, flow-sorting, tissue immunofluorescence, hepatocytes culture, cell cycle analysis were replicated at least 2 experimenters with reproducible results. All the experiments were independently repeated at least three times with reproducible results. |
| Randomization | All the samples were allocated into experimental groups randomly. |
| Blinding | In our study, most experimental findings were related to comparative analysis. For comparative experiments, the groups were treated at different conditions, the investigators can not be blinded to the group allocation. However, most data collection and analysis were automatically completed by the software, there is low likelihood of investigator's bias in the final readout. For the experiment of leukaemic model, the investigators were blinded to the group allocation of FetuA ko and wild type control mice, but the data analysis were not blinded because we needed the genotype to identify different group. |

# Behavioural & social sciences study design

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

| | |
|---|---|
| Study description | *Briefly describe the study type including whether data are quantitative, qualitative, or mixed-methods (e.g. qualitative cross-sectional, quantitative experimental, mixed-methods case study).* |
| Research sample | *State the research sample (e.g. Harvard university undergraduates, villagers in rural India) and provide relevant demographic information (e.g. age, sex) and indicate whether the sample is representative. Provide a rationale for the study sample chosen. For studies involving existing datasets, please describe the dataset and source.* |
| Sampling strategy | *Describe the sampling procedure (e.g. random, snowball, stratified, convenience). Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient. For qualitative data, please indicate whether data saturation was considered, and what criteria were used to decide that no further sampling was needed.* |
| Data collection | *Provide details about the data collection procedure, including the instruments or devices used to record the data (e.g. pen and paper, computer, eye tracker, video or audio equipment) whether anyone was present besides the participant(s) and the researcher, and whether the researcher was blind to experimental condition and/or the study hypothesis during data collection.* |
| Timing | *Indicate the start and stop dates of data collection. If there is a gap between collection periods, state the dates for each sample cohort.* |
| Data exclusions | *If no data were excluded from the analyses, state so OR if data were excluded, provide the exact number of exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.* |
| Non-participation | *State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no participants dropped out/declined participation.* |
| Randomization | *If participants were not allocated into experimental groups, state so OR describe how participants were allocated to groups, and if allocation was not random, describe how covariates were controlled.* |

# Ecological, evolutionary & environmental sciences study design

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

| | |
|---|---|
| Study description | *Briefly describe the study. For quantitative data include treatment factors and interactions, design structure (e.g. factorial, nested, hierarchical), nature and number of experimental units and replicates.* |
| Research sample | *Describe the research sample (e.g. a group of tagged Passer domesticus, all Stenocereus thurberi within Organ Pipe Cactus National Monument), and provide a rationale for the sample choice. When relevant, describe the organism taxa, source, sex, age range and any manipulations. State what population the sample is meant to represent when applicable. For studies involving existing datasets, describe the data and its source.* |
| Sampling strategy | *Note the sampling procedure. Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.* |
| Data collection | *Describe the data collection procedure, including who recorded the data and how.* |
| Timing and spatial scale | *Indicate the start and stop dates of data collection, noting the frequency and periodicity of sampling and providing a rationale for these choices. If there is a gap between collection periods, state the dates for each sample cohort. Specify the spatial scale from which the data are taken* |
| Data exclusions | *If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.* |
| Reproducibility | *Describe the measures taken to verify the reproducibility of experimental findings. For each experiment, note whether any attempts to repeat the experiment failed OR state that all attempts to repeat the experiment were successful.* |
| Randomization | *Describe how samples/organisms/participants were allocated into groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.* |
| Blinding | *Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.* |

Did the study involve field work? ☐ Yes ☐ No

## Field work, collection and transport

| | |
|---|---|
| Field conditions | *Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).* |
| Location | *State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).* |
| Access & import/export | *Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).* |
| Disturbance | *Describe any disturbance caused by the study and how it was minimized.* |

