

Reporting Summary

Nature Research 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 Research policies, see [Authors & Referees](#) 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 (<i>n</i>) 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. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> 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 <i>d</i> , Pearson's <i>r</i>), 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	Bio RAD CFX Connect Real-Time System; High Speed Spinning Disk Confocal Microscope (Dragonfly); BD FACSVerser; Cytation 5; Illumina HiSeq 2500; BioTek Cytation5; Bowtie2 version 2.3.5; Genome Analysis Toolkit version 4 ; Control-FREEC v11.6; Trimmomatic-0.39; Picard-tools; Homer version 4.10.3.
Data analysis	R v3.6.1 and Seurat3 was used to analyze the sequencing data, R package clusterProfiler and ComplexHeatmap were used for GSEA analysis. DESeq2 and g:Profiler bioinformatics tool were used for transcriptome analysis. Graphpad Prism 8 was used to analyze statistical data and draw graphs in this study. iMaris 9.7 was used to export the Dragonfly data. Flowjo V10 was used to analyze the flow cytometry data. Bowtie2 version 2.3.5, samtools version 1.10, MACS2-2.2.6, Homer version 4.10.3, clusterProfiler, deepTools-3.5.0 were used to analyze ATAC sequencing data. Genome Analysis Toolkit version 4, Control-FREEC, Fastqc and Trimmomatic-0.39, Picard-tools were used to analyze whole genome sequencing data.

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/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [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 list of figures that have associated raw data
- A description of any restrictions on data availability

Source data are provided with this paper. The accession number of the sequencing data reported in this paper is NCBI GEO: GSE156920 [<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE156920>] (single cell RNA-seq), GSE147839 [<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE147839>] (bulk RNA-seq) and GSE157237 [<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE157237>] (ATAC-seq), the WGS data number is SRP282164 [<https://trace.ncbi.nlm.nih.gov/Traces/sra/?study=SRP282164>]. Additional data is available upon request.

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 sizes were determined based on preliminary experiments, commonly used sample sizes in comparable publications within this field, such as papers by Yang Y et al., 2017 and Chen G et al., 2010. We independently performed all experiments at least 3 times, allowing to judge the degree of variability between replicates. The sample sizes and number of repeats are defined in each figure legends.
Data exclusions	No data was excluded from the analysis.
Replication	All experiments presented here been repeated several times with similar results, the number of repeated experiments is indicated in the figure legends.
Randomization	For chimeric experiments, mice and embryos were randomly distributed amongst groups at the start of each experiment. For in vitro experiments using cells, cells were randomly assigned to either experiment or control groups.
Blinding	The investigators were blinded to allocation during experiments and outcome assessment.

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 and 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
☐	☒ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

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

Antibodies

Antibodies used	OCT4 (BD, 611203; ThermoFisher Scientific, MA5-14845; 1:500 for IF, 1:1000 for FACS); OCT4 (Abcam, ab181557; 1:200 for IF); NANOG (Abcam, ab21624; 1:200 for IF); KLF4 (Abcam, ab75486; 1:200 for IF); SOX2 (Stemgent, 09-0024; 1:500 for IF); rH2AX (Millipore, 16-202A; 1:1000 for FACS); GATA3 (Abcam, ab199428; 1:200 for IF, 1:1000 for FACS); KRT7 (Dako, M7018; 1:200 for IF, 1:1000 for FACS); HLA-G (Abcam, ab7758; 1:200 for IF); hCG (Abcam, ab9376; 1:500 for IF); FOXA2 (Abcam, ab108422; 1:200 for IF); SOX17 (R&D, AF1924; 1:200 for IF); GATA4 (R&D, AF2606; 1:200 for IF); GATA6 (R&D, AF1700; 1:200 for IF); TUJ1 (Abcam, ab107216; 1:500 for IF); NESTIN (R&D, MAB1259c; 1:200 for IF); CDX2 (Biogenex, Mu392a); GFP (Abcam, ab 13970; 1:200 for IF); CK18 (Abcam, ab668; 1:200 for IF); CK8 (Sigma, HPA049866; 1:200 for IF); Neu5Gc (Biolegend, 146901; 1:1000 for FACS); XIST (Biosearch Technologies, SMF-2038-1; 1:50 for FISH) Anti-NR2F2 (Abcam, ab211776; 1:100 for IF).
Validation	Antibodies were validated according to manufacturer's instruction.

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	Human EPS cell lines H1-EPS, ES1-EPS and ES2-EPS were published previously in reference 6. Human ES cell lines H1 was obtained from WiCell and authenticated by karyotype analysis. Human iPSC line was established in our laboratory. OT H1-EPS, iPS-EPS cell lines were established in this study.
Authentication	H1 was obtained directly from the WiCell. All the cell lines used in the paper were authenticated by karyotyping.
Mycoplasma contamination	All cultured cells were tested negative for mycoplasma contamination by PCR.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified lines were used in this study.

Palaeontology

Specimen provenance	<i>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).</i>
Specimen deposition	<i>Indicate where the specimens have been deposited to permit free access by other researchers.</i>
Dating methods	<i>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.</i>
☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.	

Animals and other organisms

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

Laboratory animals	Mice was used in this study. The immunodeficient NPG mice were 6-8 weeks old. The male and female ICR mice were 6-8 weeks old. The mice were housed with a 12 hr light/dark cycle between 06:00 and 18:00 in a temperature controlled room (22 ± 1°C) with free access to water and food.
Wild animals	This study did not involve wild animals.
Field-collected samples	This study did not involve field-collected samples.
Ethics oversight	The Ethics Committee of Peking University Health Science Center approved the protocols used in this study.

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

Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics	<i>Describe the covariate-relevant population characteristics of the human research participants (e.g. age, gender, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."</i>
Recruitment	<i>Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.</i>

Ethics oversight

Identify the organization(s) that approved the study protocol.

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.

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

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

(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

Cells were harvested and fixed by BD Cytofix/Cytoperm solution (BD, 554714) for 20 min at 4°C, and washed 3 times with PBS, stained at least 4 hours with primary antibody and 1 hour with secondary antibody. Then cells were washed again and detected. All the samples were analysed using BD FACS Verse cytometer. 561/488/647 channels were used to detect signals. Negative controls (no staining sample and isotype antibodies control) were used.

Instrument	BD FACS Verse cytometry
Software	FlowJo V10
Cell population abundance	Purity was computer determined by the software.
Gating strategy	Because of the various sizes of the dissociated cells analyzed, the majority of starting cells were included in FSC/SSC gates. The boundary between positive and negative is defined according to the negative controls.

☒ 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 (See Eklund et al. 2016)	Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.
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	Involvement 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.