Supplementary Materials:

Cell culture

Generation of stem cell derived beta-like cells from human embryonic stem cells and from induced pluripotent stem cells

MEL1 human embryonic stem cells (hESC) containing the $INS^{GFP/W}$ reporter [18] (referred to as pINSGFP throughout the manuscript) and sub-clones thereof [19,20], type 1 diabetes patient induced pluripotent stem cells (T1D-iPSC) [21] (Supp. Fig 8.) and iPSC from a healthy donor [22] (Supp. Fig. 9), as well as pNKX6.1 GFP reporter iPSC[23] were employed in this study. All iPSC lines were maintained on hESC qualified Matrigel (Corning #354277) in mTeSR+ media (STEMCELL Technologies #05826). MEL1 Subclones were only used for bulk RNA sequencing analysis experiments. Differentiation to stem cell-derived beta-like cells (sBCs) was carried out in suspension-based, low attachment suspension culture plates as described [19] or in a bioreactor magnetic stirring system (Reprocell #ABBWVS03A-6, #ABBWVDW-1013, #ABBWBP03N0S-6) as follows: Confluent hESC cultures were dissociated into single-cell suspensions by incubation with TrypLE (Gibco #12-604-021) for 6 min at 37C. Detached cells were quenched with mTESR media. Live cells were counted using a MoxiGo II cell counter (Orflow), followed by seeding 0.5×10^6 cells per ml in mTeSR media supplemented with 10 μM ROCK inhibitor (Y-27632, R&D Systems #1254-50) (cluster media). Bioreactors were placed on a magnetic stirring system set at 60 RPM in a cell culture incubator at 5 % CO_2 to induce sphere formation for 48 h. To induce definitive endoderm differentiation, spheres were collected in a 50 mL Falcon tube, allowed to settle by gravity, washed once with RPMI

(Gibco #11-875-093) + 0.2 % FBS, and re-suspended in d 0 media (RPMI containing 0.2 % FBS, 1:5,000 ITS (Gibco #41400-045), 100 ng/mL Activin-A (R&D Systems #338-AC-01M), and 3 μM CHIR (STEMCELL Technologies #72054)). Differentiation media was changed daily by letting spheres settle by gravity for 3-10 min. ~80 % of spent supernatant was removed by aspiration; fresh media was added, and bioreactors were placed back on stirrer system. sBC differentiation was based on published protocol [4] with modifications as outlined below. Differentiation medias are as: d 1 and 2, RPMI containing 0.2 % FBS, 1:2,000 ITS, and 100 ng/LmL Activin A; d 3 and 4, RPMI containing 2% FBS, 1:1,000 ITS, and 25 ng/LmL KGF (Peprotech #100-19-1MG);); d 5, DMEM with 4.5 g/L D-glucose (Gibco #11960-044) containing 1:100 SM1 (STEMCELL Technologies #5711), 1:100 NEAA (Gibco #11140-050), 1 mM Sodium Pyruvate (Gibco #11360-070), 1:100 GlutaMAX (Gibco #35050-061), 3 nM TTNPB, (R&D Systems #0761), 250 nM Sant-1 (R&D Systems #1974), 250 nM LDN (STEMCELL Technologies #72149), 30 nM PMA (Sigma Aldrich #P1585-1MG), 50 μg/mL 2-phospho-L-ascorbic acid trisodium salt (VitC) (Sigma #49752-10G); d6, DMEM with 4.5 g/L D-glucose containing 1:100 SM1, 1:100 NEAA, 1 mM Sodium Pyruvate, 1:100 GlutaMAX, 3 nM TTNPB and 50 μg/mL VitC; d 7, addition of 100 ng/mL EGF (R&D Systems #236-EG-01M) and 50 μg/mL VitC to existing media; d 8 and 9, DMEM containing 1:100 SM1, 1:100 NEAA, 1 mM Sodium Pyruvate, 1:100 GlutaMAX, 100 ng/mL EGF, 25 ng/mL KGF, and 50 μg/mL VitC; d 10- 16 DMEM containing 2 % fraction V BSA, 1:100 NEAA, 1 mM Sodium Pyruvate, 1:100 GlutaMAX, 1:100 ITS, 10 μg/ml Heparin (Sigma #H3149-250KU), 2 mM N-Acetyl-L-cysteine (Cysteine) (Sigma #A9165-25G), 10 μM Zinc sulfate heptahydrate (Zinc) (Sigma #Z0251-100g), 1x BME, 10 μM Alk5i II RepSox (R&D Systems #3742/50), 1 μM 3,3',5-Triiodo-L-

