nature research

Corresponding author(s): Jose M. Polo

Last updated by author(s): Jan 25, 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 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 (*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	Motic Image Plus 2.0, LightCycler 480 software (Roche, version 1.5.1.62), BD FACS TM software, version 1.2.0.142 for influx sorters, Leica application suite X version 3.5.6.21594 for SP8, Leica application suite X version 3.7.1.21655 for DMi8, ZEISS Efficient Navigation version 6.0.0.485 for ZEISS LSM 780.
Data analysis	Data was analyzed using: GraphPad Prism (v7), ImageJ (v1.8.0_112), Imaris x64 version 9.5.0, QuPath (version 0.2.3), Adobe Photoshop (version 21.2.3), Cellranger (v2.1.0, v3.0.2, v3.1.0), base R(3.6.3), Seurat 3.1.5, Cellranger 3.1.0, STAR aligner (v2.5.1b), limma 3.42.2, tidyverse 1.3.0, RColorBrewer 1.1-2, ggplot2 3.3.1, egg 0.4.5, shiny 1.2.0, cellSNP 0.3.0, vireoSNP 0.3.2, pheatmap (v1.0.12). All data were analysed with standard programs and packages as detailed above. Scripts can be found at http://github.com/jibsch/iBlastoids_scripts.

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

All processed data sets are available at (http://blastoid.ddnetbio.com/). Raw and processed next generation sequencing datasets of iBlastoids were deposited at the Gene Expression Omnibus (GEO) repository with the following accession numbers: GSE156596. Published human embryo datasets used in this study were obtained from Blakeley et al and Petropoulos et al under accession numbers GSE66507 and E-MTAB-3929 (ArrayExpress) respectively. For Extended Data Fig. 6h, datasets

were abstained from Okae et al Table S1, Linneberg-Agerholm et al: GSE138012 for the RNA-seq integration. All source data related to this study are available in the source data table related to each figure, which are provided within the online content of this paper. There are no restrictions on data availability.

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	We did not involve statistical methods to pre-determine the sample size, this was determined based on previous experience and other similar studies. We generated iBlastoids using three different donor fibroblasts in multiple runs of reprogramming experiments (at least 4 independent replicates). Furthermore, we performed transcriptional profiling of iBlastoids using scRNA-seq libraries (from 2 donors). Similar results/findings were observed. These suggest that our sample size is sufficient.
Data exclusions	No data were excluded.
Replication	Each experiment was reproduced at least in 4 biological replicates (from two donors) if not otherwise stated. Please refer to figure legends and methods for details. All replications were successful.
Randomization	The iBlastoids generated were pre-selected based on morphological assessment, however for all downstream analyses the pre-selected iBlastoids were chosen at random. The experiments were not randomized.
Blinding	The investigators were not blinded during data collection and analysis, as neither human/animal studies or specific grouping were involved in this manuscript.

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

Details of all antibodies used in this study were provided in Supplementary Table 9 with catalog number, lot number and commercial sources supplied.

For Immunostaining:
Antibody, Company, Catalogue Number, Lot Number, Dilution:
Mouse anti-GATA2 IgG2a Sigma-Aldrich Cat# WH0002624M1 Lot G2151-2D11 1:100
Rabbit anti-CK7 (KRT7) IgG abcam Cat# ab181598 Lot GR3214132-7 1:100
Mouse anti-CDX2 IgG1 abcam Cat# ab157524 Lot GR3301072-2 1:50
Rabbit anti-CDX2 IgG Cell Signaling Technology Cat# 12306 Lot 2 1:100
Rabbit anti-NANOG polyclonal abcam Cat# ab21624 Lot GR3260262-26 1:100
Mouse anti-NANOG IgG1 Invitrogen Cat# 14-5768-82 Lot 2003073 1:100
Mouse anti-OCT3/4 IgG2b Santa Cruz Biotechnology Cat# sc-5279 Lot I3015 1:50
Mouse anti-SOX17 IgG1 abcam Cat# ab84990 Lot GR3277101-3 1:100
Rat anti-SOX2 IgG2a Invitrogen Cat# 14-9811-82 Lot 2124609 1:200
Goat anti-GATA6 polyclonal R&D Systems Cat# AF1700 Lot KWT0418111 1:200
Rabbit anti-MMP2 IgG Cell Signaling Technology Cat# 40994 Lot 3 1:800
Mouse anti-hCG IgG1 abcam Cat# ab9582 Lot GR3285769-1 1:100
Alexa Fluor 488 Phalloidin ThermoFisher Cat# A-12379 Lot 2069633 1:200

