Electronic supplementary material (ESM)

ESM Methods

Bioinformatic clustering analysis

Single cell RNA-seq islet data were downloaded from Gene Expression Omnibus database (accession number GSE101099 [1], E17.5 timepoint, batches '1SW', '2SW' and 'Batch 2'; GSE155742 [2], E16.5 newly synthesised beta cells batches 'Green 1' and 'Green 2'; GSE128565 [3, 4], adult islets, batch 'No STZ', and GSE110648 [5], adult islets, 'B6_28wks_1' and 'B6_28wks_2' (chow fed mice)), filtered according to the respective published work and hypervariability of gene expression was assessed by variance analysis using the scran package v1.28.2 (https://bioconductor.org/packages/release/bioc/html/scran.html) [6]. This method avoids artefacts resulting from technical dropouts or of prioritising low abundance genes with large variances due to technical noise. The R package RaceID v0.3.0 (https://cran.r-project.org/web/packages/RaceID/index.html) [7] was used for further analysis. Mitochondrial genes were removed from the expression matrix prior to downstream analysis and cells were filtered to have at least 5000 transcripts with an expression value of >2. A distance matrix was then constructed using Pearson correlation as the distance metric. A K-nearest neighbour graph layout utilising the Fruchterman-Rheingold algorithm was performed to cluster cells into different groups, with a value of k = 3. This algorithm also identifies outlier populations, and so a total of six clusters were identified using the clustering method, with differential expression analysis performed comparing individual clusters to all other cells. Further analysis of E17.5 mouse scRNA-seq data in the form of Seurat R object was downloaded from https://doi.org/10.6084/m9.figshare.c.4158458 [1]. Normalisation was performed using the sctransform function [8] in Seurat v4 [9], followed by principal component analysis (PCA) and clustering in the top 30 PC space using the 'leiden' algorithm. The clusters were assigned their appropriate cell-types based on the expression of known cell-type specific marker genes. For visualisation purpose, cells were mapped to the UMAP space using

Seurat's built-in function, with extraction of islet clusters identifying two distinct subpopulations of beta cells. Based on this distribution, cells with *Nnat* expression >3 were designated as *Nnat*[+]. For geneset enrichment analysis (GSEA), differentially expressed genes were ranked according to their fold change. GSEA was performed for the Gene Ontology biological process terms (www.gsea-msigdb.org/gsea/index.jsp) [10-13] and (org.Mm.eg.db: Genome wide annotation for Mouse. R package version 3.8.2) using the R package ClusterProfiler [14]. The enrichment and pathway module plots were prepared using the DOSE [15] and enrichplot (package version 1.14.2. 2022) R packages respectively. Volcano plots were prepared using the R package EnhancedVolcano (R package version 1.12.0, https://github.com/kevinblighe/EnhancedVolcano).

Pseudo-time analysis

scRNA-seq datasets collected from islets during mouse embryogenesis (E12.5-E17.5) were obtained from [1] with Seurat v4.3 used for quality control and batch correction [16, 17]. The dataset was filtered to select for islet cells by visualising the expression of *Cdh1* and *Chga* as well as the lack of expression of *Vim* and *Col3a1* [1], which resulted in 4556 cells for downstream analyses. Following filtering, clustering was achieved by running the Louvain community detection algorithm on a shared nearest neighbours (SNN) graph generated from the single cell expression profiles [17]. Using the developmental and cell type markers identified in [1] we labelled each cluster as a known islet cell type or precursor of which 818 cells were identified as 'early' or 'late' beta cells. This also allowed us to evaluate the change in beta cell precursor/state proportion across embryonic stages. After labelling, we used the Monocle3 package to identify cell fate trajectories using a pseudo-time analysis [18]. After defining the development trajectory of beta cells during embryogenesis, we integrated our data with a dataset from adult mouse islets [3, 4].

Study approval

All animal procedures were in accordance with the UK Animals (Scientific Procedures) Act 1986 and approved by the UK Home Office (PPL: PP4712156). All animal studies have been approved by the Imperial College Animal Welfare and Ethical Review Body. Animals were euthanised by cervical dislocation. Animal work was designed and reported according to the Animal Research: Reporting of In Vivo Experiments (ARRIVE) guidelines. Power calculations were based on known effect sizes. Mutant mouse lines and the appropriate control mice were housed together to minimise cage bias. We did not use randomisation to allocate mice due to the absence of experiments requiring treatment groups prior to islet isolation. Where possible, blinding for experimental mice and tissues was used and all studies were replicated in at least three independent cohorts or using litters from three different breeding cages.

Animal models

Transgenic mice lines expressing *Cre* recombinase under the control of the rat insulin promoter (RIP) [19] and with a tdTomato reporter downstream of a stop codon flanked by *loxP* sites [20] have been described previously. Mice with *Nnat*-driven eGFP expression were purchased from the Mutant Mouse Resource and Research Centre (MMRRC, USA) repository (Tg(Nnat-eGFP)EA106Gsat/Mmucd, #010611-UCD) with the presence of the *Egfp* allele assessed by standard PCR using DNA from ear biopsy (primers 5'-CATCACCCCTCCTTCTCAAC-3' and 5'-GAACTTCAGGGTCAGCTTGC-3'). Mice with global deletion of *Nnat* were used as previously described [21]. Conditional *Dnmt3a* null mice [22], generated by crossing to *Pdx1-Cre* [23] have been described elsewhere [24-26]. Mice containing the *Dnmt3a* floxed allele but lacking *Pdx1-Cre* were used as controls. Islet transcriptomic data from 7 week old Zucker diabetic fatty (ZDF) rats vs ZDF/+ controls was according to [27]. Similarly, islet transcriptomic data in the mouse *ob/ob* strain vs *+/ob* control mice (4 and 10 weeks old on a C57BL/6 background) was obtained from a publicly available source from the Alan Attie laboratory (http://diabetes.wisc.edu/search). 6-8 week old *db/db* and control C57BLKS/J mice on the BKS background were obtained from the Jackson

Laboratory (JAX 000642 and JAX 000662, respectively). All other mouse strains were on a C57BL/6J background and experimental cohorts were group-housed and maintained on a 12-hour light/dark cycle. Animals were fed standard mouse chow (RM3, Special Diet Services) with access to drinking water *ad libitum* in specific pathogen-free barrier facilities. Male mice were 10-12 weeks old at the time of experiments unless otherwise stated. For experiments involving the *Nnat*-eGFP reporter line, mice were hemizygous for the reporter via paternal transmission, unless otherwise stated. Sperm was isolated from young adult male C57BL/6J mice.

Primary islet isolation and FACS

To isolate primary islets, the pancreas was inflated with Liberase TM (Roche), digested at 37°C, purified using a Histopaque 1119/1083/1077 gradient (Sigma-Aldrich) and hand-picked. For FACS-based experiments, cells from purified islets were isolated immediately following isolation using Accutase (Sigma-Aldrich) for 5 min. at 37°C followed by gentle disaggregation by pipetting. Islet cells were centrifuged and resuspended in FACS buffer (25 mmol/l HEPES, 5 mmol/l EDTA, 1% fetal bovine serum (FBS) in PBS, pH 7.4) prior to cell sorting using an FACSAria III flow cytometer (BD Biosciences, UK) equipped with 488 nm and 561 nm lasers for excitation of eGFP and tdTomato, respectively. Emission from eGFP and tdTomato expressing cells was collected using a 530/30 and 580/30 bandpass filter, respectively. DAPI (0.2 μg/mL) was used to exclude dead cells and was excited using a 405 nm laser with emission collected using a 450/50 bandpass filter. Forward scatter area *vs* forward scatter height was used to offset and remove doublets. Total insulin content in FACS-sorted populations was assessed using ultra-sensitive insulin homogeneous time-resolved fluorescence (HTRF) assay kits (Cisbio).

Intracellular calcium imaging

Pancreatic islets were isolated as above and left to settle overnight in RPMI medium (Gibco) supplemented with 10% (v/v) FBS, (Sigma-Aldrich), 2 mM L-glutamine and 100 U/mL

penicillin/100 μg/mL streptomycin (both Gibco) at 37°C and 5% CO_2. For experiments involving reporter mice expressing *Nnat*-eGFP, Ca^{2+} imaging of whole islets was performed after loading the cytosol with 2 μmol/l Cal-590 AM (Stratech) in Krebs-Ringer bicarbonate buffer (140 mmol/l NaCl, 3.6 mmol/l KCl, 0.5 mmol/l NaH_2PO_4, 25 mmol/l $NaHCO_3$ (saturated with CO_2), 1.5 mmol/l $CaCl_2$, 0.5 mmol/l $MgSO_4$, 10 mmol/l HEPES, pH 7.4) containing 3 or 11 mmol/l glucose. Images were captured at 0.5 Hz on an Axiovert microscope (Zeiss, Germany) equipped with a 10x 0.3–0.5 NA objective, a Hamamatsu ImagEM camera coupled to a Nipkow spinning-disk head (CSU-10, Yokogawa UK) and illuminated at 490 nm or 530 nm. Data were analysed using ImageJ and changes in fluorescence intensity over time were measured after drawing a region of interest around individual cells for each islet. Pearson-based connectivity and correlation analyses were performed as previously described [28]. The first 50 datapoints were removed to account for the initial spike in response to the addition of 11 mmol/l glucose. Signal binarisation-based connectivity and correlation analyses were performed as previously described [29]. Connectivity between cells was explored using Pearson's R-based correlation analyses in MATLAB as described [25]. We note that although the Ca^{2+} dye used here was not selectively targeted towards beta cells, at the concentration of glucose used (11 mmol/l), pancreatic alpha cells are expected to be largely inactive. We do not exclude a contribution of delta cells, although these represent <5% of the overall endocrine cell population.

For experiments involving global *Nnat* null mice [21], islets were incubated at 37°C for 40 min in Krebs-Ringer bicarbonate buffer containing 4.5 µmol/l Cal-520 AM (Stratech) and either 3 or 11 mmol/l glucose, and then transferred into an imaging chamber. Imaging was performed on a Nikon Eclipse Ti microscope equipped with a 40x/1.2 NA oil objective and an ibidi heating system. Cal-520 AM was excited with a 491 nm laser line and emitted light filtered at 525/50 nm. Images were acquired with an ORCA-Flash 4.0 camera (Hamamatsu) and Metamorph

software (Molecular Device) was used for data capture, in streaming mode and with an interval of 230 ms.

Pearson's R-based connectivity and correlation analyses in an imaged islet were performed as previously described [28]. The first 50 datapoints were removed to account for the initial spike in response to the addition of 11 mmol/l glucose. Ca^{2+} signals were first smoothed by fitting them to a smoothing spline [30]. The smoothing spline s was constructed and minimised thus:

$$p \sum_i w_i \big(y_i - s(x_i)\big)^2 + (1-p) \int \left(\frac{d^2 s}{dx^2}\right)^2 dx$$

where p = specified smoothing parameter between 0 and 1, w = specified weight

Once the data was fitted to a smoothing spline model, a baseline trend was then set according to the means of each individual fluorescence trace. A 20% threshold above this baseline was then imposed to minimise false positives from any residual fluctuations in baseline fluorescence. Cell signals with deflection above the de-trended baseline were represented as '1' and inactivity represented as '0', thus binarising the signal at each time point. The coactivity of every cell pair was then measured as:

$$C_{ij} = \frac{T_{ij}}{\sqrt{T_i T_j}}$$

where T_{ij} = total coactivity time, T_i and T_j = total activity time for two cells

The significance at $p<0.001$ of each coactivity measured against chance was assessed by subjecting the activity events of each cell to a simulation with 10,000 iterations obtained through shuffling the existing data. Synchronised Ca^{2+}-spiking behaviour was assessed by calculating the percentage of coactivity using the binarised cell activity dataset. A topographic

representation of the connectivity was plotted in MATLAB (available from: https://www.mathworks.com/matlabcentral/fileexchange/24035-wgplot-weighted-graph-plot-a-better-version-of-gplot) with the edge colours representing the strength of the coactivity between any two cells. A 95% threshold was imposed to determine the probability of the data, which was then plotted as a function of the number of connections for each cell to determine if the dataset obeyed a power-law relationship [31] (available from: https://www.mathworks.com/matlabcentral/fileexchange/29545-power-law-exponential-and-logarithmic-fit). Finally, the data and figures were written from MATLAB to Microsoft Excel files (available from: https://www.github.com/michellehirsch/xlswritefig) for easy visualisation and dissemination.