thyronine sodium salt (T3) (Sigma #T6397), 0.5 μM LDN, 1 μM Gamma Secretase Inhibitor XX (XXi) (AsisChem #ASIS-0149) and 1:250 1 M NaOH to adjust pH to ~7.4; d 17 and up, CMRL (Gibco #11530-037) containing 1% BSA, 1:100 NEAA, 1 mM Sodium Pyruvate, 1:100 GlutaMAX, 10 μg/mL Heparin, 2 mM Cysteine, 10 μM Zinc, 1x BME, 10 μM Alk5i II RepSox, 1 μM T3, 50 μg/mL VitC, and 1:250 NaOH to adjust pH to ~7.4. All media contained 1x PenStrep (Gibco #15140-122). Media was changed every other day starting d11. For differentiations of iPSCs, 70 - 80 % confluent cultures were washed with PBS and incubated in TrypLE for 8 min at 37 °C followed by quenching with mTeSR+. 0.5×10^6 cells/mL in mTeSR media supplemented with 10 μM ROCK inhibitor were seeded and differentiated as per hESC bioreactor differentiation protocol above, with the following modifications: d 4 and 5, 50 ng/mL KGF instead of 25 ng/mL; d 7, DMEM containing 1:100 SM1, 1:100 NEAA, 1 mM Sodium Pyruvate, 1:100 GlutaMAX, 3 nM TTNPB and 50 μg/mL VitC; d8 and d9, DMEM containing 1:100 SM1, 1:100 NEAA, 1 mM Sodium Pyruvate, 1:100 GlutaMAX, 200ng/ml EGF and 50 ng/mL KGF; d 10-16, DMEM containing 2 % fraction V BSA, 1:100 NEAA, 1 mM Sodium Pyruvate, 1:100 GlutaMAX, 1:100 ITS, 10 μg/ml Heparin, 2 mM Cysteine, 10 μM Zinc, 1x BME, 10 μM Alk5i II RepSox, 1 μM T3, 0.5 μM LDN, 10μM RI, 1 μM Xxi and 1:250 1 M NaOH to adjust pH to ~7.4; d 17 and up, CMRL (Gibco #11530-037) containing 1% BSA, 1:100 NEAA, 1 mM Sodium Pyruvate, 1:100 GlutaMAX, 10 μg/mL Heparin, 2 mM Cysteine, 10 μM Zinc, 1x BME, 10 μM Alk5i II RepSox, 1 μM T3, 50 μg/mL VitC, and 1:250 NaOH to adjust pH to ~7.4 (also referred to as maturation media). All media contained 1x PenStrep. Media was changed every other day starting d11.

Human islet culture

Human islets (hIslets) were used in this study from:

1. Human islets for research were provided by the Alberta Diabetes Institute Islet Core at the University of Alberta in Edmonton (www.bcell.org/isletcore) with the assistance of the Human Organ Procurement and Exchange (HOPE) program, Trillium Gift of Life Network (TGLN) and other Canadian organ procurement organizations. Islet isolation was approved by the Human Research Ethics Board at the University of Alberta (Pro00013094) [39,40].
2. Human pancreatic islets were provided by the NIDDK-funded Integrated Islet Distribution Program (IIDP) (RRID:SCR _014387) at City of Hope, NIH Grant # 2UC4DK098085.

All donors' families gave informed consent for the use of pancreatic tissue in research (details of individual preps outlined in Methods Table. 1). hIslet were cultured for up to 24h in hIslet media (CMRL containing 1X Pen/Strep, 10 % FBS, 100 μg/mL Gentamicin (Sigma #G1914), 1X BME) before analysis.

