

Corresponding author(s): Nicole Soranzo, Jing-wen Lin, Da Jia, and Lu Chen

Last updated by author(s): Jan 6, 2026

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
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

Bulk and single-cell RNA-seq datasets were included in this study. For bulk datasets, the raw files (fastq) of human (PRJEB26969), rhesus (PRJEB26956), rabbit (PRJEB26840), mouse (PRJEB26869), rat (PRJEB26889), opossum (PRJEB27035) and chicken (PRJEB26695) of organs development datasets were downloaded from NCBI. The raw files of zebrafish were retrieved through accession number GEO: GSE120581. The raw files of human (GSE142216) and mouse (EGAD00001000745) of lineage differentiation datasets were collected from GEO and EGA databases. Besides, Illumina NovaSeq 6000 platform was used for single-cell transcriptomic sequencing. Sequencing reads and single-cell expression matrices are deposited from the GEO under accession codes GSE276911.

Data analysis

FAScore code, documentation and tutorials are available at GitHub: <https://github.com/LuChenLab/FAScore.git>.

For bulk RNA-seq, the quality of the raw sequence data was checked using FastQC software (v0.11.8). Low-quality reads were trimmed using Trimmomatic (v0.38). Trimmed reads were aligned to the respective reference genomes using STAR (v2.6.1a). The gene expression and isoform abundance were quantified by RSEM (v1.3.1). The alternative splicing events were identified and quantified by MeDAS (Codes are available at <https://github.com/LuChenLab/MeDAS>) and AS types were inferred using SUPPA2 (v2.3) with 'local AS events' mode. NA, indicating no inferred AS type. The inferred AS events present in less than two samples were filtered out.

For scRNA-seq, the sequencing data were processed using the Cell Ranger (v6.1.2). The raw gene expression matrices were merged using R (v4.0.2) and converted to a Seurat object using Seurat (v4.0.2). The doublets were identified using DoubletFinder (v2.0.3). The differentially expressed genes (DEGs) were identified using the FindMarkers function in Seurat. Cellular differentiation states were predicted by using CytoTRACE (v0.3.3). The Augur (v1.0.3) was used to identify which lineage exhibit the most transcriptional changes between E15-KO and WT. Co-expression analysis was carried out using the hdWGCNA R package (v0.2.2). The co-regulatory networks of transcription factors and potential target genes were inferred using pySCENIC (v0.12.1). The co-regulatory network of TFs and their target genes was visualized using

Gephi (v0.10.1).

For flow cytometric analysis was performed using an Attune NXT (Life Technologies) flow cytometer and cell sorting were performed using a FACS Aria III (BD Biosciences). The data were analyzed using FlowJo software (v.10.8.1, BD).

For the gray value of the WB and the Fluorescence intensity quantified by imageJ (1.52a, , github.com/imagej/ImageJ) .

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](#)

For lineage differentiation datasets, raw sequencing data of mouse are deposited in the Gene Expression Omnibus (GEO) database under the accession number GSE142216 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE142216>), and raw sequencing data of human are available from the Blueprint Consortium website (<http://www.blueprint-epigenome.eu>) and the European Genome-phenome Archive with accession number EGAD00001000745 (<https://ega-archive.org/datasets/EGAD00001000745>). For FHO development datasets, the raw sequencing data are available through the NCBI: human (<https://www.ncbi.nlm.nih.gov/bioproject/PRJEB26969>), rhesus (<https://www.ncbi.nlm.nih.gov/bioproject/PRJEB26956>), rabbit (<https://www.ncbi.nlm.nih.gov/bioproject/PRJEB26840>), mouse (<https://www.ncbi.nlm.nih.gov/bioproject/PRJEB26869>), rat (<https://www.ncbi.nlm.nih.gov/bioproject/PRJEB26889>), Opossum (<https://www.ncbi.nlm.nih.gov/bioproject/PRJEB27035>) and chicken (<https://www.ncbi.nlm.nih.gov/bioproject/PRJEB26695>). For zebrafish, the accession number for the RNA-seq data is the GEO database: GSE120581 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE120581>). The data of single-cell RNAseq, including sequencing reads and single-cell expression matrices, are available from the GEO: GSE276911 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE276911>). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (<https://proteomecentral.proteomexchange.org>) via the iProX partner repository with the dataset identifier PXD072614 and PXD072611. Source Data are provided with this paper. The intermediate data related to model training and prediction have been deposited in the Figshare database (<https://doi.org/10.6084/m9.figshare.29069336.v4>). All newly generated plasmids and other relevant materials are available upon request from the corresponding author.

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	All the data used in our study are publicly available. This information has not been collected.
Reporting on race, ethnicity, or other socially relevant groupings	All the data used in our study are publicly available. This information has not been collected.
Population characteristics	All the data used in our study are publicly available. This information has not been collected.
Recruitment	All the data used in our study are publicly available. This information has not been collected.
Ethics oversight	All the data used in our study are publicly available. This information has not been collected.

