nature portfolio

Corresponding author(s):   Jared Toettcher, Harry McNamara

Last updated by author(s):   8/9/2024

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	Imaging data was collected using NIS Elements version 4.40.00 (Nikon Inc). Cell sorting was performed using the Sony SH800 cell sorter (Sony Inc; no version number). No other software was used in data collection.
Data analysis	Image analysis was performed in ImageJ 1.54f and MATLAB R2024a. Single-cell sequencing data were processed and analyzed using the publicly available Single-Cell Analysis in Python (scanpy version 1.10) software package. All code is available on authors' Github repository (www.github.com/toettchlab/McNamara2024).

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

Single-cell RNA sequencing data generated in this study is available through the NCBI Gene Expression Omnibus (GEO) (accession number GSE274389). Plasmids

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	N/A
Reporting on race, ethnicity, or other socially relevant groupings	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

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	No statistical methods were used to pre-determine sample sizes; sample sizes were designed to be similar to those reported in recent publications (Beccari et al, Nature 2018; Merle et al, Nat Struct Mol Biol 2024).
Data exclusions	A small proportion of gastruloids (<10%) failed to elongate (consistent with field standards for gastruloid reproducibility; see Turner et al, Development 2017) and were excluded from anterior-posterior profile analysis (Fig. 3B, 7D-G, S8D-E). For bulk gastruloid reaggregation experiments (Fig. 4C; Fig. S5C-D), very small satellite ES cell aggregates and multipolar aggregates were excluded from analysis of A-P profiles.
Replication	All experiments were performed in at least two biological replicates to ensure reproducibility. All attempts at replication were successful.
Randomization	No randomization was performed because no covariates were present in the basic molecular and cell biology performed here.
Blinding	No blinding was performed because all experiments were conducted by a single experimental biologist.

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	anti-Pcdh19 (Abcam ab191198), anti-Aldh1a2 (Abcam ab156019), anti-Cdh1 (Cell Signaling Technology 3195T), Goat anti-rabbit Alexa647 (Invitrogen A27040), anti-Brachyury (R&D Systems, AF2085).
Validation	We used well established commercial antibodies from reputable vendors that have also been used in prior publications (e.g., Pcdh19: Wu et al, Mol Brain 2021; Cdh1: extensive validation on Cell Signaling website and 3,000+ citations; Brachyury: 100+ citations).

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	E14tg2a mouse embryonic stem cell line, obtained directly from ATCC for this paper.
Authentication	Cell line was not independently authenticated, but is authenticated in-house at ATCC by STR profiling.
Mycoplasma contamination	Lines were tested and confirmed negative for mycoplasma (ATCC Universal Mycoplasma Kit 30-1012K).
Commonly misidentified lines (See ICLAC register)	None.

Plants

Seed stocks	N/A
Novel plant genotypes	N/A
Authentication	N/A

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	Methods are described in the Supplementary Methods, from which the following is quoted: To assess gastruloids via flow cytometry, at least n = 30 gastruloids were first pooled into a 1.5 mL Eppendorf tube, centrifugated (500 rcf, 3 min), washed in PBS, and then trypsinized in 100 µL of TrypLE Express at 37 °C. Gastruloids were trypsinized for 3 minutes, retrieved and triturated to dislodge gastruloids, and returned to 37 for an additional 3 minutes to complete digestion. Trypsinization was then quenched with 300 µL of N2B27, after which samples were immediately assayed via flow cytometry. Cytometer events were first gated for events corresponding to cells based on forward scattering and backscattering profiles, and then for single cells based on the height and area of forward scattering events (Fig. S1D).
Instrument	Sony SH800 fluorescence-activated cell sorter
Software	Sony standard flow cytometry software
Cell population abundance	Because we worked with a single cell line, 100% of cell events corresponded to the relevant cell population.
Gating strategy	Cells were gated on size and shape (forward scatter / backscatter) and doublets were discriminated using forward scatter height vs area. See Fig. S1E.

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