Rabbit anti-Laminin IgG abcam Cat# ab11575 Lot GR3268366-6 1:100
Mouse anti-PKC ζ IgG2a Santa Cruz Biotechnology Cat# sc-17781 Lot F2520 1:100
Rabbit anti-GATA3 IgG abcam Cat# ab199428 Lot GR3266930-1 1:100
Rat anti-KRT8 IgG2a Merck Millipore Cat# MABT329 Lot 3429042 1:100
Rabbit anti-CCR7 IgG abcam Cat# ab32527 Lot GR259329-31 1:100
Mouse anti-Nestin IgG1 Merck Millipore Cat# MAB5326 Lot 2393925 1:100
Mouse anti-CXCR4 IgG2a-PE Miltenyi Biotec Cat# 130-117-690 Lot 5200108213 1:100
Rabbit anti-Brachyury IgG Cell Signaling Technology Cat# 81694 Lot 1 1:200
Mouse anti-HLA-G IgG1 [MEM-G/1] abcam Cat# ab7759 Lot GR3262011-5 1:50
Rabbit anti-SDC1 IgG Cell Signaling Technology Cat# 12922 Lot 1 1:400
Rabbit anti-Ki67 IgG abcam Cat# ab15580 Lot GR3198185-1 1:100

Goat anti-rat IgG AF555 secondary ThermoFisher Cat# A-21434 Lot 1726587 1:500
Goat anti-rabbit IgG AF555 secondary ThermoFisher Cat# A-21428 Lot 1786491 1:500
Goat anti-mouse IgG2b AF647 secondary ThermoFisher Cat# A-21242 Lot 1968702 1:500
Goat anti-mouse IgG-AF488 secondary ThermoFisher Cat# A-11029 Lot 2066710 1:500
Goat anti-mouse IgG1-AF488 secondary ThermoFisher Cat# A-21121 Lot 1964382 1:500
Goat anti-mouse IgG2a-AF488 secondary ThermoFisher Cat# A-21131 Lot 2136788 1:500
Goat anti-mouse IgG2a-AF555 secondary ThermoFisher Cat# A-21137 Lot 1899521 1:500
Goat anti-mouse IgG2a-AF647 secondary ThermoFisher Cat# A-21241 Lot 2056280 1:500
Donkey anti-mouse IgG-488 secondary ThermoFisher Cat# A-21202 Lot 1696430 1:500
Donkey anti-goat IgG-555 secondary ThermoFisher Cat# A-21432 Lot 872636 1:500
Donkey anti-rabbit IgG-647 secondary ThermoFisher Cat# A-31573 Lot 1903516 1:500

Validation

Antibodies obtained from the commercial source were validated by the suppliers, detailed validation analysis relevant literatures are provided on the company website for the products used in this study. Some antibodies were validated in a previously published study as indicated in methods or relevant literature was cited.

1. GATA2 (WH0002624M1) https://www.sigmaaldrich.com/catalog/product/sigma/wh0002624m1
2. KRT7 (ab181598) https://www.abcam.com/cytokeratin-7-antibody-epr17078-cytoskeleton-marker-ab181598.html
3. CDX2 (ab157524) https://www.abcam.com/cdx2-antibody-cdx2-88-ab157524.html
4. CDX2 (12306) https://www.cellsignal.com/products/primary-antibodies/cdx2-d11d10-rabbit-mab/12306
5. NANOG (ab21624) https://www.abcam.com/nanog-antibody-ab21624.html
6. NANOG (14-5768-82) https://www.thermofisher.com/antibody/product/Nanog-Antibody-clone-hNanog-2-Monoclonal/14-5768-82
7. OCT-3/4 (sc-5279) https://www.scbt.com/p/oct-3-4-antibody-c-10
8. SOX17 (ab84990) https://www.abcam.com/sox17-antibody-oti3b10-ab84990.html
9. SOX2 (14-9811-82) https://www.thermofisher.com/antibody/product/SOX2-Antibody-clone-Btjce-Monoclonal/14-9811-82
10. GATA6 (AF1700) https://www.rndsystems.com/products/human-gata-6-antibody_af1700
11. MMP2 (40994) https://www.cellsignal.com/products/primary-antibodies/mmp-2-d4m2n-rabbit-mab/40994
12. hCG (ab9582) https://www.abcam.com/hcg-beta-antibody-5h4-e2-ab9582.html
13. Alexa Fluor 488 Phalloidin (A-12379) https://www.thermofisher.com/order/catalog/product/A12379#/A12379
14. Laminin (ab11575) https://www.abcam.com/laminin-antibody-ab11575.html
15. PKC ζ (sc-17781) https://www.scbt.com/p/pkc-zeta-antibody-h-1
16. GATA3 (ab199428) https://www.abcam.com/gata3-antibody-epr16651-chip-grade-ab199428.html
17. KRT8 (MABT329) https://www.merckmillipore.com/AU/en/product/Anti-Cytokeratin-8-Antibody-clone-TROMA-1,MM_NF-MABT329
18. CCR7 (ab32527) https://www.abcam.com/ccr7-antibody-y59-ab32527.html
19. Nestin (MAB5326) https://www.sigmaaldrich.com/catalog/product/mm/mab5326?
20. CXCR4-PE (130-117-690) https://www.miltenyibiotec.com/AU-en/products/cd184-cxcr4-antibody-anti-human-12g5.html
21. Brachyury (81694) https://www.cellsignal.com/products/primary-antibodies/brachyury-d2z3j-rabbit-mab/81694
22. HLA-G (ab7759) https://www.abcam.com/hla-g-antibody-mem-g1-ab7759.html
23. SDC1 (12922) https://www.cellsignal.com/products/primary-antibodies/syndecan-1-d4y7h-rabbit-mab/12922
24. Ki67 (ab15580) https://www.abcam.com/ki67-antibody-ab15580.html
25. Goat anti-rat IgG AF555 (A-21434) https://www.thermofisher.com/antibody/product/Goat-anti-Rat-IgG-H-L-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-21434
26. Goat anti-rabbit IgG AF555 (A-21428) https://www.thermofisher.com/antibody/product/Goat-anti-Rabbit-IgG-H-L-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-21428
27. Goat anti-mouse IgG2b AF647 (A-21242) https://www.thermofisher.com/antibody/product/Goat-anti-Mouse-IgG2b-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-21242
28. Goat anti-mouse IgG-AF488 (A-11029) https://www.thermofisher.com/antibody/product/Goat-anti-Mouse-IgG-H-L-Highly-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-11029
29. Goat anti-mouse IgG1-AF488 (A-21121) https://www.thermofisher.com/antibody/product/Goat-anti-Mouse-IgG1-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-21121
30. Goat anti-mouse IgG2a-AF488 (A-21131) https://www.thermofisher.com/antibody/product/Goat-anti-Mouse-IgG2a-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-21131
31. Goat anti-mouse IgG2a-AF555 (A-21137) https://www.thermofisher.com/antibody/product/Goat-anti-Mouse-IgG2a-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-21137
32. Goat anti-mouse IgG2a-AF647 (A-21241) https://www.thermofisher.com/antibody/product/Goat-anti-Mouse-IgG2a-Cross-