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	At least twenty Tbc1d23ΔE15/ΔE15 (E15-KO) mice or C57/BL6 mice were used for colony-forming unit assay, blood routine, and flow cytometry analysis. At least fifty AB zebrafish strain were used for whole Embryo Staining for globin expression. At least forty Tg (gata1:DsRed) Transgenic zebrafish lines were used for quantification analysis of the Gata1: DsRed intensity. For other experiments, sample sizes were derived from established protocols in our laboratory using similar experimental designs. All sample size were indicated in the corresponding figure legends.
-------------	---

Data exclusions	No data were excluded from this study.
Replication	All experiments were replicated at least three times. We confirm that all attempts at replication were successful.
Randomization	In mouse studies, animals were not randomized because the experiments aimed to compare strain or genotype differences. To minimize variation, all mice were age-matched and maintained under identical housing conditions. For in vitro studies, we either pooled cells from multiple biological replicates or applied different treatments to cells from the same replicate, thereby enabling paired comparisons.
Blinding	The colony formation assay was performed blinded. For all other quantitative data collected by instruments, blinding was unnecessary

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

antibodies used for flow cytometry
 Fixable Viability Stain 520 BD 564407
 PE-Cy7 anti-mouse CD117 Biolegend 105814 2B8
 BV605 anti-mouse CD71 BD 563013 C2
 BV711 anti-mouse CD11b BD 563168 M1/70
 BV421 anti-mouse F4/80 Biolegend 123137 BM8
 PE anti-mouse Gr-1 Biolegend 108408 RB6-8C5
 Fixable Viability Stain 780 BD 564995
 PE anti-mouse/human CD44 Biolegend 103008 IM7
 FITC anti-mouse Ter119 Biolegend 116206 TER-119
 FITC anti-mouse CD3ε Biolegend 100306 145-2C11
 FITC anti-mouse Gr-1 Biolegend 108406 RB6-8C5
 FITC anti-mouse CD11b Biolegend 101206 M1/70
 FITC anti-mouse B220 Biolegend 103224 RA3-6B2
 BV785 anti-mouse Sca-1 Biolegend 108139 D7
 PE anti-mouse CD117 Biolegend 105808 2B8
 BV711 anti-mouse CD16/32 Biolegend 101337 93
 Percp-cy5.5 anti-mouse CD34 Biolegend 119328 MEC14.7
 PE-cy7 anti-mouse CD48 Biolegend 103424 HM48-1
 PE-CF594 anti-mouse CD127 BD 562419 SB/199
 BV421 anti-mouse CD135 BD 562898 A2F10.1
 APC anti-mouse CD150 Biolegend 115910 TC15-12F122
 BV510 anti-mouse Ter119 Biolegend 116237 TER-119
 APC-Cy7 anti-mouse CD117 Biolegend 105826 2B8
 PerCP-Cy5.5 anti-mouse CD3ε Biolegend 100328 145-2C11
 BV421 anti-mouse CD19 Biolegend 115537 6D5
 PE/Dazzle™ 594 anti-mouse F4/80 Biolegend 123146 BM8
 PE-Cy7 anti-mouse Ly6G Biolegend 127617 IA8
 BV605 anti-mouse Ly6C Biolegend 128036 AL-61
 PE anti-mouse CD71 BD 113807 RI727
 APC anti-mouse CD41 Biolegend 133914 MWRReg30

antibodies used for WB or IF
 Rabbit anti-TBC1D23 polyclonal IgG Proteintech 17002-1-AP
 Rabbit anti-TBC1D23 polyclonal IgG Sigma HPA038152
 Mouse anti-RanBP2 monoclonal IgG1 κ Santa Cruz Biotechnology sc-74518
 Rabbit Anti-RabGAP1 polyclonal IgG Proteintech 11102-1-AP
 Mouse anti-Flag-tag polyclonal IgG Proteintech 20543-1-AP
 Mouse anti-HA-tag monoclonal IgG1 Proteintech 66006-1-IG
 Mouse anti-GFP IgG2a Proteintech 66002-1-Ig

Rabbit anti-TGN46 polyclonal IgG Abcam ab50595
 Mouse anti-GM130 IgG1, κ BD 610822
 Mouse anti-CI-MPR Monoclonal IgG1 Bio-Rad MEM-238
 Rabbit anti-ZFP1 polyclonal IgG Invitrogen PA5-53593
 Rabbit anti-HDAC1 monoclonal IgG CST 34589
 Mouse anti-HDAC1 monoclonal IgG1 CST 5356
 Rabbit anti-SUMO1 monoclonal IgG HUABIO ET1606-53
 anti-SUMO2/3 monoclonal IgG HUABIO ET1701-17
 Goat anti-Mouse IgG Seconadary antibody HRP conjugated IgG(H+L) SAB L3032-2
 Goat anti-Rabbit IgG Seconadary antibody HRP conjugated IgG SAB L3012-2
 FITC affinipure goat anti-Mouse IgG Jackson ImmunoResearch 115-095-003
 FITC affinipure goat anti-Rabbit IgG Jackson ImmunoResearch 111-545-003
 Alexa Fluor 647affinipure goat anti- Mouse IgG Jackson ImmunoResearch 115-605-003

Validation

All the antibodies have been validated for the use of flow cytometry, immunofluorescence, immunoprecipitation or WB analyses. Data are available on the manufacturer's website.

