nature research

Corresponding author(s): Henrik Semb

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	Gene set enrichment Analysis was done using GSEA Java application for loading the microarray datasets and finding enrichments comparing to C2 database of MsigDB. BD FACS Diva version 8.0.1 for analyzing FACS sorted data using machines BD FACS ARIA III and LSR FORTESSA.
Data analysis	1. Graphprism for Mac OS X version 6.0h 2. Microsoft Excel for Mac Version 16.16.2 3. FCS Express 6 4. Volocity 6.2 (PerkinElmer, Waltham, MA) 5. Image J 1.50f 6. IMARIS 9.0.2 (Bitplane)

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

- The accession code for Microarray experiment is E-MTAB-6891 (ArraryExpress).
- The list of figures that have associated are Extended data 10j&k.
- There are no restrictions on data availability.

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	Sample size was chosen based on previously published literature of the same subject. Biological replicates were atleast n = 3 unless otherwise noted. These are clearly showed with the dot-plot overlaying the bar graph. For in vivo study, sample sizes were chosen based on a evaluation of animal ethics versus ability to obtain sufficient samples for statistical analysis
Data exclusions	No data were excluded.
Replication	All attempts at replication of data contained in this manuscript were successful. Biological replicates were performed at least 3 times, unless otherwise indicated. Within one biological experiment, there are numerous technical samples (e.g. a few embryos/explants for in vivo experiments or 3 wells for technical replicate for cell culture). The mean +/- S.E.M or SD and student t-test were used to see if and how the difference between two parameters (e.g. untreated vs. treated) are significant. Significance were presented as stars, calculated by Graph Pad software and described under Methods.
Randomization	There were no randomization as there were no clinical studies or patient participants were involved. Animals were allocated to groups based on genotype.
Blinding	The investigators were not blinded to allocation during experiments or analysis, as group allocation was clearly visible in the samples due to phenotypic changes. The researchers knew the type of drug and concentration used to perform the study.

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	All unique materials used in this study are readily available from the authors

Antibodies

Antibodies used

Below is a summary of commercially available antibodies used in this study: or as cited in the results/Methods section.
Antibodies used for human PPs are as follows:
anti-PDX1 Goat polyclonal - R&D systems (AF2419) 1:500
anti-SOX9 - Millipore (AB5535) 1:500
anti-NGN3 - R&D systems (AF2419) 1:400
anti-FOXA2 - Santa Cruz (sc6554) 1:500
anti-SOX17 goat - R&D systems (AF1924)
anti- HNF6 Rabbit - Santa Cruz (sc13050) 1:200
anti-YAP mouse monoclonal - Santa Cruz (sc-101199) 1:200
anti-YAP rabbit polyclonal - Cell Signaling Technologies (4912) 1:200
anti-GFP Goat polyclonal - Abcam (ab5450) 1:500
anti-INSULIN - DAKO 1:1000
anti-GLUCAGON- Linco Research (4031-01) 1:1000
Biotinylated DBA - vector Labs (B-1035 1:500)
anti-YAP rabbit mAb - Cell Signaling Technology (D8H1X) (14074) 1:200
anti-Sox9 guinea-pig - BCBC (kind gift from Ole Madsen, Novo Nordisk) 1:2000
anti-Pdx1 guinea-pig - BCBC (AB2028); 1:1000
anti-Insulin guinea-pig - DAKO (A0564) 1:800
anti-Ecadherin rat polyclonal - Takara (M108-clone ECCD-2) 1:400
anti-Nkx6.1 mouse monoclonal - DSHB 1:500 (F55A12)
anti-pFAK (Tyr397) Rabbit monoclonal - Cell Signaling Technologies (8556) 1:1000
anti-FAK Rabbit monoclonal - Cell Signaling Technologies (13009) 1:1000
DAPI - life Tech (D1306) 1:1000
anti-glucagon mouse monoclonal - Sigma Aldrich (G2654) 1:500
anti-GFP chicken polyclonal - Abcam (ab13970) 1:1,000
Biotinylated DBA - Vector Labs (B-1035) 1:500
anti-Ngn3 Rabbit polyclonal - NIH BCBC (AB2011) 1:4000,
anti-Hnf1b Rabbit polyclonal - Santa Cruz Biotechnology (sc-22840)1:2000
anti-Muc1 Armenian Hamster - Thermo Fisher (MA5-11202) 1:200
anti-GFP Chicken polyclonal - Abcam (ab13970) 1;1000
Acti-stainTM 488/Phalloidin Cytoskeleton (PHDG1) 1:200
anti-Integrin alpha 5 Rabbit monoclonal - Abcam (ab150361) 1:1000
anti-Laminin Rat polyclonal - Acris (BM6046P) 1:200
anti-Fibronectin Rabbit polyclonal - Abcam (ab2413) 1;200
Alexa Fluor (AF) 488 Donkey anti-goat - Jackson ImmunoResearch (705-545-147) 1:500
Alexa Fluor (AF) 488 Donkey anti-mouse- Jackson ImmunoResearch (715-545-150) 1:500
Alexa Fluor (AF) 488 Donkey anti-guinea-pig - Jackson ImmunoResearch (715-545-148) 1:500
Alexa Fluor (AF) 594 Donkey anti-Goat IgG - Jackson ImmunoResearch (705-585-147) 1:500
Alexa Flour (AF) 647 Donkey anti-Rabbit IgG - Life Technologies (A-31573) 1:500
Cy5 Donkey anti-mouse IgG - Jackson ImmunoResearch (715-175-151) 1:500