Fluorescence associated cell sorting (FACS)

pINSGFP/ ENTPD3 sorting

pINSGFP clusters were collected in an Eppendorf tube, allowed to settle by gravity, the supernatant removed and then washed twice with PBS containing 2 mM EDTA (KD Medical #RGF-3130). Clusters were dissociated in 0.05 % trypsin/EDTA (Lonza #cc3232) in 37 °C bead bath (Thermo Scientific) for 15 min. After 15 min cluster/trypsin solution was vortexed for 1 min, fresh trypsin added and then incubated for a further 5 min at 37 °C. Finally, the suspension was pipetted up and down using a p1000 pipette until all clusters were fully dissociated. Cells were quenched immediately with ice cold culture

media and spun down. Supernatant was removed and cells resuspend in FACS buffer (PBS containing 2 % FBS and 2 mM EDTA). Cells were filtered through a 40 μm cell strainer into FACS tubes (Falcon #352235) for staining. For pINSGFP sorting, cells were incubated for 20 min on ice with DAPI (1:1000) then analyzed on BioRad S3e Cell Sorter; gating for live cells using DAPI and then pINSGFP on 488/FITC channel as per [18]. For ENTPD3 sorting of hESC derived seBCs, cells were incubated for 20 min on ice with DAPI (1:1000) and in house conjugated ENTPD3- Alexa555 antibody. For ENTPD3 sorting of iPSC derived seBCs, cells were first incubated with biotin labelled HPi1(HICO-49F) antibody [17] for 20 min on ice, then washed with FACS buffer. Cells were then incubated for 20 min on ice with Streptavidin-PECy7, and in house conjugated ENTPD3-Alexa555 antibody and DAPI (1:1000). (NTPDase3) antibody has been described and can be obtained at (http://ectonucleotidases-ab.com) [33]. Direct conjugation of ENTPD3 AB was done as per manufacture protocol (Thermo Fisher #A20187) in house. Cells were gated for live cells, then pINSGFP expression or HPi1 and then ENTPD3 as outlined in Fig. 4 and Supp Fig 4. Antibodies were used at concentrations indicated in Methods Table. 2.

Human islet sorting

Human islets were collected in an Eppendorf tube, allowed to settle by gravity, the supernatant removed and then washed twice with PBS. The islets were dissociated in 500 μL of warm 0.05 % trypsin for 15 min in a 37 °C bead bath – islets were pipetted up and down every 3 min using a p1000 pipette to aid dissociation. Single cells were quenched with culture media and resuspended in FACS buffer, filtered through a 40 μm cell strainer into a FACS tube. Cells were first incubated with biotin labelled HPi1(HICO-

49F) antibody [17] for 20 min on ice, then washed with FACS buffer. Cells were then incubated for 20 min on ice with Streptavidin-PECy7, HIC3-2D1D-PE [17], NTPDase3-4888 antibodies and DAPI (1:1000). After incubation, cells were washed with FACS Buffer and resuspended in FACS Buffer. Populations were gated and sorted on BioRad S3e Cell Sorter as per Supp Fig. 5. Antibodies were used at concentrations indicated in Methods Table. 2.

Cell characterization

Flow cytometry

hESC and iPSC clusters were collected and dissociated as outlined above. Single cells were filtered through cell strainer into FACS tubes and incubated for 30 min on ice (or overnight at 4 °C) in conjugated antibody diluted in FACS buffer. After incubation the cells were washed and strained again through cell strainer and resuspended in FACS buffer for analyses on CYTEK Aurora.

Immunofluorescence

sBC and human islet clusters were fixed for 20 min at room temperature with 4 % paraformaldehyde then washed twice with PBS. Fixed clusters were then prepped for (i) whole mount staining or (ii) embedding and cryo-sectioning. (i) whole mount staining was performed in suspension by blocking for 30 min in CAS-block (Thermo Fisher #008102) with 0.2 % Triton X-100 (Thermo Fisher #85111) then incubation in primary antibody solution (antibody diluted in CAS-block, 0.2 % Triton X-100) overnight at 4 °C. On the following day, the clusters were washed three times for 5 min in PBS containing 0.1 % Tween-20 (PBS-T) (Sigma #P4417) and incubated in appropriate secondary antibody solution (antibody diluted in PBS-T and DAPI (1:1000)) for 2 h at room temperature.

