

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	In high-content compound screening, drug libraries were handled by Echo550 (Labcyte) and cells were imaged by confocal Opera Phenix (PerkinElmer). Real-time qPCR were performed using CFX96 and CFX384 (BIO-RAD). Confocal images were taken by LSM710 Confocal Microscope (Zeiss). Flow cytometry samples were analyzed on Aria II and Aria III Flow Cytometer System (BD Biosciences). RNA-seq, ATAC-seq and RRBS libraries were sequenced using Illumina Novaseq platform. Metabolites were extracted by 80% methanol and were analyzed by LC-MS/MS (Nexera X2 LC-30A, Shimadzu and QTRAP 6500+, SCIEX)
Data analysis	Data from high-content compound screening were analyzed by HARMONY (v4.9) analysis system and PhenoLOGIC machine learning algorithms (PerkinElmer). Analysis of flow cytometry data was performed by FlowJo (v9.3.2 and v10.4). Statistic figure were analyzed and drawn by Prism 8. Immunofluorescence staining images were adjusted by imageJ 1.53a. RNA-seq, RRBS and ATAC seq datasets were analyzed using the following tools: TopHat2 (v2.0.13), TopHat2 (v2.1.1), featureCounts (v1.6.5), TrimGalore software (with Cutadapt v1.18), Cufflinks (v2.2.1), Cell Ranger (v6.0.2), Scanpy python package (v1.4.5), Scanorama (v1.7.1), Multicore-tsne (v8.0.2), GSEA (v4.1.0), Bowtie2 (v2.3.3.1), BS-Seeker2 (v2.1.1), bedtools (v2.29.2), IGV_2.9.4, R (v 4.1.1), Python (v 3.8.2), Cell Ranger (v6.0.2), . Metabolomics were analyzed by MultiQuant software (v.3.0)(SCIEX).

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](#)

RNA-seq, scRNA-seq, ATAC-seq and RRBS raw data in this study were deposited at the NCBI Gene Expression Omnibus (GEO) repository under the accession number GSE185005. Reference data from other works can be accessed in NCBI GEO under their GSE numbers. scRNA-seq data of different embryonic stages, mESCs, D-EPSC, L-EPSC and TBLCs were downloaded from GSE100597, GSE45719, GSE78140, GSE145609 and GSM5195025. For the comparison of chromatin accessibility with mouse embryos were downloaded from GSE66581. The reference RRBS data of embryos and mESCs were downloaded from GEO database (embryo: GSE34864; mESCs: GSE133926). Source data are provided with this paper.

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	Sample size was determined based on previously published studies and ensure reproducibility. No statistical method was used to pre-determine sample size. The numbers of independent experiments for all the statistical results were described in the figure legend.
Data exclusions	No data was excluded from the analysis.
Replication	The experiments in this study were performed with at least three biological replicates unless specified. All attempts at replication were successful.
Randomization	No randomization methods were used as all samples were randomly collected from their populations.
Blinding	During embryo injection, the investigator injected the cells to the 8-cell stage embryos without knowing the cell types. In the other experiments, the investigators were not blinded to group allocation during data collection because no bias could have been removed by blinding.

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
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

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

Antibodies

Antibodies used

Primary antibodies:
 Mouse anti-MERVL-gag (Epigentek, A-2801), 1:200 dilution.
 Mouse anti-CDX2 (BioGenex, MU392A-UC), 1:200 dilution.
 Mouse anti-Elf5 (Santa-Cruz, sc-166653), 1:100 dilution.

Rabbit anti-Oct4 (Abcam, ab19857), 1:200 dilution.
 Goat anti-tdTomato (Biorbyt, orb182397), 1:100 dilution.
 Rat anti-Cytokeratin 8 (Developmental Studies Hybridoma Bank, TROMA-1, AB_531826), 1:10 dilution.
 Mouse anti-Proliferin (Santa Cruz, sc-271891), 1:50 dilution.
 Goat anti-Placental lactogen I (Santa Cruz, Sc-34713), 1:75 dilution.
 Rabbit anti-Tfap2c (Santa cruz, Sc-8977), 1:50 dilution.
 Secondary antibodies:
 Donkey anti-Mouse IgG 488 (Life technologies, A-21202), 1:500 dilution.
 Donkey anti-Mouse IgG 555 (Life technologies, A-31570), 1:500 dilution.
 Goat anti-Mouse IgG 647 (Life technologies, A-21241), 1:500 dilution.
 Goat anti-Rabbit IgG 647 (Life technologies, A-21244), 1:500 dilution.
 Donkey anti-Rabbit IgG 488 (Life technologies, A-21206), 1:500 dilution.
 Donkey anti-Goat IgG 555 (Life technologies, A-21432), 1:500 dilution.
 Donkey anti-Rat IgG 488 (Life technologies, A-21208), 1:500 dilution.