Histological techniques and immunofluorescence

Dissected tissues were fixed in 4% paraformaldehyde at 4°C, washed with PBS and cryoprotected overnight at 4°C using 30% (w/v) sucrose in PBS before embedding in optimal cutting temperature (OCT) and stored at -80°C. Sections (10 μm) were cut using a CM1950 Cryostat (Leica) and permeabilised with 0.2% Triton X-100 in PBS, followed by blocking in 2% (w/v) BSA and 5% (v/v) chicken serum in PBS containing 0.1% Tween-20. Sections were immunostained using primary antibodies diluted in blocking buffer against NNAT (clone EPR13554(B), ab181353 from Abcam, 1:1000), GFP (GFP-1010, Aves Labs, 1:3000) INS (clone K36AC10, I2018, Sigma-Aldrich, 1:5000), somatostatin (SST) (13-2366, American Research Products, 1:1000), glucagon (GCG) (clone K79bB10, G2654, Sigma-Aldrich, 1:3000) urocortin 3 (UCN3) (a kind gift from Joan Vaughan and Paul E. Sawchenko, The Salk Institute for Biological Studies, PBL #7218 [32] 1:1000), translocase of outer mitochondrial membrane (TOM20) (FL-145, sc-11415, Santa Cruz Biotechnology, 1:250), DNMT1 (clone 60B1220.1, NB100-56519, Novus Biologicals, 1:500) and DNMT3A (clone 64B1446, NB120-13888, Novus Biologicals, 1:500) with overnight incubation at 4°C followed by incubation with secondary antibody (Alexa Fluor, Thermo Fisher Scientific, 1:500) and counterstaining with

DAPI (0.1 μg/mL). Sections were imaged using a TCS SP5 confocal microscope (Leica, Germany), capturing multiple islets from at least three regions of tissue >100μm apart. Primary antibodies were first assessed individually by comparative staining *vs* IgG controls. For immunostaining of paraffin-embedded sections from *db/db* mice and mice with beta cell-selective deletion of *Dnmt3a*, slides were immunostained according to [26, 33, 34] using antigen retrieval with Tris-EDTA buffer (10 mmol/l Tris base, 1 mmol/l EDTA, 0.05% Tween-20, pH 9.0) and antibodies against NNAT (ab27266, Abcam, 1:200) and INS (ab195956, Abcam, 1:800).

Human pancreata were obtained from 9 non-diabetic (ND) heart-beating organ donors (mean glycemia during intensive care unit stay: 147±18 mg/dL (mean±SEM, see ESM Table 4 for details)), received before November 2021 and processed as previously described to assess pancreatic islet cell features [35, 36], with the approval of the Ethics Committee of the University of Pisa, upon written consent of donors' next-of-kin. Tissue samples were obtained from the body of the pancreas close to the head. Specimens (approximately 4-5 mm^3) were cleaned of any residual fat tissue, wiped and fixed for 24h either in 4% paraformaldehyde or 10% buffered formalin at 4°C. Samples were washed three times in PBS and stored in 70% ethanol solution at 4°C until further processing steps and paraffin embedding. Sequential sections (2 µm thickness) were prepared, and H&E staining performed on one section from each tissue block to observe the overall morphology of the pancreatic tissue and evaluate the quality (adequate and uniform fixation, preservation of tissue structures, absence of artifacts such as shrinkage, tearing or crumbling) before proceeding with immunofluorescence staining. H&E staining was performed using the automated stainer Dako CoverStainer (Agilent). Sections were pre-heated at 60°C for 15 min, deparaffinised with Histoclear (Sigma Aldrich) and hydrated through passages in a decreasing alcohol series. The staining was performed by incubation of the sections with Hematoxylin (Agilent) for 3.5 min followed by rinsing in distilled water for 1 min, treatment with Bluing Buffer (Agilent) for 1 min, and rinsing in tap

water for 1 min and ethanol 95% for 1 min, followed by incubation with Eosin (Agilent) for 4 min. After dehydration through passages in increasing alcohol series, sections were mounted using mounting medium (Agilent), covered with a coverslip and allowed to dry. Immunosections were deparaffinised with xylene, rehydrated through a graded alcohol series and antigen retrieval carried out using a citrate-based unmasking solution (pH 6.0, Vector Laboratories). Sections were then permeabilised with 0.4% Triton X-100 in PBS, followed by blocking in 2% (w/v) BSA and 5% (v/v) chicken serum in PBS containing 0.1% Tween-20. Sections were immunostained using primary antibodies diluted in blocking buffer against neuronatin (NNAT) (PA5-115646 from Thermo Fisher Scientific, 1:1000), insulin (INS) (IR002, Dako, 1:100), somatostatin (SST) (13-2366, American Research Products, 1:1000) and glucagon (GCG) (clone K79bB10, G2654, Sigma-Aldrich, 1:3000) with overnight incubation at 4°C followed by incubation with secondary antibody (Alexa Fluor, Thermo Fisher Scientific, 1:500) and counterstaining with DAPI (0.1 mg/mL). Experiments to validate primary antibodies targeting NNAT (PA5-115646 from Thermo Fisher Scientific) used for experiments involving human pancreatic sections were carried out in HEK293T cells (cultured in DMEM, 10% FBS, 2 mmol/l L-glutamine, 37 °C, 5% CO_2 on poly-L-lysine coated coverslips). Cells were transfected with plasmids expressing human *NNAT* (Origene, RC200328) using Lipofectamine 2000. Transfected cells were fixed with 4% paraformaldehyde and immunostained as above using primary antibodies targeting NNAT (PA5-115646 from Thermo Fisher Scientific, 1:500). Transfected cells were visualised using co-transfection with a GFP-expressing plasmid (a kind gift from Connie Cepko, Addgene #11150 [37]). Sections were imaged using a TCS SP5 confocal microscope (Leica, Germany).

Cell classification was carried out with a custom-written Fiji script [38] and CellProfiler pipeline (software version 4.1.3). Nuclei segmentation was batch processed in Fiji using the Stardist plugin, with subsequently generated nuclei label images imported into CellProfiler for cell type classification following measurement of fluorescence intensity in acquired fluorescent channels.

Cell culture, RNA silencing and GSIS

Human EndoC-βH1 [39] and rat INS1E [40] beta cells were mycoplasma free and cultured as previously described [21]. EndoC-βH1 beta cells were incubated with lentiviruses expressing an *NNAT*-targeting shRNA (clone ID: TRCN0000163961) *vs* scrambled shRNA control (multiplicity of infections (MOI) = 2, MISSION transduction particles, Sigma-Aldrich) for 72h in complete media without BSA. INS1E cells were transfected with Silencer Select siRNAs (30 nmol/l; *Nnat*, s137247; *Sec11a*, s80747; *Spcs1*, s226238; *Spcs2*, s18919; *Spcs3*, s184336; all Ambion) using Lipofectamine 2000 (Invitrogen) in OptiMEM (Gibco) for 48h. GSIS assays were performed as previously described [39]. Briefly, EndoC-βH1 cells were cultured in low (2.8 mmol/l) glucose complete media for 16h prior to GSIS assay. Cells were incubated in Krebs-Ringer bicarbonate buffer (115 mmol/l NaCl, 5 mmol/l KCl, 24 mmol/l NaHCO$_3$ (saturated with CO$_2$), 1 mmol/l CaCl$_2$, 1 mmol/l MgSO$_4$, 10 mmol/l HEPES, 0.2% BSA, pH 7.4) containing 0.5 mmol/l glucose for 1h prior to incubation with low (3 mmol/l) or high (16.7 mmol/l) glucose for 1h. Cells were collected in lysis buffer (20 mmol/l Tris-HCl pH 7.5, 137 mmol/l NaCl, 0.1% Triton X-100, 2 mmol/l EGTA, 1% glycerol with protease inhibitors (Complete Mini from Roche)), centrifuged for 5 min at 3,000 rpm and 4°C, with supernatants and collected media stored at -80°C until insulin quantification with ultra-sensitive insulin HTRF assay kits (Cisbio).

Immunoprecipitation and mass spectrometry

Cells were homogenised in protein lysis buffer (50 mmol/l Tris-HCl pH 7.5, 150 mmol/l NaCl, 1% Triton X-100, 1 mmol/l EDTA with protease inhibitors, Complete Mini from Roche), solubilised for 30 min on ice, clarified and normalised for protein content. Supernatants were incubated with protein A-agarose beads (Calbiochem) and antibodies against NNAT (ab27266, Abcam, 1:500) or rabbit IgG overnight at 4°C. Agarose beads were washed seven times with protein lysis buffer, with the final three washes not containing Triton X-100. For mass spectrometry, beads were sedimented and resuspended in 100 μL of 20 mmol/l HEPES

(pH 8.0), 1mol/l urea, 1 mmol/l DTT and 1.5 µg of trypsin gold (Promega), transferred to protein lo-bind tubes, and digested sequentially for 30 min and overnight at 37 °C (EndoC-βH1) or overnight at 37 °C only (INS1E). Digests were subsequently acidified with trifluoroacetic acid, followed by de-salting using reversed-phase spin tips (Glygen Corp) according to the manufacturer's recommendations. Eluents were subsequently dried to completion using vacuum centrifugation. Dried peptides were re-suspended in 0.1% trifluoroacetic acid (TFA) and analysed by LC-MS/MS analysis using an Ultimate 3000 nano-HPLC coupled to a Q-Exactive mass spectrometer (Thermo Scientific) via an EASY-Spray source. Peptides were loaded onto a trap column (Acclaim PepMap 100 C18, 100µm × 2cm) at 8 µL/min in 2% acetonitrile, 0.1% TFA. Peptides were then eluted on-line to an analytical column (Acclaim Pepmap RSLC C18, 75 µm × 75 cm). A stepped gradient separation was used with 4-25% of 75% acetonitrile, 0.1% formic acid for 75-90 min, followed by 25-45% for 10-30 min, with subsequent column conditioning and equilibration. Eluted peptides were analysed by the Q-Exactive operating in positive polarity and data-dependent acquisition mode. Ions for fragmentation were determined from an initial MS1 survey scan at 70000 (EndoC-βH1) or 120,000 resolution (INS1E), followed by higher-energy collisional dissociation of the top 12 (EndoC-βH1) or 20 (INS1E) most abundant ions at a resolution of 17,500 (EndoC-βH1) or 15,000 (INS1E). MS1 and MS2 scan AGC targets set to 3e6 and 5e4 for a maximum injection time of 50ms and 50ms (EndoC-βH1) and 25ms and 110ms (INS1E), respectively. A survey scan m/z range of 400–1800 m/z (EndoC-βH1) or 350-1750 (INS1E) was used, with a normalised collision energy set to 27%, minimum AGC of 1e3, charge state exclusion enabled for unassigned and +1 ions and a dynamic exclusion of 30 (EndoC-βH1) or 50 (INS1E) seconds. Data was processed using the MaxQuant [41] software platform (v1.6.1.0) for EndoC-βH1 and (v1.6.10.43) for INS1E with database searches carried out by the in-built Andromeda search engine against the Swissprot Homo Sapiens database (downloaded 4th January 2018, entries: 20,256) or Ratttus Norvegicus (Downloaded 31st May 2022, entries: 22,831). A reverse decoy database approach was used at a 1% FDR for peptide spectrum

matches and protein identifications. Search parameters included: maximum missed cleavages set to 2, variable modifications of methionine oxidation, protein N-terminal acetylation, asparagine deamidation and cyclisation of glutamine to pyroglutamate. Label-free quantification was enabled with an label-free quantification (LFQ) minimum ratio count of 2 (EndoC-βH1) or 1 (INS1E). For post processing of raw protein intensities for experiments with EndoC-βH1, technical intensities per sample were averaged, with the matrix containing an intensity column per replicate for each digest and condition. Intensity per digest was then generated by averaging raw intensities, with the matrix containing an intensity per digest for each condition (four columns). Intensities per digest for each condition were then summed, with the final matrix containing an intensity per condition. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [42] partner repository with the dataset identifier PXD048465.

