nature research

Corresponding author(s): Jeffrey R. Millman

Last updated by author(s): February 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	The BD FACSDiva software was used to acquire flow cytometry data. The StepOne Software v2.3 was used to collect qRT-PCR data.
Data analysis	FlowJo (10.5.3) was used to analyze flow cytometry data. ImageJ/Fiji (1.52e) was used to process microscopy images. GraphPad Prism v8 was used for creating graphs and statistical 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

The original data for Figures 5 and 6 as well as sample calculations for these data are available in the Source Data section. Any other data are available from the corresponding author upon reasonable 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 size for the representative experiments showcasing the ability of each cell line to differentiate into beta cells was determined by sample availability for each cell line.
Data exclusions	N/A
Replication	In this manuscript, we show this protocol to work reproducibly across 10 different cell lines. Furthermore, we have published on five of these lines previously, and here we present new data sets for each of these lines that show similar values. For all data in this manuscript, each replicate was a separate well or a group of cell clusters that was collected and processed (e.g., separate RNA extractions and PCR reactions).
Randomization	Randomization of samples was not applicable due to the nature of the in vitro experiments.
Blinding	The experiments in this manuscript collected specific quantitative data, and therefore a blinding method was not used.

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

- rat anti-C-peptide (Developmental Studies Hybridoma Bank – University of Iowa, GN-ID4-S, RRID:AB_2255626)
- mouse anti-NKX6-1 (Developmental Studies Hybridoma Bank – University of Iowa, F55A12-S, RRID:AB_532379)
- goat anti-PDX1 (R&D Systems, AF2419, RRID:AB_355257)
- sheep anti-NEUROG3 (R&D Systems, AF3444, RRID:AB_2149527)
- rabbit anti-CHGA (ABCAM, ab15160, RRID:AB_301704)
- rabbit anti-SST (ABCAM, ab64053, RRID:AB_1143012)
- (histology only) mouse anti-SST (Santa Cruz Biotechnology, sc-55565, RRID:AB_831726)
- mouse anti-GCG (ABCAM, ab82270, clone IMD-7, RRID:AB_1658481)
- (histology only) rabbit anti-GCG (Cell Marque, 259A-18, RRID:AB_1158356)
- mouse anti-SOX17 (R&D Systems, MAB1924, clone # 245013, RRID: AB_2195646)
- rabbit anti-FOXA2 (MilliporeSigma, 07-633, RRID:AB_390153)
- mouse anti-OCT-3/4 (Santa Cruz Biotechnology, sc-5279, clone C-10, RRID:AB_628051)
- goat anti-NANOG (R&D Systems, AF1997, RRID:AB_355097)

- anti-rat alexa fluor 488 (Invitrogen, A21208, RRID:AB_141709)
- anti-mouse alexa fluor 488 (Invitrogen, A21202, RRID:AB_141607)
- anti-mouse alexa fluor 647 (Invitrogen, A31571, RRID:AB_162542)
- anti-mouse alexa fluor 594 (Invitrogen, A21203, RRID:AB_141633)
- anti-goat alexa fluor 488 (Invitrogen, A11055, RRID:AB_2534102)
- anti-goat alexa fluor 647 (Invitrogen, A21447, RRID:AB_2535864)
- anti-sheep alexa fluor 594 (Invitrogen, A11016, RRID:AB_10562537)
- anti-rabbit alexa fluor 647 (Invitrogen, A31573, RRID:AB_2536183)

	• anti-rabbit alexa fluor 488 (Invitrogen, A21206, RRID:AB_2535792) • anti-rat PE (Jackson ImmunoResearch, 712-116-153, RRID:AB_2340657)
Validation	We have previously published data using these antibodies (references 37-39). Further immunocytochemistry validation and references can be found on the manufacturers' websites.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	HUES8, 1013-4FA, 1016SeVA, 1026-3FC, and 1031SeVA were provided by Harvard university. The H1 line was obtained from WiCell. AN1.1 and WS4 lines were generated at the Washington University in St. Louis Genome Engineering and iPSC Center. T2D001A was generated by us from mesenchyme with the Sendai virus kit.
Authentication	All lines have been authenticated with DNA fingerprinting.
Mycoplasma contamination	All stem cell lines test negative on routine checks for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No lines were used from this list.

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	TrypLE was used to disperse cells, and they were fixed for 30 minutes with 4% PFA. Cells were washed with PBS and incubated in ICC solution (see "Reagent Setup") for 45 minutes, incubated overnight at 4°C with primary antibodies, and then incubated for 2 hours at 4°C with the secondary antibodies. Cells were then washed with ICC and filtered. See section on "Intracellular flow cytometry" in the text for full methodology.
Instrument	LSRII flow cytometer (BD Biosciences)
Software	Acquisition: BD FACSDiva Analysis: FlowJo (10.5.3)
Cell population abundance	Flow cytometry analysis for the indicated markers was often more than 30% for differentiating cells but less than 1% for stem cells stained with the same markers.
Gating strategy	Debris and doublets were gated out. Stem cells that negatively stained for the selected markers were then used as a negative control. Examples of this gating strategy are shown in Figure 3 and Supplemental Figure 4.

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