# 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 |
| ☒ | ☐ Plants |

### Methods

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

## Antibodies

| | |
|---|---|
| Antibodies used | CD34 FITC(581) BD 555821, Biotin-labled lineage markers ThermoFisher 88-7774-75, Sca-1 PE-cy7(D7) ThermoFisher 25-5981-82, c-Kit APC (2B8) ThermoFisher 17-1171-82, CD150 PE(mShad150) ThermoFisher 12-1502-82, CD48 FITC(HM48-1) ThermoFisher 11-0481-82, Phosphor-histone H2A.X(ser139)(20E3) rabbit mAb cell signaling technology(CST) 9718s, Tlr4 antibody abcam ab13556, FetuinA antibody(EPR17839-163) abcam ab187051, p-RPA(phopho s33) abccam ab211877, Tlr4 antibody(UT41) ThermoFisher 53-9041-80, huamn FetuinA antibody(1F6B9) Proteintech 66094-1-Ig, Myd88(E11) Santa Cruz sc-74532, anti-mouse CD117(ACK2) ThermoFisher 14-1172-85, Ter119 APC Biolegend 116212, Gr-1 APC Biolegend 108412, Mac-1 APC Biolegend 101212, B220 APC Biolegend 103212, CD3 APC Biolegend 100236, CD150 BV421(SLAM) Biolegend 115925, CD41 APC(MWRag30) Biolegend 133913, c-Kit goat mAb R&D AF1356, E-cadherin rabbit monoclonal antibdoy(24E10) CST 3195T, anti-huamn albumin antibody(MAB1455) R&D 188835, laminin monoclonal antibody Abcam ab11575, Ki67-FITC(SolA15) ThermoFisher 11-5698-80, CD34 antibody(EP373Y) Abcam ab81289, anti-p-junb(Thr102/Thr104) CST 8053S, anti-p-fosl1(s265) CST 3880S, anti-p-cjun(phopho s63), Abcam ab32385, anti-junb (EPR6518) Abcam ab128878, anti-fosl1 Abcam ab232745, anti-c-jun(EP693Y) Abcam ab40766, anti-laminB1(EPR8985) Abcam ab133741, anti-FetuinA(EPR17839-163) Abcam ab187051, anti-Blm(B-4) santa cruz sc-365753, c-Jun rabbit mAb (60A8) CST 9165T, Junb rabbit mAb(C37F9) CST 3753S, Fosl1 mouse mAb(c-12) santa cruz sc-28310, dRNH1 antibody, 2xHBD antibody, CD45 FITC(104) Invitrogen MCD45201, Nestin antibody Beyotime AN205-1, E-cadherin antibody(DECAM-1) Santa Cruz sc-59778, CD144 antibody BD 550548, Sca-1 antibody (D7) BD 557403, FITC-labelled Sca-1(E13-161.7) biolegend 122506, APC-cy7-labelled CD48 (HM48-1) BioLegend, 103431, FITC-labelled Ki-67(SolA15) ThermoFisher, 11-5698-80. |
| Validation | The dRNH1 antbody and 2xHBD antibody were validated by previous articles with the following links:<br>dRNH1 antibody:https://doi.org/10.1083/jcb.202101092<br>2xHBD antibody:https://www.science.org/doi/10.1126/sciadv.abe3516<br>Antibody validation information can be found at the manufacturer's website with the following links:<br>CD34 FITC(581) BD 555821:https://www.bdbiosciences.com/zh-cn/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/fitc-mouse-anti-human-cd34.555821<br>Biotin-labled lineage markers ThermoFisher<br>Sca-1 PE-cy7(D7) ThermoFisher 25-5981-82: https://www.thermofisher.cn/cn/zh/antibody/product/Ly-6A-E-Sca-1-Antibody-clone-D7-Monoclonal/25-5981-82<br>c-Kit APC (2B8) ThermoFisher 17-1171-82:https://www.thermofisher.cn/cn/zh/antibody/product/CD117-c-Kit-Antibody-clone-2B8-Monoclonal/17-1171-82<br>CD150 PE(mShad150) ThermoFisher 12-1502-82:https://www.thermofisher.cn/cn/zh/antibody/product/CD150-Antibody-clone-mShad150-Monoclonal/12-1502-82<br>CD48 FITC(HM48-1) ThermoFisher 11-0481-82:https://www.thermofisher.cn/cn/zh/antibody/product/CD48-Antibody-clone-HM48-1-Monoclonal/11-0481-82<br>Phosphor-histone H2A.X(ser139)(20E3) rabbit mAb cell signaling technology(CST) 9718s:https://www.cellsignal.cn/products/primary-antibodies/phospho-histone-h2a-x-ser139-20e3-rabbit-mab/9718<br>Tlr4 antibody abcam ab13556:https://www.abcam.cn/products/primary-antibodies/tlr4-antibody-ab13556.html |