Western immunoblotting, RT-PCR, RNA sequencing

For Western blotting, EndoC-βH1 cells were homogenised in protein lysis buffer as above, clarified and normalised for protein content using the BCA method (Bio-Rad) as previously described [21]. Briefly, proteins were transferred from SDS-PAGE to PVDF and immunoblotted using antibodies against NNAT (ab27266, Abcam, 1:2000) and β-tubulin (clone 9F3, #2128, Cell Signaling, 1:5000), HRP-conjugated secondary antibodies (Cell Signaling, 1:3000) and Immobilon Forte Western HRP substrate (Millipore). Scanned blots were assessed by densitometry using Image J. For RT-PCR, mRNA was purified using an Allprep DNA/RNA/protein mini kit or by direct homogenisation into Trizol reagent with purification using chloroform and RNeasy kits (both Qiagen). Normalised quantities of mRNA were reverse transcribed, and cDNA assessed by quantitative RT-PCR using Taqman reagents on a QuantStudio 7 Real Time PCR cycler with *Hprt*/*HPRT1* as an internal mRNA control (all Applied Biosystems, USA, *NNAT* Hs00193590_m1; *HPRT1* Hs02800695_m1; *Egfp* Mr04097229_mr; *Mafa* Mm00845209_s1; *Neurod1* Mm01280117_m1; *Nkx6.1*

Mm00454962_m1; *Pdx1* Mm0435565_m1; *Nnat* Mm00440480_m1; *Hprt* Mm00446968_m1; *Ins2* Mm00731595_gH; *Slc2a2* (*Glut2*) Mm00446229_m1; *Sst* Mm00436671_m1; *Gcg* Mm00801714_m1).

For RNA sequencing, RNA from FACS-purified cells (four biological replicates each from an individual mouse) was quantified and assessed for integrity (RNA integrity number > 9) using a Bioanalyzer 2100 and a RNA 6000 pico assay (Agilent Technologies). and 20 ng of RNA was processed for RNA sequencing using a NEBNext Ultra II Directional RNA library Prep Kit for Illumina paired with a Poly(A) mRNA magnetic isolation module and AMPure XP SPRIselect beads (Beckman Coulter), all according to manufacturer's instructions. Libraries were assessed on a Bioanalyzer 2100 using a High-Sensitivity DNA assay (Agilent Technologies), with subsequent cDNA quantification using a Quant-iT dsDNA High Sensitivity Assay kit (Thermo Fisher Scientific). Libraries were sequenced using a NextSeq 500 High Output sequencer (Illumina, USA), with 2 x 75 bp length at 50 million reads per sample. Quality assessment of paired end reads was performed using the FASTQC tool (version 0.11.2, http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and raw reads were aligned to the mouse genome (mm10) using STAR aligner (version 2.7.1a) [43]. Gene counts were obtained using STAR and gencode gene annotations (vM25) and differential expression analysis was performed using the DESeq2 Bioconductor package (version 1.32.0) [44]. Differentially expressed genes were obtained by comparing NNAT[+] beta cells, NNAT[-] beta cells and the non-beta cell fraction. Gene set enrichment analysis was performed using GSEA (version 4.1.0) [10-13] and the Molecular Signatures Database v7.2.0, with Wald statistical analysis from the differential expression data used to create preranked files as an input to the GSEA preranked analysis.

Bisulphite sequencing

Genomic DNA was extracted from FACS-purified cells using Allprep DNA/RNA/protein mini kits from Qiagen. Sperm cells were lysed using differential extraction, first by incubation in 100

mmol/l Tris-HCl pH 7.5, 200 mmol/l NaCl, 5 mmol/l EDTA, 0.2% SDS, 0.1 mg/mL proteinase K at 55°C for 2h to lyse epithelial cells, followed by pelleting and resuspension of sperm cells in the above lysis buffer plus 10 mmol/l dithiothreitol (DTT) and subsequent DNA extraction using phenol/chloroform/isoamyl alcohol (25:24:1) and standard isopropanol precipitation. Bisulphite conversion and desulphonation was performed using an EZ DNA Methylation-Gold kit from Zymo Research. Eluted DNA was used for amplification of PCR products with primers spanning the *Nnat* promoter (5'-TTTAGGTGGTAAGAGGGTATTTAAGGTA-3' and 5'-AATACATACTCACCTACAACA-3') and gametic DMR (5'-TTGATTGGTGGATAAGTTGTGTTT-3' and 5'-CCACCCTTAAAAAAATACCCATAAT-3'), designed by BiSearch [45] and an EpiTaq hot start kit (TaKaRa Bio). PCR fragments were gel purified (Qiagen), ligated into pJET1.2/blunt using a CloneJET PCR cloning kit (Thermo Scientific) and antibiotic-resistant plasmids prepared in 96 well format using a Wizard SV 96 plasmid DNA purification system (Promega) before Sanger sequencing (Genewiz, now Azenta Life Sciences). Sequencing data was quantified as percentage of methylated CpGs across >12 clones, using Bisulphite Sequencing DNA Methylation Analysis (BISMA) software [46].

Statistical analysis

Data are shown as mean±SEM in all panels. Data were assessed for normal distribution using the D'Agostino-Pearson normality test with statistical significance between biological groups determined using GraphPad Prism 9 with ANOVA, 2-tailed Student's *t* test (parametric) or Mann-Whitney *U*/Kruskal-Wallis tests (nonparametric) used where appropriate. Sidak's (or Dunn's as nonparametric equivalent) post hoc tests were utilised to correct for multiple corrections. Paired analyses were assessed using paired Student's *t* tests or the Wilcoxon matched-pairs signed rank test. <5% error probability was considered significant (i.e., $p < 0.05$) and further statistical information, such as *n* numbers and *p* values, are provided in the figure legends.

ESM results

Heterogeneous NNAT expression in beta cells during postnatal development is associated with altered CpG methylation

To extend the findings in Fig. 1-4 using an orthogonal approach, and to facilitate subsequent functional studies, we utilised a Bacterial Artificial Chromosome (BAC) reporter mouse line expressing eGFP under the control of the *Nnat* promoter and upstream enhancers. Aligning with studies of the endogenous gene (Fig. 1-4), adult reporter mice expressed eGFP in a subpopulation (~15% of total) of beta cells (Fig. 5a). eGFP expression was allele-specific with expression from the paternal, but not maternal allele, indicating the preservation of *Nnat* imprinting in the reporter (ESM Fig. 10a, b). Furthermore, eGFP expression was restricted to tissues known to express *Nnat* endogenously (pancreatic islets, hypothalamus) with no evidence of expression in *Nnat*-deficient tissue systems (liver, kidney, muscle) (ESM Fig. 10a-c). Possession of the transgene did not affect endogenous *Nnat* expression (ESM Fig. 10b) nor the expression of key beta cell functional genes, and eGFP staining overlapped with endogenous NNAT reactivity (ESM Fig. 10d, e). Furthermore, through multiple (>6) generations of mice, eGFP reporter expression was silenced via maternal inheritance and re-established through paternal inheritance. Thus, *Nnat* imprinting status for the transgene was restored transgenerationally via the germline. Taken together, these data demonstrate that the *Nnat*-eGFP reporter mouse line mimics endogenous *Nnat* expression.

The NNAT⁺ beta cell population shows impaired glucose-stimulated Ca^{2+} dynamics and is de-enriched for highly connected 'hub' cells

To determine whether NNAT may influence beta cell connectivity and membership of the 'hub' cell subgroup [47], we studied glucose-induced Ca^{2+} dynamics in *Nnat* deficient (*Nnat*⁺/⁻ᵖ) islets [21] by high-speed confocal imaging of the Ca^{2+} probe Cal-520 AM and off-line analysis using Pearson's R-based correlation analyses [48]. Ca^{2+} increases in response to high (11 mmol/l)

glucose were significantly higher in $Nnat^{+/-p}$ than control islets (ESM Fig. 14a). The mean Pearson's coefficient of correlation in wild type islets under low glucose conditions was 0.72±0.05, a correlation not significantly different when compared with high glucose conditions (0.79±0.03) (ESM Fig. 14a, b). Similarly, the mean Pearson's coefficient of correlation in $Nnat$ deficient islets under low glucose conditions was 0.70±0.039, with this correlation increased significantly under high glucose conditions (0.80±0.020; p=0.03) (ESM Fig. 14a, b). Connectivity was further analysed using signal binarisation-based connectivity and correlation analyses. Surprisingly, the pooled data over 31 wild type islets under low glucose conditions did not conform to a scale-free network, with 53.3% of all beta cells identified to be highly connected (R^2=0.17) (ESM Fig. 14c, d). However, these same 31 wild type islets displayed an obedience to a power-law distribution only when in under high glucose conditions, where a small proportion of beta cells (15.4%) were identified to be highly-connected 'hub' cells, as assessed by counting the number of beta cells displaying coordinated Ca^{2+} responses with at least 50% of all beta cells (R^2=0.003), and this behaviour was similar in the $Nnat$ deficient islets (ESM Fig. 14c, d). Pooled data over 38 $Nnat$ deficient islets revealed an adherence to a power law only in the high glucose condition, where 9.9% of beta cells were identified as 'hub' cells (R^2=0.11). This is in contrast to the 51.7% of identified highly-connected beta cells in the low glucose condition (R^2=0.15) (ESM Fig. 14c, d).

The mean connectivity of highly-connected wild type beta cells in low glucose was significantly higher than 'follower' cells (70.2±0.6% vs 14.1±0.9%; p<0.001) (ESM Fig. 14e). This was true also under high glucose conditions (66.7±0.9% vs 13.4±0.5%; p<0.001) (ESM Fig. 14e). When looking at $Nnat$ deficient islets, beta cell 'hubs' were also significantly more connected than 'followers' in both low glucose (69.1±0.5% vs 15.3±0.8%; p<0.001) and high glucose (59.7±0.6% vs 12.2±0.5%; p<0.001) (ESM Fig. 14e). Interestingly, while there were no significant differences in the mean connectivity between the highly-connected wild type vs $Nnat$ deficient 'hub' cells in low glucose (p=0.16), the wild type 'hub' cells were significantly more connected than the $Nnat$ deficient 'hub' cells in high (11 mmol/l) glucose (p<0.001) (ESM

16

Fig. 14e). The proportion of highly-connected 'hubs' to 'followers' were not significantly different between the wild type and *Nnat* deficient groups (ESM Fig. 14f). However, this proportion decreased significantly in both conditions under high glucose (ESM Fig. 14f). In wild type islets, the 53.5% of highly-connected beta cells decreased significantly to only 13.6% when in high glucose ($p<0.001$) (ESM Fig. 14f). Similarly, in the *Nnat* deficient islets, the 53.1% of highly-connected beta cells also decreased significantly to 9.3% under high glucose conditions ($p<0.001$) (ESM Fig. 14f). Interestingly however, there were no significant differences in the proportion of 'hub'/'follower' cells between the wild type and *Nnat* deficient islets under high glucose conditions ($p=0.36$) (ESM Fig. 14f). Thus, NNAT is not involved in setting beta cell connectivity but likely reflects a marker of less well-connected cells.

We further explored this question using islets from *Nnat*-eGFP reporter mice using the red-shifted calcium probe Cal-590 AM [25]. *Nnat*-GFP+ and *Nnat*-GFP- cells responded similarly to challenge with high (11 mmol/l) glucose (Fig. 8a). Beta cell-beta cell connectivity (Fig. 8b) was not different between NNAT+ and NNAT- populations and considered across the whole population, the mean Pearson's coefficient of correlation was 0.87±0.01 (Fig. 8c), with 14.9% of beta cells being significantly connected.