Eukaryotic cell lines

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

Cell line source(s)

HEK-293T and K562 was purchased from Type Culture Collection of the Chinese Academy of Sciences. c-kit⁺ cells were isolated from the bone marrow of 8? to 12?week?old female wild?type C57BL/6J mice or Cas9 knock?in mice.

Authentication

No additional authentication was performed.

Mycoplasma contamination

Cell lines were tested negative for mycoplasma.

Commonly misidentified lines
(See [ICLAC](#) register)

No commonly misidentified cell lines were used in the study.

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

Tbc1d23ΔE15/ΔE15 (E15-KO) mice and C57/BL6 mice were purchased from GemPharmatech (China); Rosa-Cas9 knock-in mice (stock no. 024858) were from The Jackson Laboratory; All mice were housed and bred under SPF (Specific Pathogen Free) conditions with a 12 hr light/dark cycle at the Laboratory Animal Center of West China Second University Hospital, with free access to food and water. AB zebrafish strain were purchased from China Zebrafish Resource Center, CZRC; Tg (gata1:DsRed) Transgenic zebrafish lines were gifted by Dr.Chengjian Zhao at Sichuan University. Both adult fish and embryos were maintained at 28.5°C in Aquatic Ecosystems.

Wild animals

The study did not involve wild animals.

Reporting on sex

Female mice were used in the primary experiments of this paper, and the major finding regarding the impact of Tbc1d23-E15 deletion on erythroid differentiation was validated in age-matched male mice to determine its sex-specificity. Mouse sex was determined by gonadal inspection at 4 weeks of age.

Field-collected samples

The study did not involve samples collected from the field.

Ethics oversight

Mice experiments were carried out following the protocols approved by the ethics committee of West China Second University Hospital[(2018) Animal Ethics Approval No.004]. Zebrafish experiment procedures were approved by the Animal Welfare Center of West China Hospital of Sichuan University.

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

Plants

Seed stocks

The study did not involve Plants.

Novel plant genotypes

The study did not involve Plants.

Authentication

The study did not involve Plants.

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

Whole BM cells were collected from the tibia, femur and ilium, filtered through 70-µm nylon mesh, quantified by PBS and counted by Countstar Altair (Counstar). A portion of the cells was set aside to assess the erythropoiesis stage of the RI, RII, RIII, RIV and RV, while the remaining cells were lysed with red blood cell buffer (Solarbio, #R1010), washed with PBS and filtered through 70-µm nylon mesh. The BM cells were re-quantified by PBS and counted. All single cell suspensions were stained with antibodies at 4 °C for 30 min.

Instrument

Flow cytometric analysis was performed using an Attune Nxt (Life Technologies) flow cytometer and cell sorting were performed using a FACS Aria III (BD Biosciences)

Software

The data were analyzed using FlowJo software (v.10.8.1, BD).

Cell population abundance

Cell numbers were determined using a Countstar Altair (Counstar). The volume analyzed for each flow cytometry sample was recorded, thereby allowing the calculation of total cell numbers in the original whole bone marrow samples.

Gating strategy

BM cells are first gated on the basis of cell size (FSC) and cell granularity (SSC). The stage of erythropoiesis (referred to in the article as RI, RII, RIII, RIV, RV) are characterized by Ter119, CD44 and cell size (FSC); RI are the population of ter119+CD44hiFSCChi; RII are the population of ter119+CD44medFSCChi; RIII are the population of ter119+CD44medFSCmed; RIV are the population of ter119+CD44lowFSClow; RV are the population of ter119+CD44-FSClow. LT-HSC cells are the population of Lin-CD117+Scal+CD48+CD150-; ST-HSC cells are the population of Lin-CD117+Scal+CD48-CD150-; MPP cells are the population of Lin-CD117+Scal+CD48-CD150+; MEP cells are the population of Lin-CD117+Scal-CD34-CD16/32-; GMP cells are the population of Lin-CD117+Scal-CD34+CD16/32med; CMP cells are the population of Lin-CD117+Scal-CD34+CD16/32hi. CLP cells are the population of Lin-CD127+CD135+; B cells are the population of Ter119-CD117-CD3-CD19+; T cells are the population of Ter119-CD117-CD19-CD3+; Neutrophil cells are the population of Ter119-CD117-CD3-CD19-CD11b+Ly6G+; monocyte are the population of Ter119-CD117-CD3-CD19-Ly6G-CD11b+F4/80-; erythroid are the population of Ter119+; MK are the population of Ter119-CD71-CD41+.

CFU colonies cells are first gated on the basis of cell size (FSC) and cell granularity (SSC). BFU-E cells are the population of CD117+CD41-CD71+; MK are the population of CD117+CD71-CD41+; Granulocytes are the populations of CD117-Gr1+; macrophage cells are the populations of CD117-Gr1-CD11b+F4/80+.

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