FetuinA antibody(EPR17839-163) abcam ab187051:https://www.abcam.cn/products/primary-antibodies/ahsg-antibody-epr17839-163-ab187051.html
p-RPA(phopho s33) abccam ab211877:https://www.abcam.cn/products/primary-antibodies/rpa32rpa2-phospho-s33-antibody-ab211877.html
Tlr4 antibody(UT41) ThermoFisher 53-9041-80:https://www.thermofisher.cn/antibody/product/53-9041-80.html?CID=AFLLS-53-9041-80
huamn FetuinA antibody(1F6B9) Proteintech 66094-1-Ig:https://www.ptglab.com/Products/AHSG-Antibody-66094-1-Ig.htm
Myd88(E11) Santa Cruz sc-74532:https://www.scbt.com/zh/p/myd88-antibody-e-11
anti-mouse CD117(ACK2) ThermoFisher 14-1172-85:https://www.thermofisher.cn/cn/zh/antibody/product/CD117-c-Kit-Antibody-clone-ACK2-Monoclonal/14-1172-85
Ter119 APC Biolegend 116212:https://www.biolegend.com/en-us/products/apc-anti-mouse-ter-119-erythroid-cells-antibody-1863
Gr-1 APC Biolegend 108412:https://www.biolegend.com/en-us/products/apc-anti-mouse-ly-6g-ly-6c-gr-1-antibody-456
Mac-1 APC Biolegend 101212:https://www.biolegend.com/en-us/products/apc-anti-mouse-human-cd11b-antibody-345
B220 APC Biolegend 103212:https://www.biolegend.com/en-us/products/apc-anti-mouse-human-cd45r-b220-antibody-442
CD3 APC Biolegend 100236:https://www.biolegend.com/en-us/products/apc-anti-mouse-cd3-antibody-8055
CD150 BV421(SLAM) Biolegend 115925:https://www.biolegend.com/en-us/products/brilliant-violet-421-anti-mouse-cd150-slam-antibody-7162
CD41 APC(MWRag30) Biolegend 133913:https://www.biolegend.com/en-us/products/apc-anti-mouse-cd41-antibody-7592
c-Kit goat mAb R&D AF1356:https://www.rndsystems.com/cn/products/human-mouse-cd117-c-kit-antibody_af1356
E-cadherin rabbit monoclonal antibdoy(24E10) CST 3195T:https://www.cellsignal.cn/products/primary-antibodies/e-cadherin-24e10-rabbit-mab/3195
anti-huamn albumin antibody(MAB1455) R&D 188835:https://www.rndsystems.com/cn/products/human-serum-albumin-antibody-188835_mab1455
laminin monoclonal antibody Abcam ab11575:https://www.abcam.cn/products/primary-antibodies/laminin-antibody-ab11575.html
Ki67-FITC(SolA15) ThermoFisher 11-5698-80:https://www.thermofisher.cn/cn/zh/antibody/product/Ki-67-Antibody-clone-SolA15-Monoclonal/11-5698-80
CD34 antibody(EP373Y) Abcam ab81289:https://www.abcam.cn/products/primary-antibodies/cd34-antibody-ep373y-ab81289.html
anti-p-junb(Thr102/Thr104) CST 8053S:https://www.cellsignal.cn/products/primary-antibodies/phospho-junb-thr102-thr104-d3c6-rabbit-mab/8053
anti-p-fosl1(s265) CST 3880S:https://www.cellsignal.cn/products/primary-antibodies/phospho-fra1-ser265-antibody/3880
anti-p-cjun(phopho s63) Abcam ab32385:https://www.abcam.cn/products/primary-antibodies/c-jun-phospho-s63-antibody-y172-ab32385.html
anti-junb (EPR6518) Abcam ab128878:https://www.abcam.cn/products/primary-antibodies/junb-antibody-epr6518-ab128878.html
anti-fosl1 Abcam ab232745:https://www.abcam.cn/products/primary-antibodies/fra1-antibody-ab232745.html
anti-c-jun(EP693Y) Abcam ab40766:https://www.abcam.cn/products/primary-antibodies/c-jun-antibody-ep693y-ab40766.html
anti-laminB1(EPR8985) Abcam ab133741:https://www.abcam.cn/products/primary-antibodies/lamin-b1-antibody-epr8985b-nuclear-envelope-marker-ab133741.html
anti-FetuinA(EPR17839-163) Abcam ab187051:https://www.abcam.cn/products/primary-antibodies/ahsg-antibody-epr17839-163-ab187051.html
anti-Blm(B-4) santa cruz sc-365753:https://www.scbt.com/zh/p/blm-antibody-b-4
c-Jun rabbit mAb (60A8) CST 9165T:https://www.cellsignal.cn/products/primary-antibodies/c-jun-60a8-rabbit-mab/9165
Junb rabbit mAb(C37F9) CST 3753S:https://www.cellsignal.cn/products/primary-antibodies/junb-c37f9-rabbit-mab/3753
Fosl1 mouse mAb(c-12) santa cruz sc-28310:https://www.scbt.com/zh/p/fra-1-antibody-c-12
CD45 FITC(104) Invitrogen MCD45201:https://www.thermofisher.cn/cn/zh/antibody/product/CD45-2-Antibody-Monoclonal/MCD45201?imageId=2517
Nestin antibody Beyotime AN205-1:https://www.beyotime.com/product/AN205.htm
E-cadherin antibody(DECAM-1) Santa Cruz sc-59778:https://www.scbt.com/zh/p/e-cadherin-antibody-decma-1
CD144 antibody BD 550548:https://www.bdbiosciences.com/zh-cn/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/purified-rat-anti-mouse-cd144.550548
Sca-1 antibody (D7) BD 557403:https://www.bdbiosciences.com/zh-cn/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/purified-rat-anti-mouse-ly-6a-e.557403
FITC-labelled Sca-1(E13-161.7) biolegend 122506: https://www.biolegend.com/ja-jp/products/fitc-anti-mouse-ly-6a-e-sca-1-antibody-3894?GroupID=BLG5162
APC-cy7-labelled CD48 (HM48-1) BioLegend, 103431:https://www.biolegend.com/ja-jp/products/apc-cyanine7-anti-mouse-cd48-antibody-8054
FITC-labelled Ki-67(SolA15) ThermoFisher, 11-5698-80: https://www.thermofisher.cn/cn/zh/antibody/product/Ki-67-Antibody-clone-SolA15-Monoclonal/11-5698-80

## Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

| | |
|---|---|
| Cell line source(s) | *State the source of each cell line used and the sex of all primary cell lines and cells derived from human participants or vertebrate models.* |
| Authentication | *Describe the authentication procedures for each cell line used OR declare that none of the cell lines used were authenticated.* |
| Mycoplasma contamination | *Confirm that all cell lines tested negative for mycoplasma contamination OR describe the results of the testing for mycoplasma contamination OR declare that the cell lines were not tested for mycoplasma contamination.* |
| Commonly misidentified lines (See ICLAC register) | *Name any commonly misidentified cell lines used in the study and provide a rationale for their use.* |

## Palaeontology and Archaeology

| | |
|---|---|
| Specimen provenance | *Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information). Permits should encompass collection and, where applicable, export.* |
| Specimen deposition | *Indicate where the specimens have been deposited to permit free access by other researchers.* |
| Dating methods | *If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.* |

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

| | |
|---|---|
| Ethics oversight | *Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.* |

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

## 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 | C57BL/6J, Cg-Tg(Alb-cre)21Mgn/J, B6.Cg-Gt(ROSA)26Sortm14(CAG-tdTomato)Hze/J, B6.129P2-Gt(ROSA)26Sortm1(DTA)Ly/J, C57BL/6JGpt-FetuinA ko(cas9).<br>Mice were maintained and bred in a specific pathogen-free facility in ventilated cages, a maximum of 5 mice per cage, on a 12-hour day-night cycle, at 20-26 °C and 30-70% humidity. For embryo collection, 8–10-week-old male and female were used. For the proportion analysis of LSK, CFC, whole genome sequencing, chromosome FISH and leukaemic model, 3-week-old mice were used. |
| Wild animals | No wild animals were used in this study. |
| Reporting on sex | The sex was not considered in our study design. About 500 mice were used in the study. |
| Field-collected samples | No field-collected samples were used in this study. |
| Ethics oversight | Animal procedures for mice were approved by the Scientific Investigation Board of Shanghai Jiao Tong University School of Medicine, Shanghai, china. |

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

## Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

| | |
|---|---|
| Clinical trial registration | *Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.* |
| Study protocol | *Note where the full trial protocol can be accessed OR if not available, explain why.* |
| Data collection | *Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.* |
| Outcomes | *Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.* |

## Dual use research of concern

Policy information about dual use research of concern

### Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

| No | Yes | |
|---|---|---|
| ☐ | ☐ | Public health |
| ☐ | ☐ | National security |
| ☐ | ☐ | Crops and/or livestock |
| ☐ | ☐ | Ecosystems |
| ☐ | ☐ | Any other significant area |

## Experiments of concern

Does the work involve any of these experiments of concern:

| No | Yes | |
|---|---|---|
| ☐ | ☐ | Demonstrate how to render a vaccine ineffective |
| ☐ | ☐ | Confer resistance to therapeutically useful antibiotics or antiviral agents |
| ☐ | ☐ | Enhance the virulence of a pathogen or render a nonpathogen virulent |
| ☐ | ☐ | Increase transmissibility of a pathogen |
| ☐ | ☐ | Alter the host range of a pathogen |
| ☐ | ☐ | Enable evasion of diagnostic/detection modalities |
| ☐ | ☐ | Enable the weaponization of a biological agent or toxin |
| ☐ | ☐ | Any other potentially harmful combination of experiments and agents |