Pooled data over 45 islets uncovered a scale-free network topography (Fig. 8d) where a small proportion of beta cells (4.4%) were identified as 'hub' cells (Ca^{2+} responses with at least 30% of all beta cells; $R^2=0.31$). Of the islets analysed, 34 (75.6%) displayed obedience to a power law, and this adherence pertained for both NNAT+ (Fig. 8d; $R^2 = 0.38$) and NNAT- populations ($R^2=0.30$). Cells could be categorised as belonging to one of four distinct groups (Fig. 8e). Of all cells examined, 4.27% were NNAT- 'hub' cells connected to an mean of 39.0% of all beta cells (Fig. 8e, f). The fraction of these (1.15%, Fig. 8e) identified as NNAT+ 'hub' cells had an mean of 38.3% coactivity with other cells (Fig. 8f). There were no significant differences in connectivity between these two groups (Fig. 8f; $p=0.76$) nor in connectivity between NNAT+ and NNAT- 'follower' cells ($p=0.72$). However, when considered within each individual islet, a

significantly lower ratio of 'hubs' to 'followers' was observed in NNAT$^+$ *versus* NNAT$^-$ cells (Fig. 8g). Overlap with 'leader' cells [29, 48] was not explored in the absence of readily identifiable Ca^{2+} waves.

ESM Tables

ESM Table 1: Analysis of highly variable genes between individual beta cells at the late embryonic and adult stage

ESM Table 2: Transcriptomic assessment of *Nnat$^+$* vs *Nnat$^-$* beta cells at E17.5

ESM Table 3: Unbiased clustering analysis of individual beta cells at the late embryonic and adult stage

Unique identifier	ID	Sex	Age	Body mass index (BMI, KG/m^2)
18/40	HP 25/08/2012	M	17	24.5
26/88	HP 11/08/2020	F	15	24.7
27/119	HP 22/11/2021	M	18	20.4
20/46	HP 25/08/2014B	M	15	19.5
20/60	HP 27/09/2014	M	13	25.4
18/29	HP 20/06/2012	F	60	20.8
17/71	HP 18/11/2011	M	71	25.4
17/17	HP 18/03/2011	M	78	24.2
16/58	HP 20/12/2010	F	75	23.4

ESM Table 4: Donor characteristics for human pancreatic sections

ESM Table 5: Bulk RNA-seq analysis of FACS-sorted adult NNAT$^+$ and NNAT$^-$ beta cells

ESM Figures and Figure Legends
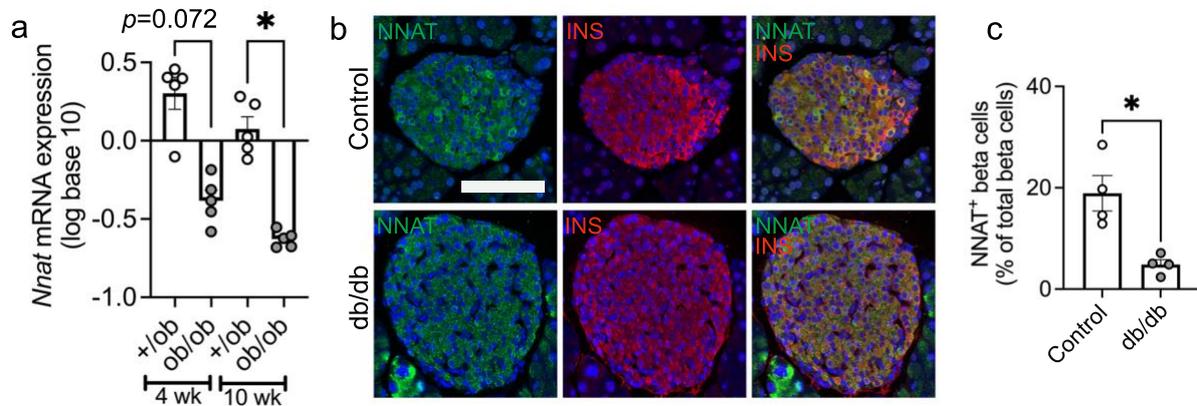
ESM Fig. 1. (a) Meta-analysis of *Nnat* mRNA expression in isolated islets of 4 week and 10 week old *ob/ob* vs *+/ob* control mice (n = 5 mice per genotype and age, Kruskal-Wallace test with Dunn's multiple comparisons). (b) Representative confocal microscopy of pancreatic sections from *db/db* vs control mice immunostained with antibodies against NNAT (green) and insulin (INS, red). Nuclei are visualised with DAPI. Scale bar = 100μm (n = 4 mice per genotype). (c) Quantification of NNAT$^+$ beta cells from images shown in b, expressed as NNAT/INS co-positive cells as a percentage of total INS-positive cells. *$p<0.05$.
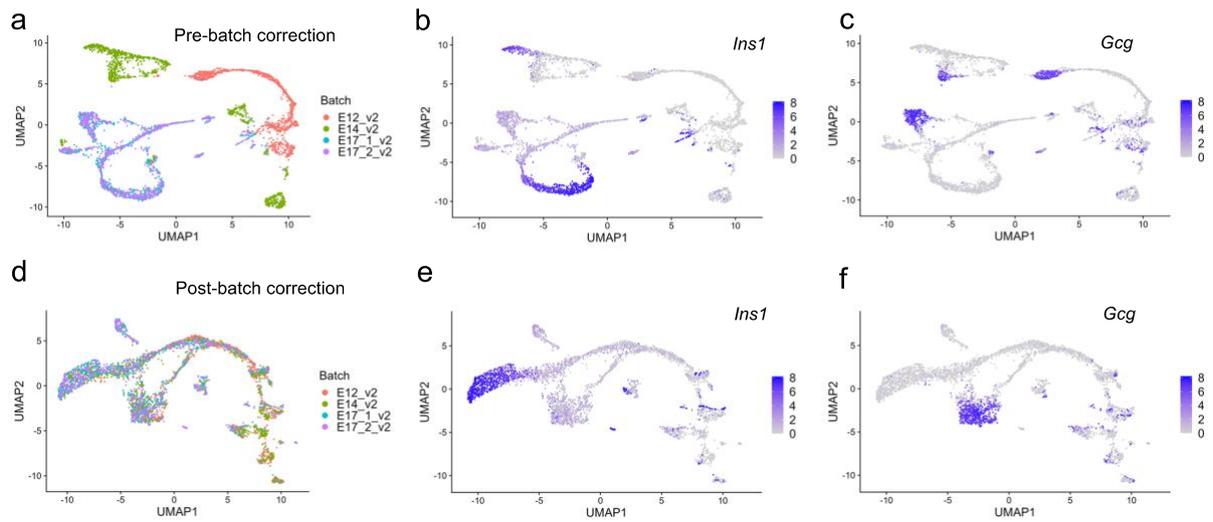
ESM Fig. 2. (a-c) UMAP projection of scRNA-seq data from embryonic mouse islets (E12.5-17.5) before batch correction labelled by embryonic stage for *Ins1* expression (b), and *Gcg* expression (c). (d-f) UMAP projection of the same islet cells following batch correction.
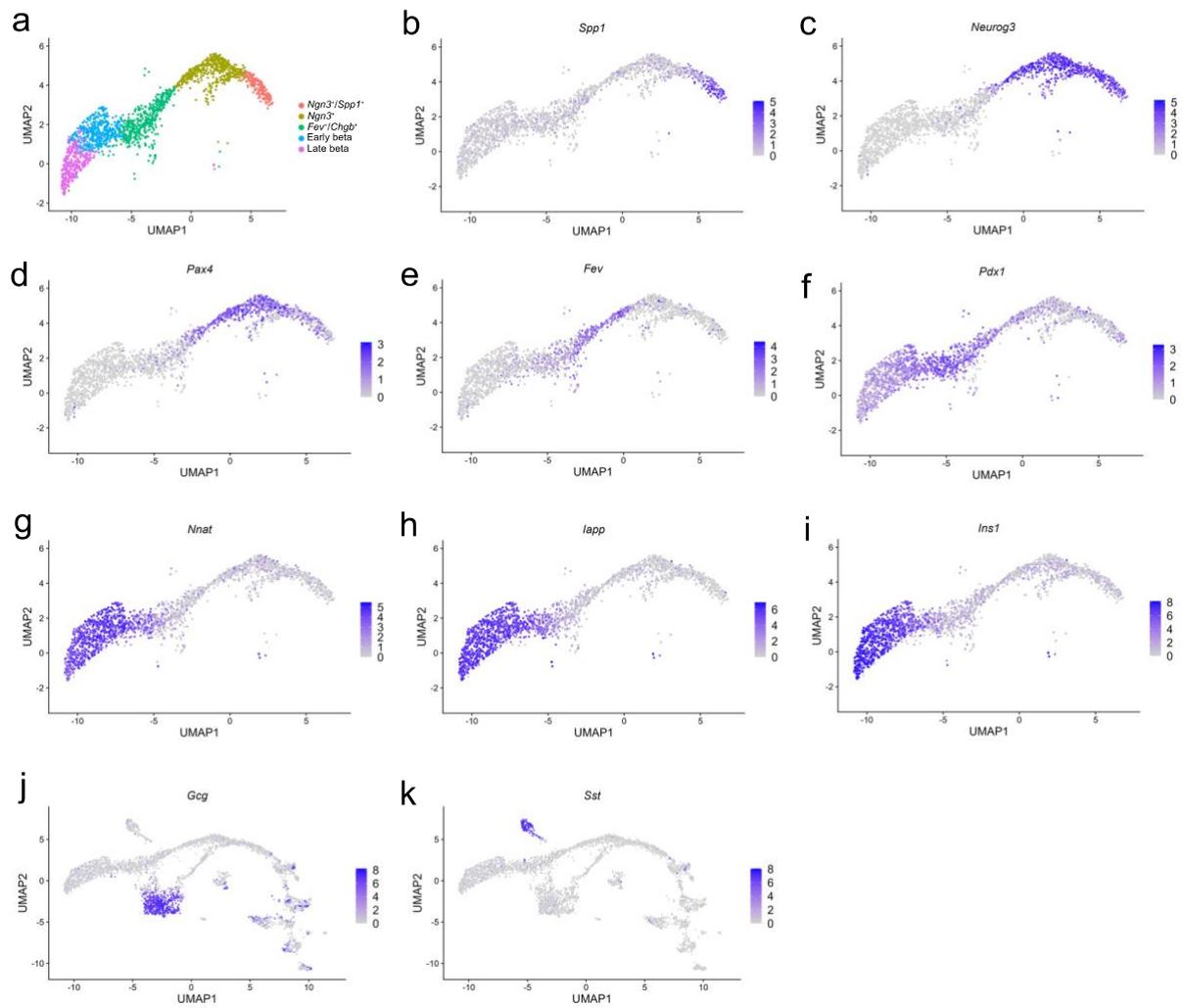
ESM Fig. 3. UMAP projections of beta cell precursors and identity markers, and other islet cell markers, from pseudo-time analysis of embryonic scRNA-seq islet data. (a) Plot representing the five beta cell states found in the embryonic data and (b-i) plots representing the shift in expression of markers across precursor states from E12.5 to E17.5 (b-i) and markers of other islet cell types (j, k).
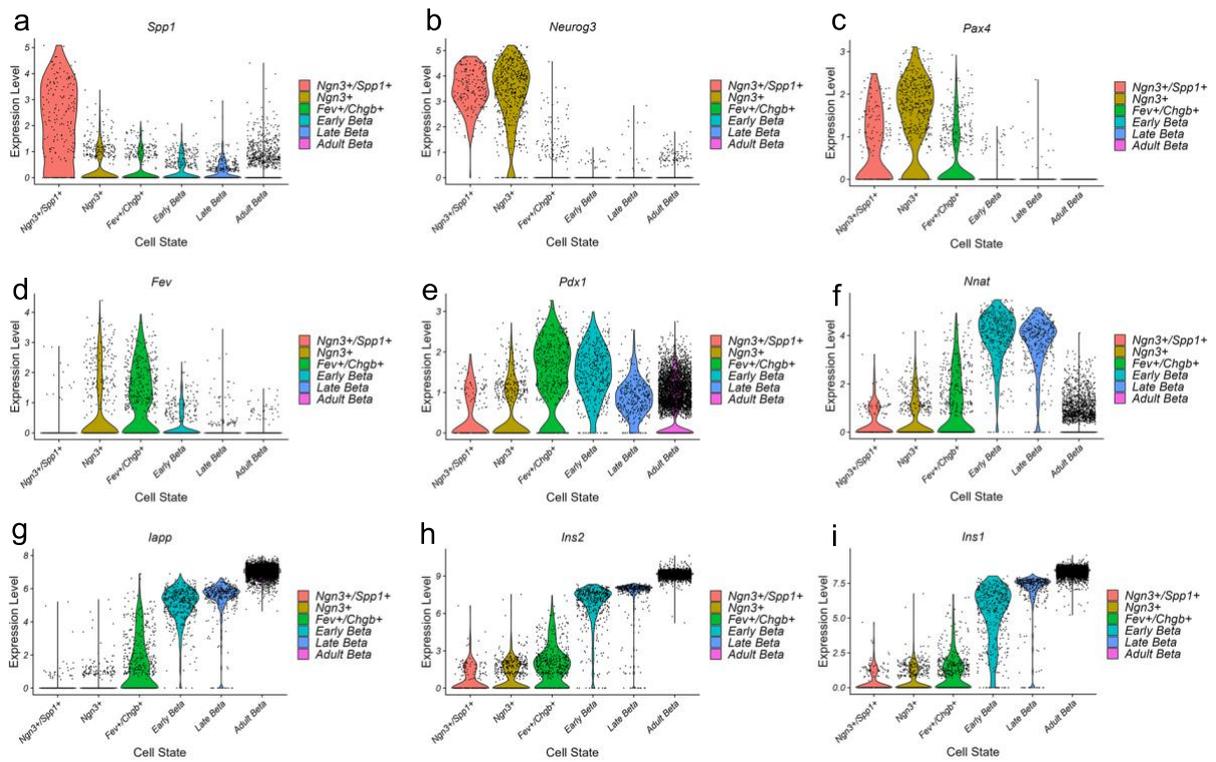
ESM Fig. 4. (a-i) Violin plots representing the progressive change in marker expression from early beta cell precursors during embryogenesis (E12.5 – E17.5) to mature beta cells in adulthood.
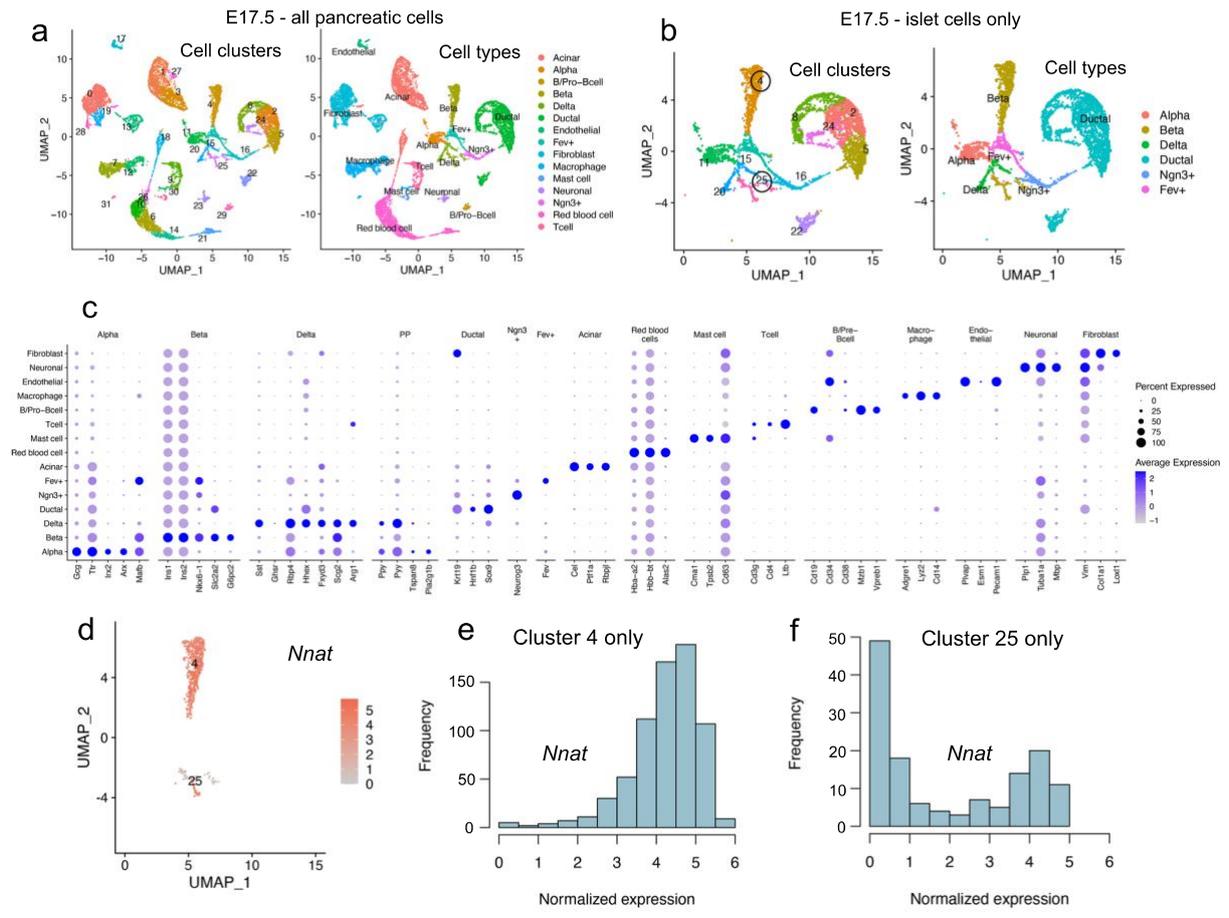
ESM Fig. 5. (a, b) Cell clusters and assigned cell types within the E17.5 pancreas for all cells (a) and islet cells only (b), displayed in the UMAP space. In panel b, the beta cell clusters are labelled using black circles. (c) Dot plot showing the expressions of known marker genes within each assigned cell type. Each dot is coloured according to the mean gene expressions, while the dot diameters are proportional to the percentage of cells expressing the gene. (d) UMAP showing *Nnat* expression in cell clusters 4 and 25 only (beta cell clusters). (e, f) Distribution of *Nnat* expression in individual beta cell clusters 4 (e) and 25 (f).
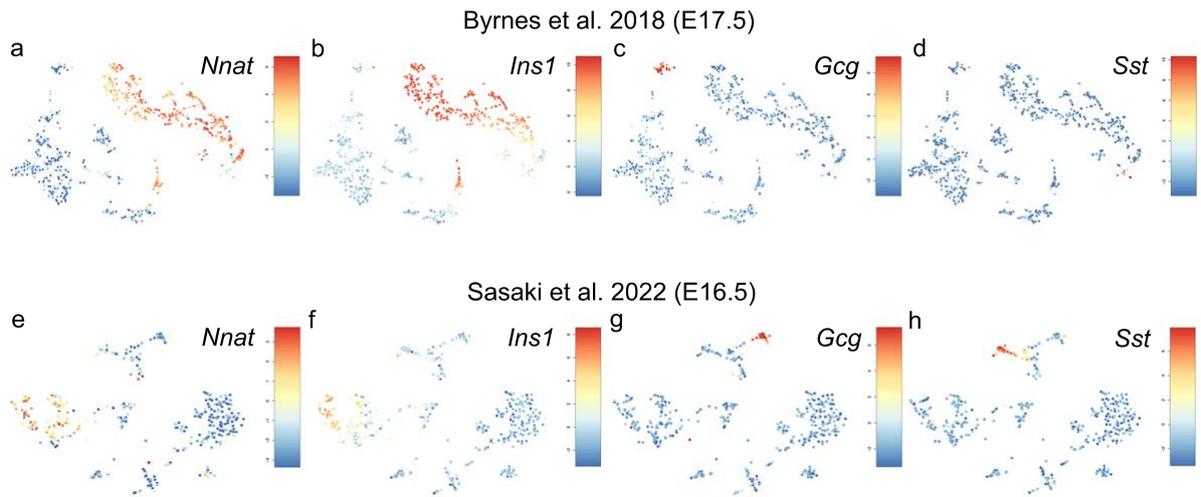
ESM Fig. 6. UMAP visualisation of gene expressions of *Nnat* and key islet cell markers in islet

