nature research

Corresponding author(s): DBPR- NCB-H37400A

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.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. 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
☐	☒	An indication of 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 statistics 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
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	SimplePCI software (Hamamatsu Corp. location) was used for image acquisition and analysis for the islet perifusion and calcium/ Rhodamine-123 imaging experiments
Data analysis	Tuxedo Suite tools (TopHat version 2.1.0, Cufflinks and Cuffdiff) and DESeq2 were used to analyze RNA-seq data. FlowJo v1.1.0 was used to analyze flow cytometry data. ImageJ v 1.49u was used to determine the length of the inner mitochondrial membrane. Statistical analyses were conducted with GraphPad Prism 7. Additional details are provided in the 'Methods' section. Bisulfite sequencing data were aligned and analyzed using the Sequencher software. SimplePCI software was used for image acquisition and analysis for the calicum and rhodamine-123 imaging experiments. R studio v1.1.453 was used to plot scatter plots of RNA-seq data. GSEA tool was used for gene set enrichment analysis and Morpheus was used for generating heatmaps.

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 upon request. 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

Materials, raw data and protocols that were used in this study are available from the corresponding author upon request. RNA-Sequencing data has been deposited to NCBI GEO (Accession GSE109795) for private viewing for the reviewers. The datasets will be made public upon acceptance and publication of the manuscript.

Field-specific reporting

Please select 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/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Statistical methods were not used to determine sample size. Sample sizes were chosen based on availability of experimental samples. These sample sizes are sufficient since the study is very thorough and we have used a variety of experimental techniques to confirm the findings
Data exclusions	Stem Cell differentiations that did not pass a minimum criteria of 85% SOX17+/FOXA2+ double positive cells at the early stage of defintive endoderm were excluded from further analysis. This criteria was preselected because the formation of definitive endoderm is the first step toward pancreatic differentiation. When cells don't pass the above criteria, the spheres do not form pancreatic endoderm efficiently and start falling apart on further culture.
Replication	Differentiations and analysis were repeated several times in our laboratory. We also sent our cells for functional analysis by other groups that specialize in beta-cell physiology in various parts of United States to ensure transparency. These groups have been included as authors on the manuscript. Hence, replication was successful.
Randomization	No Randomization was performed with samples. This is not applicable to our study. Since the study reports directed stem cell differentiation, randomization was not a factor. For the in vivo experiments, animals were randomnly chosen to get different experimental cells
Blinding	Our collaborators were blinded to the treatment groups in calcium signaling, mitochondrial energization, DNA methylation experiments and evaluation of mitochondrial inner membrane length.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☐	☒ Unique biological materials
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☐	☒ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Unique biological materials

Policy information about availability of materials

Obtaining unique materials	The human ES-cell line, Mel1 INSGFP/W used in the study is available from Prof. Ed Stanley at Murdoch Children's Research Institute. We are happy to provide an aliquot of the INS-GFP cells cultured in our

laboratory given appropriate MTAs are in place with Murdoch Children's Research Institute. We are also happy to train people in learning beta-cell differentiation.

Antibodies

Antibodies used

Antigen Species Dilution Manufacturer
Human C-peptide Mouse 1:200 Chemicon
Human C-peptide Rat 1:1000 DSHB
Insulin Mouse 1:1000 Sigma
Insulin Guinea pig 1:500 DAKO
Somatostatin Rat 1:500 Millipore
Glucagon Guinea pig 1:500 Linco
Glucagon Mouse 1:1000 Sigma
NKX6.1 Mouse 1:100 DSHB
NKX2.2 Mouse 1:20 DSHB
PDX1 Goat 1:200 R&D systems
SOX9 Rabbit 1:2000 Prestige (Sigma)
Amylase Rabbit 1:500 Sigma
Synaptophysin Mouse 1:200 BioGenex
CK19 Rat 1:1000 DHSB
Tomm20 Rabbit 1:1000 Santa Cruz

Antibody Manufacturer
Human PAX6- Alexa647 BD Bioscience
Islet-1- PE BD Bioscience
NKX6.1- Alexa647 BD Bioscience
NKX6.1-PE BD Bioscience
ChromograninA- PE BD Bioscience
NeuroD1- Alexa647 BD Bioscience
PDX1-PE BD Bioscience
Glucagon-Alexa647 Sigma (Mouse antibody)
conjugated in house
Somatostatin Dako (Rabbit antibody) at 1:250 dilution used
with Alexa-568 secondary

Validation

The primary antibodies for immunofluorescent staining were validated on Human pancreatic sections using various dilutions (including the manufacturer's recommended dilution). Primary antibody only control and secondary antibody only controls were used to evaluate and confirm specificity of the antibodies to the respective epitopes. Markers such as Insulin,C--peptide,glucagon,Somatostatin,PDX1,NKX6.1,NKX2.2, synaptophysin stained islet cells of the human pancreas. Amylase stained exocrine tissue whereas CK19 and SOX9 stained ducts. Tomm20 was specifically localized to compartments distinct from insulin compartment within the cells. Antibodies used for flow cytometry were similarly validated on human islet preparations. Specificity was confirmed to either alpha, beta or delta cells. Moreover, all of the antibodies used in this study have been used and reported in prior studies.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

Mel1 INSGFP/W human embryonic stem cells were used in the study and were obtained from S. J. Micallef and E. G. Stanley

Authentication

Cell line authentication was not done.

Mycoplasma contamination

We routinely check mycoplasma levels in human ESC lines in the lab. The cell lines tested negative for mycoplasma

Commonly misidentified lines
(See ICLAC register)

No cell lines used in this study were found in the database of commonly misidentified cell lines that is maintained by ICLAC and NCBI Biosample

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals

NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ mice (NSG) were obtained from Jackson Laboratories and bred in our facility. Male mice between the age group of 10-20 weeks were used in this study. Both genders were used

Wild animals

The study did not involve wild animals

Field-collected samples

Field samples were not collected for the study

Human research participants

Policy information about studies involving human research participants

Population characteristics	Human islet preparations from male and female cadavers were used in the study. The age ranged from 32 to 62 and BMI varied from 15.4 to 43.77. HbA1C values ranged between 5.1 to 6.3.
Recruitment	UCSF islet isolation facility and Prodo laboratories were involved in the recruitment of donors. We were not part of this process

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	Clusters at indicated stages were dissociated by adding accumax and incubating in a bead bath between 5-8 min. After incubating, samples were titurated to break apart the spheres. Cells were then fixed with 4% PFA for 15 min. Permeabilization and staining was carried out overnight in buffer containing triton-X and donkey serum for various intracellular markers. Samples were then washed with PBS and resupended in FACS buffer (2%BSA in PBS) for analysis
Instrument	LSRFortessa X20 Dual, Aria II
Software	FlowJo software
Cell population abundance	The main population isolated in the study is INS-GFP+ cells. Those account for 45-50% of starting population. The purity of the sort was confirmed by re-running the sample on the Aria II and determined to be >98%
Gating strategy	Supplementary Figure 1 has the gating strategy used in the study indicating where the boundaries between positive and negative cells were defined

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