# Plants

| | |
|---|---|
| Seed stocks | *Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.* |
| Novel plant genotypes | *Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.* |
| Authentication | *Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.* |

# 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<br>*May remain private before publication.* | *For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.* |
| Files in database submission | *Provide a list of all files available in the database submission.* |
| Genome browser session<br>(e.g. UCSC) | *Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.* |

## Methodology

| | |
|---|---|
| Replicates | *Describe the experimental replicates, specifying number, type and replicate agreement.* |
| Sequencing depth | *Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.* |
| Antibodies | *Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.* |
| Peak calling parameters | *Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.* |

| | |
|---|---|
| Data quality | *Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.* |
| Software | *Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.* |

# 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 | The murine mononuclear cells used for sorting or cell cycle analysis were isolated from placenta, foetal liver or bone marrow at different developmental stages. The tissues were dissected into single-cell suspensions, then incubated with antibodies. After staining, the cells were washed and resuspended in ice cold MACS for sorting or analysis.<br>The human mononuclear cells used for sorting were isolated from foetal liver or cord blood. The foetal livers were dissected into single cell suspensions, then the mononuclear cells from foetal liver or cord blood were separated using Ficoll density gradient centrifugation. Lineage-positive cells were depleted using the MagniSort Human haematopoietic lineage depletion kit. Lineage-negative cells were then incubated with antibodies. After staining, the cells were washed and resuspended in ice cold MACS for sorting. |
| Instrument | BD FACSAria3, MoFlo Astrios, cytoFLEX LX. |
| Software | FlowJo_v10 |
| Cell population abundance | Purity of sorted populations was assessed by post-sort analysis. Purity of over 95% was routinel achieved. |
| Gating strategy | For murine HSPCs: The first gate excluded any cellular debris based on FSC-A vs SSC-A. These cells were then sub-gated to identify only single cells, based on removal of outliers from the SSC-W vs SSC-H plot. lineage-negative cells were isolated by gating for cells with the lowest expression of lineage panel of antibodies (B220, CD3, Gr-1, Ter119, lineage-PE, or lineage Percp-cy5.5 or Lineage-APC-cy7). Sca-1+c-Kit+ cells were gated based on a high expression of Sca-1 and c-Kit (Sca-1-PE-cy7 vs c-Kit-APC). Within the Sca-1+c-Kit+ population, the gates for LT-HSCs (CD150+CD48-), ST-HSCs (CD150-CD48-), MPPs (CD150-CD48+) (CD150-PE vs CD48-FITC) were defined.<br>For cell cycle analysis: the HSPCs used for the cell cycle analysis were gated as above. Then the final gate was defined based on the Hochest vs Pyronin Y, Hochest vs EdU-FITC, Ki67-PE vs SSA.<br>For human HSPCs: The first gate excluded any cellular debris based on FSC-A vs SSC-A. These cells were then sub-gated to identify only single cells, based on removal of outliers from the SSC-W vs SSC-H plot. Lineage-negative cells were isolated by gating for cells with lowest expression of lineage panel antibodies(CD2, CD3, CD10, CD11b, CD14, CD16, CD19, CD56, CD123, CD235a, lineage-PE), within lineage-negative population, the CD34+ cells were gated based on a high expression of CD34(CD34-FITC vs SSC-A). |

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

# Magnetic resonance imaging

## Experimental design

| | |
|---|---|
| Design type | *Indicate task or resting state; event-related or block design.* |
| Design specifications | *Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.* |
| Behavioral performance measures | *State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).* |

## Acquisition

Imaging type(s) *Specify: functional, structural, diffusion, perfusion.*

Field strength *Specify in Tesla*

Sequence & imaging parameters *Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.*

Area of acquisition *State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.*

Diffusion MRI ☐ Used ☐ Not used

## Preprocessing

Preprocessing software *Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).*

Normalization *If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.*

Normalization template *Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.*

Noise and artifact removal *Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).*

Volume censoring *Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.*

## Statistical modeling & inference

Model type and settings *Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).*

Effect(s) tested *Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.*

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference *Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.*

(See Eklund et al. 2016)

Correction *Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).*

## Models & analysis

| n/a | Involved in the study |
|---|---|
| ☐ | ☐ Functional and/or effective connectivity |
| ☐ | ☐ Graph analysis |
| ☐ | ☐ Multivariate modeling or predictive analysis |

Functional and/or effective connectivity *Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).*

Graph analysis *Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).*

Multivariate modeling and predictive analysis *Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.*