scRNA-seq datasets at the embryonic stage (E17.5, a-d and E16.5, e-h).
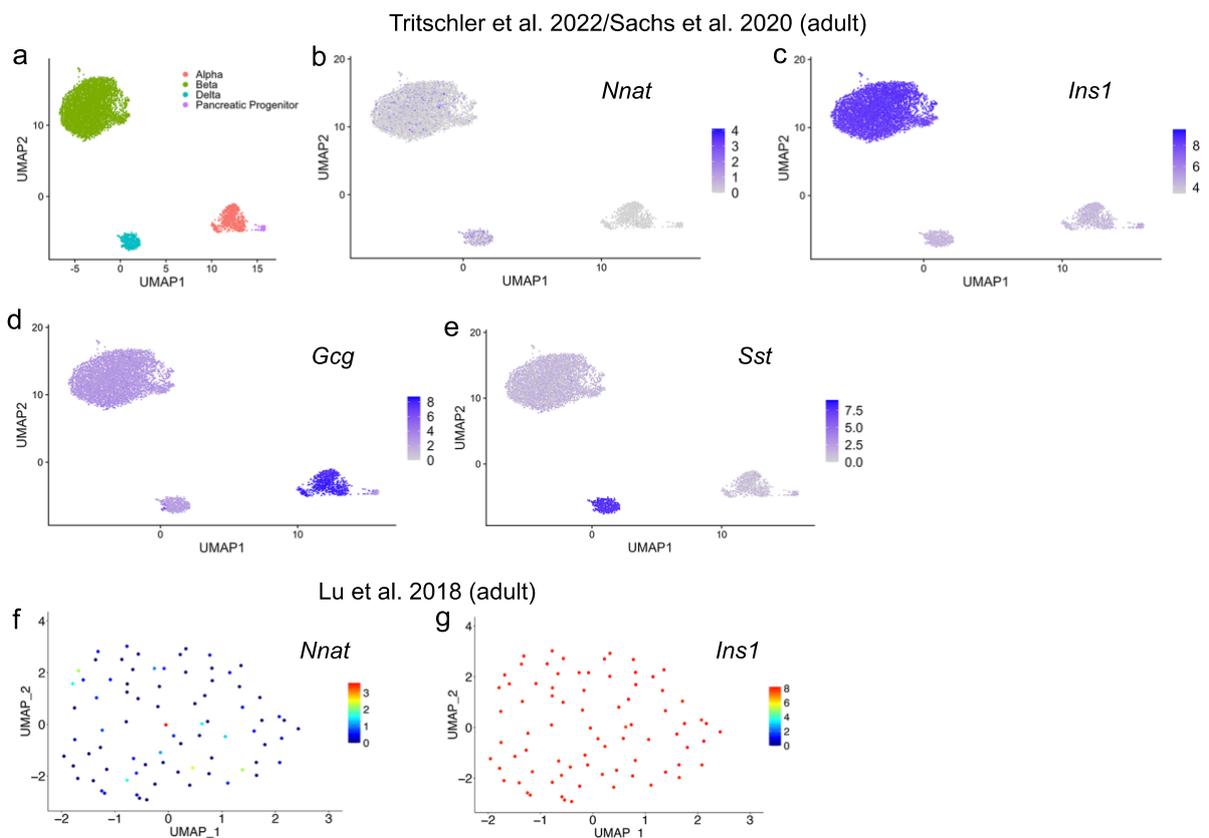
ESM Fig. 7. UMAP visualisation of gene expressions of *Nnat* and key islet cell markers in islet