Clusters were then washed 2 times for 5 min in PBST and 1 time for 5 min in PBS and mounted with Vectashield (Vector #H2000) on glass slides. (ii) fixed clusters for cryo-sectioning were incubated overnight in 30 % sucrose (Sigma #S0389) before embedding in tissue-tek OCT (Sakura #4583) and storing at -80 °C for minimum 2 h. OCT-blocks containing fixed clusters were cryo-sectioned (10 μm thickness) and transferred to glass slides. Blocking and staining of cryo-sections proceeded as per whole mount staining protocol above. Antibody dilutions were prepared as indicated in Table 3. Images were acquired using confocal microscopy (Carl Zeiss LSM 800) using 10, 20 and 40 X objectives. Where appropriate, mean fluorescence intensity of individual clusters was calculated using Image J.

Single cell and bulk RNA-seq

Single cell RNA-seq libraries were generated using the 10x Genomics 3' end platform. Sequencing reads were processed using Cell Ranger (version 2.2.0) with the GRCh38p13 genome assembly to generate unique molecular identifier (UMI) gene count matrices per sample. The genome reference was supplemented with the eGFP coding sequence to enable detection of the pINS-eGFP transgene (GenBank U55761.1). Matrices were next processed using Seurat (version 2.3.0-3.0) to perform quality control filtering, normalization, tSNE projection, and clustering [41]. Cells were removed if the UMI count was less than 250, greater than 75,000, or if the proportion of UMIs mapped to mitochondrial genes was greater than 20% (Supp. Fig. 10). Genes were excluded if they were detectable in fewer than 5 cells. Following filtering, the UMI counts were normalized to library size (total number of UMIs detected), scaled by 10,000, and log-transformed. Principal component analysis was performed on the *Z* scores of the normalized

expression values, and the top 20 dimensions were selected for tSNE projection using a perplexity of 30. Graph-based clustering was performed using the top 20 principal components, with the 30 nearest neighbors, and a resolution of 0.5. Genes differentially expressed in each cluster compared with other clusters in each tested comparison were determined using a wilcox rank sum test and corrected for multiple hypothesis testing using Bonferroni correction (Seurat FindAllMarkers function). Cells were ordered in pseudotime using Monocle2 with the DDRTree method for dimensionality reduction (v2.10.0) [42]. RNA velocity estimates were computed using the velocyto Python package [43]. Canonical correlation analysis was performed using the RunCCA and AlignSubspace Seurat commands.

Bulk RNA seq

Total RNA was isolated from cell cultures using RNeasy kits from Qiagen. Sequencing libraries were generated using the NEBNext Ultra II Directional RNA Library kit with NEBNext rRNA depletion. Paired-end sequencing reads were trimmed using cutadapt (v1.16 [24], aligned using STAR (v 2.5.2a [25]), and exonic read counts quantified using featureCounts from the subread package (v1.6.2 [26]). Differentially expressed genes were identified using DESeq2 (v1.24.0 [27]). Heatmaps were generated using ComplexHeatmap and ordered using hierarchical clustering of Euclidean distances with the complete method [28].

GO Enrichment

Gene Ontology enrichment for single cell and bulk RNA-seq was conducted using gProfiler ([44]) using an ordered query with genes ranked by adjusted p-values.

Data and Code availability

All sequencing data was deposited in the NCBI's Gene Expression Omnibus database (GSE142290). Analysis scripts and an interactive UCSC cell browser are provided at a GitHub repository (https://github.com/rnabioco/sebeta).