Validation

Antibodies were validated as noted on manufacturer's website. Additional validation was done by the use of negative control and control tissue samples.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

Human embryonic stem cell lines used in this study were published:
1. PDX1-eGFP (Ameri et al Cell Reports, 2017 vo.19 (1) pp.36-49. Undifferentiated HUES4, obtained from D.A. Melton, Howard Hughes Medical Institute (Harvard University, Cambridge, MA).
2. NGN3-eGFP reporter (Zaah M Lof-Ohlin et al Nature Cell Biology 2017, vol.127 pp. 1313-1325). The targeted NGN3-reporter cell line was generated using DEF-SA121 cells (human embryonic stem cells from source cell line SA121- Cellartis by Takara Bio), using BAC recombineering.

Authentication

The cell lines were not authenticated

Mycoplasma contamination

The cell line has been tested negative for mycoplasma contamination

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

Species: Mus musculus. Strains: C57BL/6 and other genetically modified animals on a mixed background: Pdx1-tTA, tetO-Yap1, Pdx1Cre, Sox9CreERT2, Yap1fl/fl, Foxa2T2AiCre, Rosa26dnMaml1-eGFP, Ngn3-tRFP, Hes1-eGFP, ROSA26YFP. Gender was mixed and unknown as experiments were performed on embryos.

Wild animals	The study did not involve wild animals.
Field-collected samples	The study did not involve samples collected from the fields.

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	Differentiated human ES cell or E15.5 mouse pancreata were dissociated with Accutase or Liberase/0. 125% Trypsin, and then resuspended in FACS buffer to stop dissociation (PBS/BSA).
Instrument	Sony SH800, LSR Fortessa and BD FACS Aria III cell sorter
Software	Cells were sorted on a 4 laser (488nm, 561nm, 375nm and 633nm). Acquisition was done using BD FACSDiva for the Aria III and the Sony White Rose software for the SH800. Data was analysed using FCS Express 6.
Cell population abundance	For differentiated human ES cells, about 20-70% of the live cells. Purity of post-sort fractions were checked by briefly re-running a small amount of sample as well as QPCR form reporter marker (e.g. EGFP) and other signature genes (provided in the supplement). Additionally, eGFP beads (Clontech) were used to control consistent gating between experiments. For mice pancreata, due to extremely low numbers on the sorted populations, it was not possible to check purity on the cells after the sort. However, the flow cytometry core facility routinely does purity and Rmax measurements on the cell sorters using beads to ensure that the right populations are being sorted. E15.5 pancreata were sorted on a Sony SH800 directly into RLT buffer for immediate Qiagen microElute RNA purification according to manufacturer's instructions.
Gating strategy	Cells were first gated based on similar light scatter properties, doublets gated out using the width parameter on forward scatter, and then live cells gated based on 7AAD or DAPI. The gating for High, Low, or negative eGFP population were consistently set based on 6-peaks eGFP beads (Clontech). For conjugated Ab cell surface marker staining, negative (isotype control) cells were used as gating parameter. For E15.5 pancreata, negative fluorescent cells were used as gating for double reporter mice (GFP+/RFP+).

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