scRNA-seq datasets at the adult stage (a-e, all islet cells) and (f, g, beta cells only).
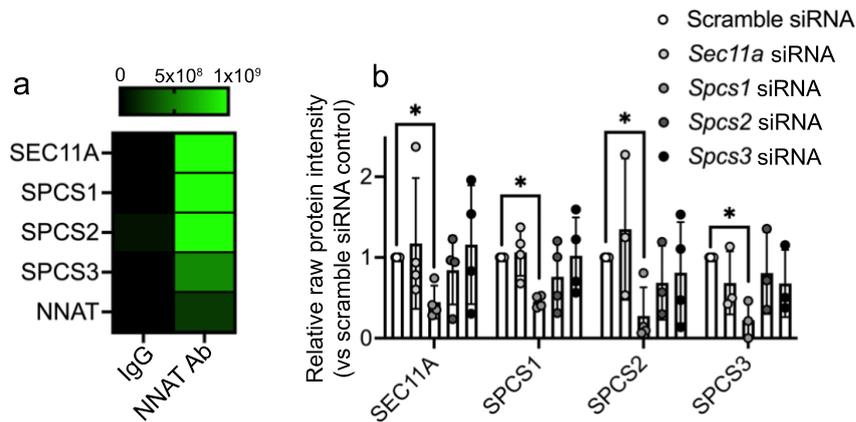
ESM Fig. 8. (a) Heatmap from affinity purification/mass spectrometry (AP/MS) analysis showing co-immunoprecipitation of signal peptidase complex subunits with endogenous NNAT in human EndoC-βH1 beta cells using antibodies against NNAT (NNAT Ab) vs control IPs with rabbit immunoglobulins (IgG). (b) INS1E cells transiently transfected with siRNA targeting subunits of the signal peptidase complex followed by NNAT immunoprecipitation as above and detection of signal peptidase complex components by affinity purification/mass spectrometry analysis, with relative protein abundance measured vs scramble siRNA controls (n = 4 independent cultures per group, Kruskal-Wallace test with Dunn's multiple comparisons).
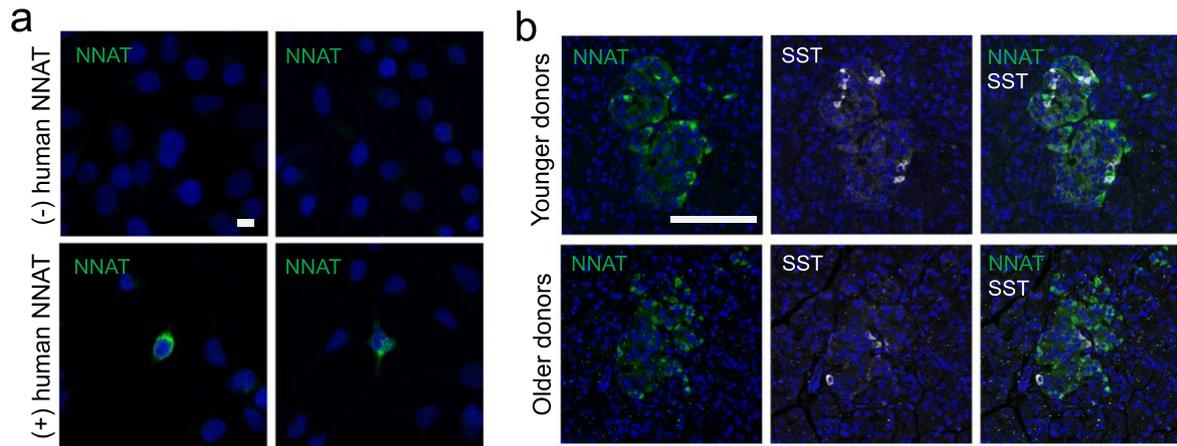
ESM Fig. 9. (a) Representative confocal microscopy of HEK293T cells transfected with a plasmid expressing human NNAT and immunostained with antibodies against NNAT (green). Cells transfected with empty vector were used as a control and nuclei are visualised with DAPI. Scale bar = 10μm (three independent experiments, each in triplicate per group). (b) Representative confocal microscopy of human pancreatic cryosections from younger (15.6±0.9 years, n = 5) and older (71.0±3.9 years, n = 4) donors. Sections were immunostained with antibodies against endogenous neuronatin (NNAT, green) and somatostatin (SST, grey). Nuclei are visualised with DAPI. Scale bar = 100 μm.
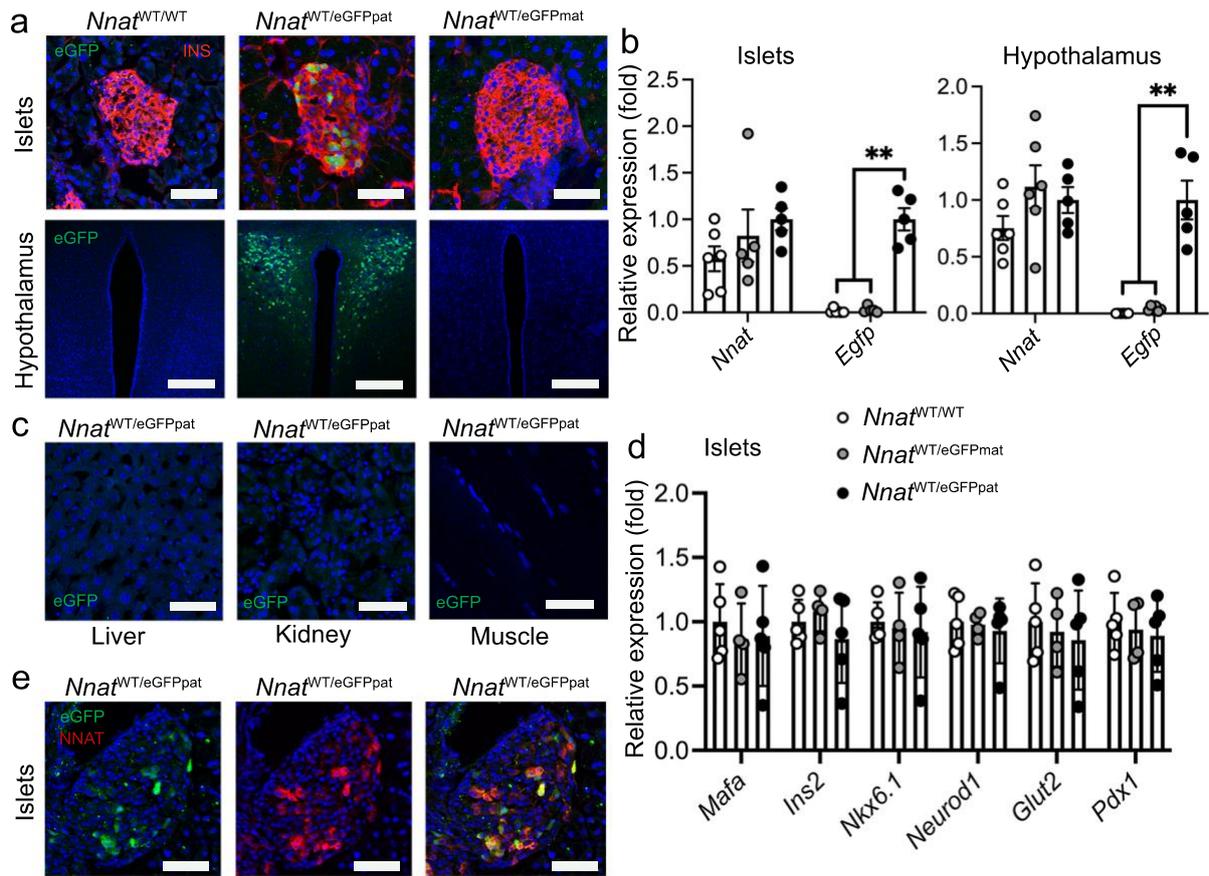
ESM Fig. 10. (a) Representative confocal microscopy of pancreatic and hypothalamic cryosections from P56 mice with *Nnat*-driven eGFP expression from the paternal (*Nnat*^WT/eGFPpat^), or maternal *(Nnat*^WT/eGFPmat^) allele or wild type (*Nnat*^WT/WT^) at this locus. Sections were immunostained with antibodies against eGFP (green) and insulin (INS, red). Nuclei are visualised with DAPI. Scale bar = 100μm (n = 3 mice per genotype). (b) RT-PCR analysis of *Nnat* and *Egfp* mRNA expression in isolated islets and hypothalamus of *Nnat*^WT/eGFPpat^, *Nnat*^WT/eGFPmat^ or *Nnat*^WT/WT^ mice. *Hprt* mRNA was used as an internal control and data are expressed relative to *Nnat*^WT/eGFPpat^ mice (n = 5-6 mice per genotype, one way ANOVA with Bonferroni test for multiple comparisons, **$p < 0.01$). (c) Representative confocal microscopy of cryosections using multiple tissues from P56 *Nnat*^WT/eGFPpat^ mice (n = 4 mice per genotype). Sections were immunostained with antibodies against eGFP (green) and nuclei are visualised with DAPI. Scale bar = 50μm. (d) RT-PCR analysis of mRNAs encoding key beta cell genes in isolated islets of *Nnat*^WT/eGFPpat^, *Nnat*^WT/eGFPmat^ or *Nnat*^WT/WT^ mice. *Hprt* mRNA was used as an internal control and data are expressed relative to *Nnat*^WT/WT^ mice (n = 4-5

mice per genotype). (e) Representative confocal microscopy of pancreatic cryosections from P56 $Nnat^{WT/eGFPpat}$ mice (n = 3 mice per genotype). Sections were immunostained with antibodies against eGFP (green) and endogenous NNAT (red). Nuclei are visualised with DAPI. Scale bar = 100μm. Representative images from three independent experiments and breeding pairs.
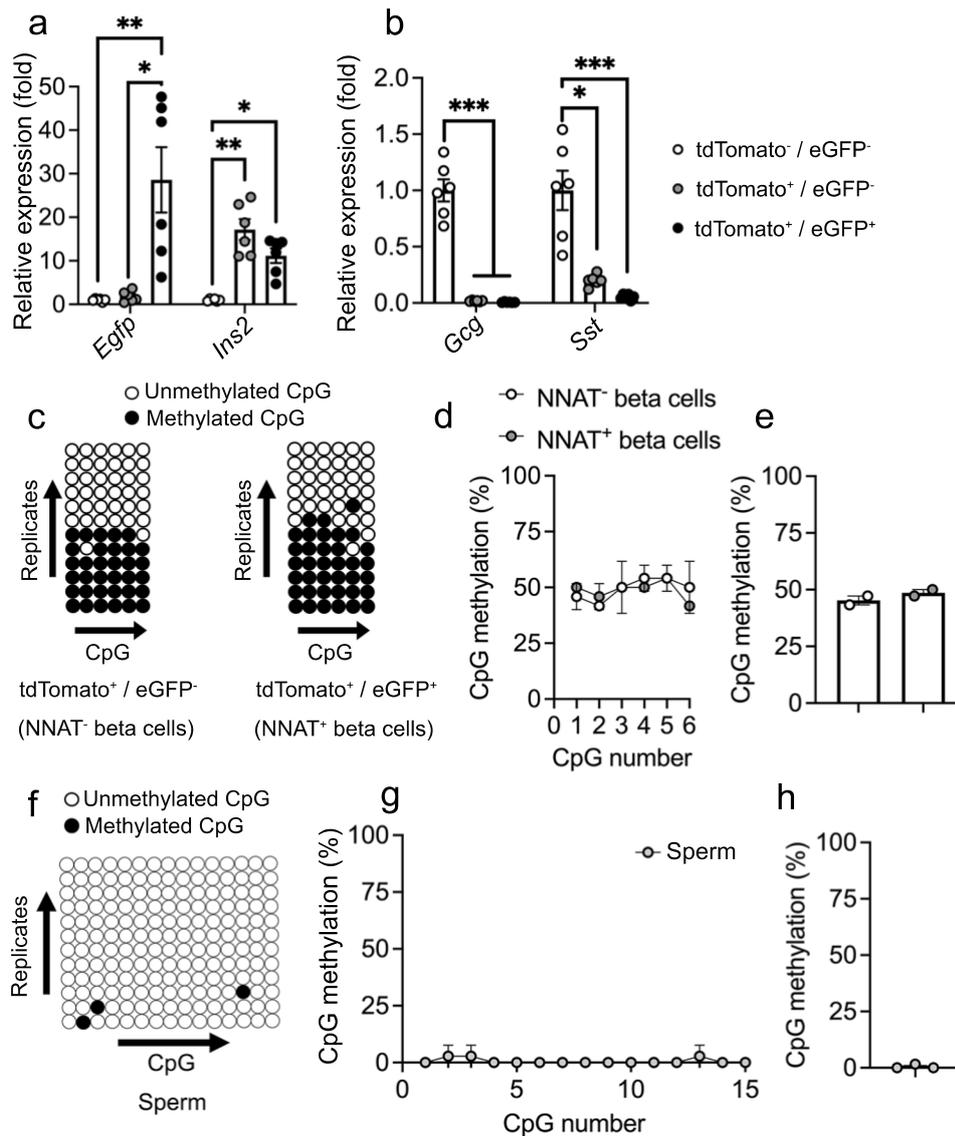
ESM Fig. 11. (a, b) RT-PCR analysis of *Egfp*, *Ins2*, *Gcg*, and *Sst* mRNA expression in FACS-sorted primary islet cells from reporter mice with insulin-driven expression of tdTomato (to label beta cells) and *Nnat*-driven eGFP expression from the paternal (*Nnat*[WT/eGFPpat]) (n = 6 mice, *$p<0.05$, **$p<0.01$, ***$p<0.001$, one way ANOVA with Bonferroni test for multiple comparisons). *Hprt* mRNA was used as an internal control and data are expressed relative to tdTomato⁻/eGFP⁻ cells. (c) Representative bisulphite analysis of CpG methylation at the *Nnat* gametic DMR in FACS-purified islet cell populations (n = 2 *Nnat*[WT/eGFPpat] mice with paternally expressed *Nnat*-driven eGFP, n = 12 clones each, two independent experiments and breeding pairs). (d, e) Quantification of data in c expressed as percentage CpG methylation across the