Validation

All the antibodies are commercially available and have been validated by the manufacturer.
 Mouse anti-MERVL-gag (Epigentek, A-2801), <https://www.epigentek.com/catalog/muervl-gag-polyclonal-antibody-p-5383.html>
 Mouse anti-CDX2 (BioGenex, MU392A-UC), <https://biogenex.com/product/anti-cdx-2/?v=1c2903397d88>
 Mouse anti-Elf5 (Santa-Cruz, sc-166653), <https://www.scbt.com/p/elf-5-antibody-g-2?requestFrom=search>
 Rabbit anti-Oct4 (Abcam, ab19857), <https://www.abcam.cn/oct4-antibody-ab19857.html>
 Goat anti-tdTomato (Biorbyt, orb182397), <https://www.biorbyt.com/tdtomato-antibody-orb182397.html>
 Rat anti-Cytokeratin 8 (Developmental Studies Hybridoma Bank, TROMA-1, AB_531826), <https://dshb.biology.uiowa.edu/TROMA-1>
 Mouse anti-Proliferin (Santa Cruz, sc-271891), <https://www.scbt.com/p/proliferin-antibody-e-10>
 Goat anti-Placental lactogen I (Santa Cruz, sc-34713), <https://www.scbt.com/p/placental-lactogen-i-antibody-p-17?requestFrom=search>
 Rabbit anti-Tfap2c (Santa cruz, Sc-8977), <https://www.scbt.com/p/ap-2gamma-antibody-h-77?requestFrom=search>
 Donkey anti-Mouse IgG 488 (Life technologies, A-21202), <https://www.thermofisher.cn/cn/zh/antibody/product/Donkey-anti-Mouse-IgG-H-L-Highly-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-21202>
 Donkey anti-Mouse IgG 555 (Life technologies, A-31570), <https://www.thermofisher.cn/cn/zh/antibody/product/Donkey-anti-Mouse-IgG-H-L-Highly-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-31570>
 Goat anti-Mouse IgG 647 (Life technologies, A-21241), <https://www.thermofisher.cn/cn/zh/antibody/product/Goat-anti-Mouse-IgG2a-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-21241>
 Donkey anti-Rabbit IgG 488 (Life technologies, A-21206), <https://www.thermofisher.cn/cn/zh/antibody/product/Donkey-anti-Rabbit-IgG-H-L-Highly-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-21206>
 Donkey anti-Goat IgG 555 (Life technologies, A-21432), <https://www.thermofisher.cn/cn/zh/antibody/product/Donkey-anti-Goat-IgG-H-L-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-21432>
 Donkey anti-Rat IgG 488 (Life technologies, A-21208), <https://www.thermofisher.cn/cn/zh/antibody/product/Donkey-anti-Rat-IgG-H-L-Highly-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-21208>

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)

tdTomato-labeled mESCs were derived from E3.5 inner cell mass (ICM) of C57BL/6-Gt(ROSA)26Sortm1(CAG-tdTomato)/Vst (Vitalstar) mouse.
 OG2-mESCs were derived from E3.5 inner cell mass (ICM) of C57BL/6-CBA-Tg(Pou5f1-EGFP)2Mnn/J (The Jackson Laboratory, Cat# 004654; RRID: IMSR_JAX:004654) mouse.
 R1-mESCs (ATCC, Cat# SCRC-1011; RRID: CVCL_2167)
 ciTotiSCs were derived from mESCs.

Authentication

All cell lines were confirmed with normal morphology and expression of cell specific genes. OG2-mESCs were validated by RNA-seq. tdTomato-labeled mESCs and OG2-mESCs were authenticated by chimera experiments (inject the cells into 8-cell stage mouse embryos). R1-mESCs were purchased from ATCC. ciTotiSCs were well characterized with the details listed in the manuscript.

Mycoplasma contamination

All cell lines were tested negative for mycoplasma contamination.

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

No commonly misidentified cell lines were used herein.

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals

All mice were housed in individually ventilated cages (maxima six mice per cage) in barrier facility at Tsinghua University. The mice were maintained on a 12/12-hour light and dark cycle, 22-26°C with sterile pellet food and water ad libitum.
 ICR mice purchased from Beijing Vital River Laboratory Animal Technology were used as embryo donor (6-8 weeks old) and the pseudopregnant recipients (4-6 weeks old) for embryo transplantation.

Wild animals

No wild animals were used in the study.

Field-collected samples

No field-collected samples were used in the study.

Ethics oversight

All the animal experiments in this study were performed in accordance with the guidelines and regulations of IACUC (Institutional

Ethics oversight

Animal Care and Use Committee) of Tsinghua University, Beijing, China. All the animal protocols used in this study have been approved by IACUC (Institutional Animal Care and Use Committee) of Tsinghua University.

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

Cultured cells were trypsinized and resuspended in appropriate medium, followed by filtering through 40 μ M cell strainers. To assess the contribution of tdTomato+ cells to different parts of conceptuses of E13.5 chimera, embryos, yolk sacs and placentas were cut into small pieces (~1 mm diameter) on ice. For placentas, samples were digested with accutase (Gibco, A1110501) for 10 min at 37 °C on the shaker. For embryos and yolk sacs, samples were digested with collagenase IV (addition of 1 U/mL DNase) for 30 min or 5 min at 37 °C, respectively, followed by incubation with TrypLE (Gibco, 12605028) for 3-5 min. After digestion, all tissues were dissociated into single cells by pipetting and filtered through 70 μ m cell strainers. Then, cells were centrifuged, resuspended and sorted by flow cytometer.

Instrument

Aria II and Aria III Flow Cytometer System (BD Biosciences).

Software

Data analysis was performed by FlowJo 10 software.

Cell population abundance

Sample purity of sorted cells were verified by reload and FACS analyse.

Gating strategy

Preliminary FSC/SCC gates were used to remove debris, doublets and other aggregated particles. The boundary between positive and negative was defined according to the negative control.

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