nature portfolio

Corresponding author(s): Thomas Gregor

Last updated by author(s): Nov 10, 2023

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	Zen Blue 3.3 (Zeiss) was used to acquire data on both custom designed microscopes and a Zeiss LSM 980 confocal microscope.
Data analysis	Python (versions 3.6.13 and 3.9) codes were used to analyze all of our data; they can be obtained upon request. Some parts using cell identification use Cellpose 2.0 software. Code can be found here: https://gitlab.pasteur.fr/tglab/gastruloids_precisionandscaling

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

Processed immunofluorescence staining data is available as maximum projection images for individual gastruloids, organized by figure number. All images have been deposited on the Zenodo repository under doi: 10.5281/zenodo.8108188. Raw images are available upon request.

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	All of our statistical estimates are largely within the 95% confidence interval.
Data exclusions	no outliers removed unless clear experimental flaws were identified
Replication	Experiments were repeated by two experimentalists and analyzed independently leading to the same conclusions.
Randomization	Randomization happens automatically upon seeding of cell aggregates, as a random subset of cells is chosen from a vastly larger cell population.
Blinding	Data reproducibility across three different experimenters was checked.

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	Antibodies against murine proteins SOX2(rat) CDX2(rabbit) BRA(rabbit) FOXC1(rabbit) Provider: eBioscience(14-9811-80), Invitrogen(EPR2764Y), Abcam(ab209665), Abcam(ab223850), respectively Secondary antibodies: Anti-Rat AF488(Intitrogen A21208) and Anti-Rabbit AF647(Invitrogen A31573)

Validation	all antibodies were commercially acquired and used in multiple studies before; their linearity was assessed in our control experiments. Antibody, reference, provider, concentration: SOX2, 14-9811-82, eBioscience, 1/200 CDX2, EPR2764Y, Invitrogen, 1/200 T/BRACHYURY, ab209665, abcam, 1/150 FOXC1, ab223850, abcam, 1/500 Anti-Rat IgG-488nm, A-21208, Invitrogen, 1/500 Anti-Rabbit IgG-647nm, A-31573, Invitrogen, 1/500 All the antibodies, were validated by the companies, they are stated to react to mouse and were validated by western blot. They are routinely used (data from CiteAb): 14-9811-82: 91 citations EPR2764Y: 9 citations (+151 citation when purchased from Abcam) ab209665: 43 citations ab223850: 6 citations

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	mouse embryonic stem cells mESCs (129/svev, EmbryoMax) are commercially available from EmbryoMax
Authentication	cell lines were not authenticated
Mycoplasma contamination	all cell lines tested negative for mycoplasma
Commonly misidentified lines (See ICLAC register)	no misidentified cell lines used