Nnat gametic DMR at individual CpGs (d) and across the entire *Nnat* gametic DMR (e). (f) Representative bisulphite analysis of CpG methylation at the *Nnat* promoter in purified sperm cells from male C57BL/6J mice (n = 3, n = 12 clones each, three independent experiments and breeding pairs). (g, h) Quantification of data in f expressed as percentage CpG methylation across the *Nnat* promoter at individual CpGs (g) and across the entire *Nnat* promoter (h). Closed circles = methylated CpG, open circles = unmethylated CpG.
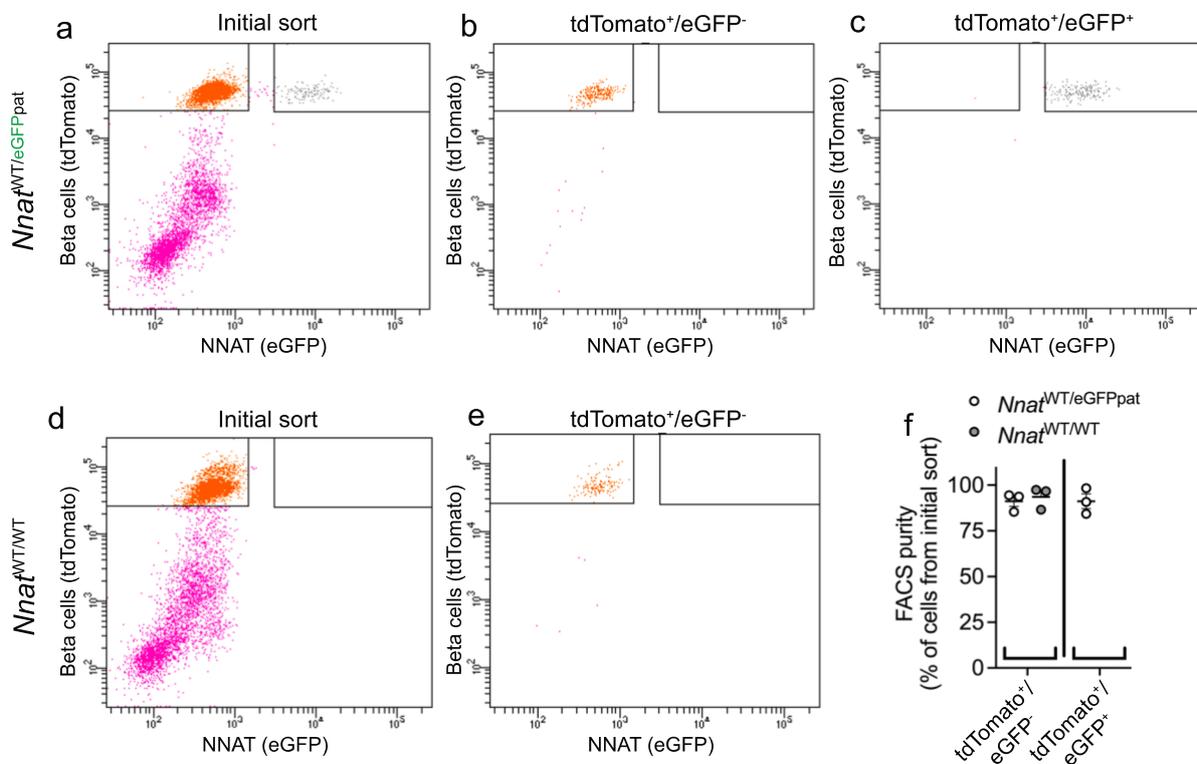
ESM Fig. 12. Separation of dispersed primary islet cells by FACS from reporter mice with insulin-driven expression of tdTomato (to label beta cells) and *Nnat*-driven eGFP expression from the paternal (*Nnat*$^{WT/eGFPpat}$) allele or wild type (*Nnat*$^{WT/WT}$) at this locus (representative image of one dispersed islet preparation). The initial sort (a, d) collecting separate populations of tdTomato$^+$/eGFP$^-$ (b, e) and tdTomato$^+$/eGFP$^+$ (c) cells were then re-sorted to assess purity (three independent experiments). (f) Quantification of data in a-e to expressed as purity of specific cell populations as a percentage of total cells.
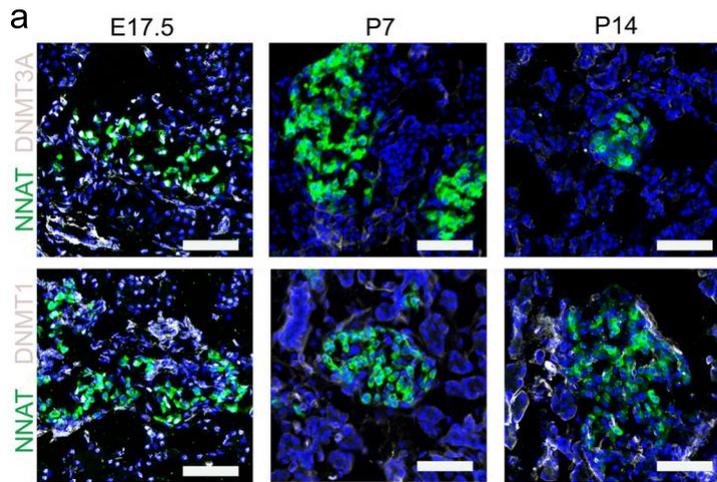
ESM Fig. 13. (a) Representative confocal microscopy of pancreatic cryosections from wild type mice on a C57BL/6J background at embryonic (E) day 17.5 and postnatal (P) day 7 and 14 (n = 3 mice per developmental stage). Sections were immunostained with antibodies against endogenous neuronatin (NNAT, green) and DNMT3A or DNMT1 (both grey). Nuclei are visualised with DAPI. Scale bar = 50μm.
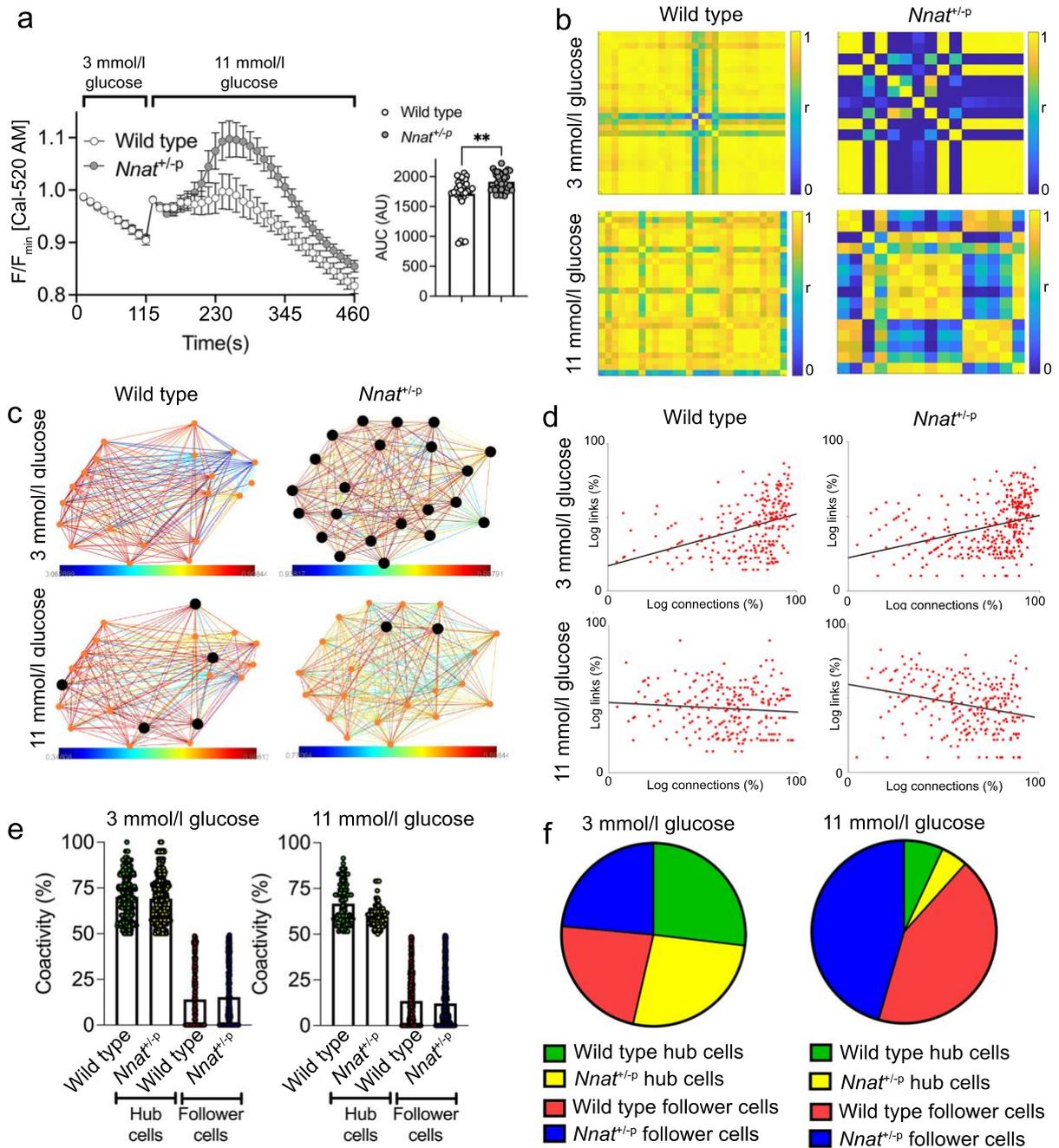
ESM Fig. 14. (a) Ca^{2+} bound Cal-520 AM fluorescence in response to high (11 mmol/l) vs low (3 mmol/l) glucose in primary islets from wild type and *Nnat*$^{+/-p}$ mice expressed as normalised intensity over time (F / F$_{min}$) (n = 31 wild type and 38 *Nnat* deficient islets total from 9 mice per genotype, inset shows quantification of area under the curve (AUC), **$p<0.01$, Mann-Whitney U test). (b) Representative heatmaps depicting connectivity strength (r) of all cell pairs (colour coded r values from 0 to 1, blue to yellow). (c) Representative cartesian map of cells with

colour coded lines connecting cells according of the strength of coactivation (colour coded R values from 0 to 1, blue to red). Beta cells are represented by differently coloured nodes depending on their coactivity with the other beta cells, where black nodes indicate coactivity with ≥80% of the remaining beta cells, while grey and white nodes represent coactivity with ≥60% and ≥40%, respectively. (d) Log-log graphs of beta cell-beta cell connectivity distribution. (e) Percentage coactivity of beta cells between all cells for identified 'hub' and 'follower' cells. (f) The proportion of cells designated as 'hub' cells vs 'follower' cells. Representative images from three independent experiments and breeding pairs.

