nature research

Corresponding author(s): Xiao Wang, David A. Brafman

Last updated by author(s): Feb 4, 2021

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 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	FACSDiva (BD Biosciences) for flow cytometry, and NIS-Elements (Nikon) for microscopy.
Data analysis	FlowJo (FlowJo LLC) for flow cytometry analysis, NIS-Elements (Nikon) for microscopy, EditR (https://moriaritylab.shinyapps.io/editr_v10/) and CLC main (Qiagen) for Sanger sequencing analysis.

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 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

Relevant materials are available through respective vendors and plasmids are available through Addgene.

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	Example data shown in figures 4, 5, 6, 7, 8 is n=1.
Data exclusions	n/a
Replication	n/a
Randomization	n/a
Blinding	n/a

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	Anti-Nanog Rabbit Antibody (Thermo Fisher Scientific Cat# PA1-097, RRID:AB_2539867; https://scicrunch.org/resolver/AB_2539867) Anti-Oct4 Rabbit Antibody (Thermo Fisher Scientific Cat# PA5-27438, RRID:AB_2544914; https://scicrunch.org/resolver/AB_2544914) Anti-Sox2 Rabbit Antibody (Thermo Fisher Scientific Cat# PA1-094, RRID:AB_2539862; https://scicrunch.org/resolver/AB_2539862) Donkey Anti-Rabbit Alexa Fluor 488 Antibody (Thermo Fisher Scientific Cat# A-21206, RRID:AB_2535792; https://scicrunch.org/resolver/AB_2535792) Donkey Anti-Mouse Alexa Fluor 488 Antibody (Thermo Fisher Scientific Cat# A-21202, RRID:AB_141607; https://scicrunch.org/resolver/AB_141607) Rabbit Anti-Alpha-fetoprotein Antibody (Santa Cruz Biotechnology Cat# sc-8399, RRID:AB_626665; https://scicrunch.org/resolver/AB_626665) Mouse Anti-Actin Antibody (Santa Cruz Biotechnology Cat# sc-53015, RRID:AB_628683; https://scicrunch.org/resolver/AB_628683) Mouse Anti-TUJ1 Antibody (Fitzgerald Industries International Cat# 10R-T136A, RRID:AB_1289248; https://scicrunch.org/resolver/AB_1289248)
Validation	Antibody validation was conducted by the manufacturer and is available on their website.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	human induced pluripotent stem cells (hiPSCs); Brookhouser et al. 2020; Stem Cell Reports. 2020 Feb 11;14(2):184-191
Authentication	(i) Expression of pluripotency markers OCT4, NANOG, and SOX2, (ii) ability to differentiate in vitro into cell types representative of the three germ layers (i.e. endoderm, mesoderm, and ectoderm), and (iii) a normal euploid karyotype

Mycoplasma contamination	All cell lines were confirmed to be free from mycoplasma contamination
Commonly misidentified lines (See ICLAC register)	None

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 dissociated with accutase and sorted in mTeSR1 media containing ROCKi.
Instrument	BD FACSAria II Sorter
Software	BD FACSDiva (BD Biosciences)
Cell population abundance	GFP+ cell abundance can vary depending on transfection conditions but is often 5-10% of gated cell populations.
Gating strategy	Cells are gated by FSC:SSC to remove extracellular debris. For fluorescent channels: Mock transfected cells are used to define double negative cells, reporter plasmid only transfections are used to define BFP or mCherry positive populations.

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