

Corresponding author(s): Yawei Gao, Xiao-ou Zhang, Shaorong Gao

Last updated by author(s): Aug 27, 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	Images were acquired using the Olympus FV3000 and processed using ImageJ software (version 1.8.0_172). The software used for collecting RT-qPCR data was the 7500 software v2.0. All sequencing data are collected from NovaSeq 6000 instrument.
Data analysis	All statistical analyses were performed using the R programming language (v4.1.1). Pearson correlation coefficient was calculated using the "cor.test" function with the default parameters. The "wilcox.test" function was used to perform the two-sided Wilcoxon rank-sum test for group comparisons. Plots were generated with the ggplot2 package (v3.5.0) and Graphpad Prism (v9). Other softwares or packages used for analysis was listed as follows: TrimGalore (v0.6.10) HISAT-3N (v2.2.1-3n-0.0.3) HISAT2 (v2.2.1) featureCounts (v2.0.3) RSeQC (v5.0.1) DESeq2 (v1.34.0) Bowtie (v2.5.1) R package vegan (v 2.6.4) StringTie2 (v2.2.1) R programming language (v4.1.1) ggplot2 package (v3.5.0) R package clusterProfiler (v4.2.2) R package Complexheatmap (v2.10.0)

MEME suite (v.5.0.2)
FIMO algorithm (v5.5.4)
bedtools (v2.30.0)

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 sequencing data reported in this paper have been deposited in the Genome Sequence Archive under project PRJCA044842 and PRJCA029644, and the accession code is CRA02943 and HRA008614. Mouse embryo Pol II Stacc-seq data, MNase-seq data and LET-seq data were downloaded from GSE135457 [<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE135457>], CRA011730 [<https://ngdc.cncb.ac.cn/gsa/browse/CRA011730>] and GSE235547 [<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE235547>]. The Nobox knockout RNA-seq data were downloaded from GSE143218 [<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE143218>]. HEK293T PRO-seq data were downloaded from GSE218128 [<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE218128>]. The mm10 reference genome (<https://hgdownload.soe.ucsc.edu/goldenPath/mm10/bigZips/mm10.fa.gz>) and hg38 reference genome (<https://hgdownload.soe.ucsc.edu/goldenPath/hg38/bigZips/hg38.fa.gz>) were used for data analysis. Repeat elements in the mm10 genome were identified with RepeatMasker from the UCSC Genome Browser (<https://hgdownload.soe.ucsc.edu/goldenPath/mm10/database/rmsk.txt.gz>). Source data are provided with this paper.

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

Reporting on race, ethnicity, or other socially relevant groupings

Population characteristics

Recruitment

Ethics oversight

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

Data exclusions

Replication

Randomization

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
☒	☐ 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

The following antibodies were used in this study. The dilution ratio is given at the end.
 Goat anti-SOX2 antibody R&D Systems AF2018-SP RRID:AB_355110 1:100
 Mouse anti-CDX2 antibody BioGenex MU392A-5UC RRID:AB_2923402 1:100
 Alexa Fluor 594 donkey anti-mouse Thermo Fisher A21203 RRID:AB_2535789 1:500
 Alexa Fluor 488 donkey anti-goat Thermo Fisher A11055 RRID:AB_2534102 1:500

Validation

All antibodies were obtained from commercial sources. Detailed validation statements are available on the corresponding manufacturers' websites, as provided below:

Goat anti-SOX2 antibody R&D Systems AF2018-SP ,cited in 191 publications: https://www.rndsystems.com/cn/products/human-mouse-rat-sox2-antibody_af2018
 Mouse anti-CDX2 antibody BioGenex MU392A-5UC ,cited in 4 publications: <https://www.labome.com/product/Biogenex/MU392A-5UC.html>
 Anti-FLAG antibody ,Millipore, F7425, <https://www.sigmaaldrich.cn/CN/zh/product/sigma/f7425>
 Anti-Alpha Tubulin Monoclonal antibody ,Proteintech, 1E4C11, <https://www.ptgcn.com/products/tubulin-Alpha-Antibody-66031-1-Ig.htm>
 Alexa Fluor 594 donkey anti-mouse ,Thermo Fisher ,A21203, cited in 87 publications:
<https://www.thermofisher.cn/cn/zh/antibody/product/Donkey-anti-Mouse-IgG-H-L-Highly-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-21203>
 Alexa Fluor 488 donkey anti-goat Thermo Fisher A11055 ,cited in 865 publications:
<https://www.thermofisher.cn/cn/zh/antibody/product/Donkey-anti-Goat-IgG-H-L-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-11055>

Eukaryotic cell lines

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

Cell line source(s)

HEK293T cell lines are from ATCC company.

Authentication

The cell line was authenticated by morphology .

Mycoplasma contamination

All cell lines tested negative for mycoplasma contamination.

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

No commonly misidentified cell lines were used.

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

B6D2F1 female mice (8–10 weeks old) and C57BL/6 female mice (3-8 weeks old) were purchased from WTLH Shanghai company. B6D2F1 male mice (10-14 weeks old) were purchased from WTLH Shanghai company. PWK male mice (8-10 weeks old) were from Jackson Laboratory.

Wild animals

This study did not involve wild animals.

Reporting on sex

All oocytes and embryos are from female mice and sperm are from male mice.

Field-collected samples	This study did not involve samples collected from the field.
Ethics oversight	All animal experiments were performed in accordance with the University of Health Guide for the Care and Use of Laboratory Animals and were approved by the Biological Research Ethics Committee of Tongji University. The mouse facility was kept at 21-25°C and 40-60% humidity with a 12h light/12h dark cycle.

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

Plants

Seed stocks	No plants involved.
Novel plant genotypes	No plants involved.
Authentication	No plants involved.