Adsorbed-Secondary-Antibody-Polyclonal/A-21241
33. Donkey anti-mouse IgG-488 (A-21202) https://www.thermofisher.com/antibody/product/Donkey-anti-Mouse-IgG-H-L-Highly-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-21202
34. Donkey anti-goat IgG-555 (A-21432) https://www.thermofisher.com/antibody/product/Donkey-anti-Goat-IgG-H-L-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-21432
35. Donkey anti-rabbit IgG-647 (A-31573) https://www.thermofisher.com/antibody/product/Donkey-anti-Rabbit-IgG-H-L-Highly-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-31573

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	Human fibroblasts sourced from ThermoFisher (Catalogue number, C-013-5C and lot#1029000 for 38F and lot#1569390 for 32F) for reprogramming experiments.
Authentication	Human dermal fibroblasts were derived from tissues of donor (age and sex identified) and authenticated by ThermoFisher (via growth performance assay, testing for Hepatitis B virus, Hepatitis C virus, HIV-1 virus, mycoplasma as well as sterility testing of bacteria and fungus), as stated in the certificate of analysis. These dermal fibroblasts have also been authenticated in-house by immunostaining, qPCR, RNA-seq.
Mycoplasma contamination	Fibroblasts lines were tested by ThermoFisher. Furthermore, cell lines were regularly tested and were mycoplasma negative.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used in this study.

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	iBlastoids were dissociated with TrypLE express (ThermoFisher), and DPBS (ThermoFisher) supplemented with 2% FBS (ThermoFisher) and 10µM Y-27632 (Selleckchem) was used for final resuspension of the samples. Dissociated cells were pelleted at 400× g for 5 mins and then resuspended in a final volume of 500 µl with propidium iodide (PI) (Sigma) added to a concentration of 2µg/ml. Cell sorting was carried out with a 100 µm nozzle on an Influx instrument (BD Biosciences).
Instrument	BD Influx cell sorters (BD).
Software	Collection: FACS TM software suit (version 1.2.0.142) (BD) for influx sorters.
Cell population abundance	Abundance of distinct cell populations of interest was determined using appropriate negative controls and puritiy of sorted populations as determined by post sort reanalysis.
Gating strategy	Standard gating setting commonly utilized at flowcore facility of Monash University were used. Cell debris was excluded using a SSC vs FSC-H gate; aggregates were excluded then via a FSC-A vs FSC-H approach; dead cells were defined as PI high/positve and gated out; Live cell populations sorted were confirmed by post sort reanalysis. A figure exemplifying the FACS gating strategy is available at the Supplementary Figure 1 of the manuscript.

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