ESM References

[1]   Byrnes LE, Wong DM, Subramaniam M, et al. (2018) Lineage dynamics of murine pancreatic development at single-cell resolution. Nat Commun 9(1): 3922. 10.1038/s41467-018-06176-3

[2]   Sasaki S, Lee MYY, Wakabayashi Y, et al. (2022) Spatial and transcriptional heterogeneity of pancreatic beta cell neogenesis revealed by a time-resolved reporter system. Diabetologia 65(5): 811-828. 10.1007/s00125-022-05662-0

[3]   Sachs S, Bastidas-Ponce A, Tritschler S, et al. (2020) Targeted pharmacological therapy restores beta-cell function for diabetes remission. Nat Metab 2(2): 192-209. 10.1038/s42255-020-0171-3

[4]   Tritschler S, Thomas M, Bottcher A, et al. (2022) A transcriptional cross species map of pancreatic islet cells. Mol Metab 66: 101595. 10.1016/j.molmet.2022.101595

[5]   Lu TT, Heyne S, Dror E, et al. (2018) The Polycomb-Dependent Epigenome Controls beta Cell Dysfunction, Dedifferentiation, and Diabetes. Cell Metab 27(6): 1294-1308 e1297. 10.1016/j.cmet.2018.04.013

[6]   Lun AT, McCarthy DJ, Marioni JC (2016) A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. F1000Res 5: 2122. 10.12688/f1000research.9501.2

[7]   Herman JS, Sagar, Grun D (2018) FateID infers cell fate bias in multipotent progenitors from single-cell RNA-seq data. Nat Methods 15(5): 379-386. 10.1038/nmeth.4662

[8]   Hafemeister C, Satija R (2019) Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol 20(1): 296. 10.1186/s13059-019-1874-1

[9]   Hao Y, Hao S, Andersen-Nissen E, et al. (2021) Integrated analysis of multimodal single-cell data. Cell 184(13): 3573-3587 e3529. 10.1016/j.cell.2021.04.048

[10]   Ashburner M, Ball CA, Blake JA, et al. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25(1): 25-29. 10.1038/75556

[11]   Gene Ontology C (2021) The Gene Ontology resource: enriching a GOld mine. Nucleic Acids Res 49(D1): D325-D334. 10.1093/nar/gkaa1113

[12]   Subramanian A, Tamayo P, Mootha VK, et al. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 102(43): 15545-15550. 10.1073/pnas.0506580102

[13]   Mootha VK, Lindgren CM, Eriksson K-F et al. (2003) PGC-1-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet 34(3):267-273. https://doi.org/10.1038/ng1180

[14]   Yu G, Wang LG, Han Y, He QY (2012) clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 16(5): 284-287. 10.1089/omi.2011.0118

[15]   Yu G, Wang LG, Yan GR, He QY (2015) DOSE: an R/Bioconductor package for disease ontology semantic and enrichment analysis. Bioinformatics 31(4): 608-609. 10.1093/bioinformatics/btu684

[16]   Korsunsky I, Millard N, Fan J, et al. (2019) Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods 16(12): 1289-1296. 10.1038/s41592-019-0619-0

[17]   Butler A, Hoffman P, Smibert P, Papalexi E, Satija R (2018) Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol 36(5): 411-420. 10.1038/nbt.4096

[18]   Cao J, Spielmann M, Qiu X, et al. (2019) The single-cell transcriptional landscape of mammalian organogenesis. Nature 566(7745): 496-502. 10.1038/s41586-019-0969-x

[19]   Herrera PL (2000) Adult insulin- and glucagon-producing cells differentiate from two independent cell lineages. Development 127(11): 2317-2322. 10.1242/dev.127.11.2317

[20]   Luche H, Weber O, Nageswara Rao T, Blum C, Fehling HJ (2007) Faithful activation of an extra-bright red fluorescent protein in "knock-in" Cre-reporter mice ideally suited for lineage tracing studies. Eur J Immunol 37(1): 43-53. 10.1002/eji.200636745

[21]   Millership SJ, Da Silva Xavier G, Choudhury AI, et al. (2018) Neuronatin regulates pancreatic beta cell insulin content and secretion. J Clin Invest 128(8): 3369-3381. 10.1172/JCI120115

[22]   Kaneda M, Okano M, Hata K, et al. (2004) Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. Nature 429(6994): 900-903. 10.1038/nature02633

[23]   Hingorani SR, Petricoin EF, Maitra A, et al. (2003) Preinvasive and invasive ductal pancreatic cancer and its early detection in the mouse. Cancer Cell 4(6): 437-450. 10.1016/s1535-6108(03)00309-x

[24]   Dhawan S, Tschen SI, Zeng C, et al. (2015) DNA methylation directs functional maturation of pancreatic beta cells. J Clin Invest 125(7): 2851-2860. 10.1172/JCI79956

[25]   Georgiadou E, Muralidharan C, Martinez M, et al. (2022) Mitofusins Mfn1 and Mfn2 Are Required to Preserve Glucose- but Not Incretin-Stimulated beta-Cell Connectivity and Insulin Secretion. Diabetes 71(7): 1472-1489. 10.2337/db21-0800

[26]   Parveen N, Wang JK, Bhattacharya S, et al. (2023) DNA methylation Dependent Restriction of Tyrosine Hydroxylase Contributes to Pancreatic beta-cell Heterogeneity. Diabetes. 10.2337/db22-0506

[27]   Parton LE, McMillen PJ, Shen Y, et al. (2006) Limited role for SREBP-1c in defective glucose-induced insulin secretion from Zucker diabetic fatty rat islets: a functional and gene profiling analysis. Am J Physiol Endocrinol Metab 291(5): E982-994. 10.1152/ajpendo.00067.2006

[28]   Akalestou E, Suba K, Lopez-Noriega L, et al. (2021) Intravital imaging of islet Ca(2+) dynamics reveals enhanced beta cell connectivity after bariatric surgery in mice. Nat Commun 12(1): 5165. 10.1038/s41467-021-25423-8

[29]   Chabosseau P, Yong F, Delgadillo-Silva LF, et al. (2023) Molecular phenotyping of single pancreatic islet leader beta cells by "Flash-Seq". Life Sci: 121436. 10.1016/j.lfs.2023.121436

[30]   Wang H, Ranalli MG (2007) Low-rank smoothing splines on complicated domains. Biometrics 63(1): 209-217. 10.1111/j.1541-0420.2006.00674

[31]   Hodson DJ, Molino F, Fontanaud P, Bonnefont X, Mollard P (2010) Investigating and modelling pituitary endocrine network function. J Neuroendocrinol 22(12): 1217-1225. 10.1111/j.1365-2826.2010.02052

[32]   van der Meulen T, Xie R, Kelly OG, Vale WW, Sander M, Huising MO (2012) Urocortin 3 marks mature human primary and embryonic stem cell-derived pancreatic alpha and beta cells. PLoS One 7(12): e52181. 10.1371/journal.pone.0052181

[33]   Tschen SI, Dhawan S, Gurlo T, Bhushan A (2009) Age-dependent decline in beta-cell proliferation restricts the capacity of beta-cell regeneration in mice. Diabetes 58(6): 1312-1320. 10.2337/db08-1651

[34]   Rodnoi P, Rajkumar M, Moin ASM, Georgia SK, Butler AE, Dhawan S (2017) Neuropeptide Y expression marks partially differentiated beta cells in mice and humans. JCI Insight 2(12). 10.1172/jci.insight.94005

[35]   Gloyn AL, Ibberson M, Marchetti P, et al. (2022) Every islet matters: improving the impact of human islet research. Nat Metab 4(8): 970-977. 10.1038/s42255-022-00607-8

[36]   Marchetti P, Dotta F, Lauro D, Purrello F (2008) An overview of pancreatic beta-cell defects in human type 2 diabetes: implications for treatment. Regul Pept 146(1-3): 4-11. 10.1016/j.regpep.2007.08.017

[37]   Matsuda T, Cepko CL (2004) Electroporation and RNA interference in the rodent retina in vivo and in vitro. Proc Natl Acad Sci U S A 101(1): 16-22. 10.1073/pnas.2235688100

[38]   Schindelin J, Arganda-Carreras I, Frise E, et al. (2012) Fiji: an open-source platform for biological-image analysis. Nat Methods 9(7): 676-682. 10.1038/nmeth.2019

[39]   Ravassard P, Hazhouz Y, Pechberty S, et al. (2011) A genetically engineered human pancreatic beta cell line exhibiting glucose-inducible insulin secretion. J Clin Invest 121(9): 3589-3597. 10.1172/JCI58447

[40]   Maechler P, Wollheim CB (1999) Mitochondrial glutamate acts as a messenger in glucose-induced insulin exocytosis. Nature 402(6762): 685-689. 10.1038/45280

[41]   Tyanova S, Temu T, Cox J (2016) The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nat Protoc 11(12): 2301-2319. 10.1038/nprot.2016.136

[42]   Perez-Riverol Y, Bai J, Bandla C et al. (2022). The PRIDE database resources in 2022: A Hub for mass spectrometry-based proteomics evidences. Nucleic Acids Res 50(D1):D543-D552. 10.1093/nar/gkab1038

[43]   Dobin A, Davis CA, Schlesinger F, et al. (2013) STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29(1): 15-21. 10.1093/bioinformatics/bts635

[44]   Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15(12): 550. 10.1186/s13059-014-0550-8

[45]   Tusnady GE, Simon I, Varadi A, Aranyi T (2005) BiSearch: primer-design and search tool for PCR on bisulfite-treated genomes. Nucleic Acids Res 33(1): e9. 10.1093/nar/gni012

[46]   Rohde C, Zhang Y, Reinhardt R, Jeltsch A (2010) BISMA--fast and accurate bisulfite sequencing data analysis of individual clones from unique and repetitive sequences. BMC Bioinformatics 11: 230. 10.1186/1471-2105-11-230

[47]   Johnston NR, Mitchell RK, Haythorne E, et al. (2016) Beta Cell Hubs Dictate Pancreatic Islet Responses to Glucose. Cell Metab 24(3): 389-401. 10.1016/j.cmet.2016.06.020

[48]   Salem V, Silva LD, Suba K, et al. (2019) Leader beta-cells coordinate Ca(2+) dynamics across pancreatic islets in vivo. Nat Metab 1(6): 615-629. 10.1038/s42255-019-0075-2

Diabetologia

Checklist for reporting human islet preparations used in research

Adapted from Hart NJ, Powers AC (2018) Progress, challenges, and suggestions for using human islets to understand islet biology and human diabetes. Diabetologia https://doi.org/10.1007/s00125-018-4772-2

Islet preparation	1	2	3	4	5	6	7	8	9
MANDATORY INFORMATION									
Unique identifier	18/40	26/88	27/119	20/46	20/60	18/29	17/71	17/17	16/58
Donor age (years)	17	15	18	15	13	60	71	78	75
Donor sex (M/F)	M	F	M	M	M	F	M	M	F
Donor BMI (kg/m^2)	24.5	24.7	20.4	19.5	25.4	20.8	25.4	24.2	23.4
Donor HbA$_{1c}$ or other measure of blood glucose control**	189	95	NA	162	164	NA	63	215	140
Origin/source of islets*									
Islet isolation centre*									
Donor history of diabetes? Please select yes/no from drop down list	No	No	No	No	No	No	No	No	No
If Yes, complete the next two lines if this information is available									
Diabetes duration (years)									
Glucose-lowering therapy at time of death									
RECOMMENDED INFORMATION									
Donor cause of death	Trauma	CVD	Trauma	Drowning	Trauma	CVD	Trauma	Trauma	CVD
Warm ischaemia time (h)									

Cold ischaemia time (h)	12	17	NA	24	21	20	14	10	20
Estimated purity (%)*									
Estimated viability (%)*									
Total culture time (h)*									
Glucose-stimulated insulin secretion or other functional measurement**									
Handpicked to purity? Please select yes/no from drop down list									
Additional notes *only pancreatic tissue studied **ICU plasma glucose levels (mg/dl)									

ICU: intensive care unit ; CVD: cardiovascular disease ; NA: not available.