Functional characterization

Calcium Imaging

Isolated clusters were loaded with 2 μM Rhod-2 AM (Invitrogen) for 35 min at 37°C in imaging medium (125 mM NaCl, 5.7 mM KCl, 2.5 mM $CaCl_2$, 1.2 mM $MgCl_2$, 10 mM HEPES, 2 mM glucose, and 0.1 % BSA, pH 7.4) and were imaged in 35 mm glass bottom dishes maintained at 37°C. Rhod-2 fluorescence was imaged on a confocal microscope (Carl Zeiss LSM 800) with a 20x 0.8 NA Plan Apochromat objective, 561 nm diode laser for excitation, and band pass emission filter of 568-700 nm. GFP fluorescence was imaged on the same microscope with a 488 nm diode laser for excitation with a band pass filter of 500-560 nm. Calcium images were acquired at ~1.5–3.5 sec/frame for 3 min at 2 mM glucose and for 10 min at 11 mM glucose after 20 min of glucose stimulation. Microscope settings (integration time, scan time, gain, laser power) were constant for all images collected within the same day. While some reports have indicated Rhod2 preferentially accumulates in mitochondria and thus reflects mitochondrial Ca^{2+}, we note that prior literature has demonstrated its cytoplasmic localization via electron microscopy and use for cytoplasmic Ca2+ imaging [45]. Indeed, our prior work has also demonstrated Rhod2 shows preferential cytoplasmic labeling within islets of Langerhans and reports on beta cell Ca2+ dynamics [30].

Image Analysis

All images were analyzed similarly to previously published methods [30] with custom Matlab (Mathworks) scripts.

Activity Analysis

Images were smoothed using a 5 x 5 pixel averaging filter. Areas without significant Rhod-2 fluorescence were removed. Saturated areas were also removed by limiting the area to intensity below the maximum value. Photobleaching was adjusted for by removing any linear trend. Any islets with significant motion artifacts were removed (displacement of > 0.5 cell width). For the time course of each 5 x 5 pixel region in the image with significant fluorescence, a peak detection algorithm was used to determine if the areas had peak amplitudes significantly above background. A region was considered 'active' if the corresponding time course for each pixel region had a peak amplitude >2.4x background. The fraction of active area was calculated as the number of pixels detected as 'active' normalized to the total number of pixels that showed significant fluorescence that were not saturated. Activity maps in Figure 2 display peak amplitude normalized to average value of each 5 x 5 pixel region over time. This is only shown for areas of the islet determined as active. Fold change was determined by calculating the ratio of activity at 11 mM glucose to activity at 2 mM glucose.

Coordinated Area Analysis

Coordinated area was only calculated for active area at 11 mM glucose. Coordination was determined based on coincident timing of identified peaks, where areas were segmented by identified peaks occurring at similar time points. The cross correlation of the time courses for two 5 x 5 pixel sub-regions were taken. If the Pearson's correlation coefficient was > 0.75, then the two sub-regions were considered highly coordinated and

merged into a larger region. The coordinated area was calculated as the number of pixels in the largest area of coordination across the islet normalized to the total number of pixels that showed significant fluorescence that were not saturated.

Statistical Analysis

All statistical analysis was performed in Prism (Graphpad) or Matlab. First an F-test was used to determine if variances were equal then a student's t-test or welch t-test (for unequal variance) were utilized for assessing differences in activity, fold change in activity and coordination. A paired t-test was performed for activity when detecting differences between 2 and 11 mM glucose for the same islet. IQR outlier analysis was performed on 2 mM data for the imBC and seBC groups and outliers were removed from all data sets. Outliers were identified as any data point outside of [Q1 - 1.5 x IQR, Q3 + 1.5 x IQR] for each group, where Q1 and Q3 are the first and third quartiles.

Perifusion assay

Dynamic insulin secretion was measured using a BioRep Technologies perifusion machine (PERI4-02-0230-FA-ORB). 20-30 sBC clusters or human islets were placed on a filter in the perifusion chamber and various solutions were perifused through the system at 100 μL/min by a peristaltic pump; cells and solutions were kept at 37 °C at ambient atmosphere. The perifusion program consisted of a 1.5 h preincubation step with KRB buffer containing 0.5 mM Glucose followed by alternating high (16.7 mM) glucose, low (0.5 mM) glucose, exendin-4 (10 nM or 10 mM), and KCl (30 mM) solutions. Perifusion flow-through was collected in 96 well plates and stored at 4 °C for future analysis. Cell pellets were recovered from the chamber after perifusion and lysed with acid/ethanol solution over night at 4 °C. We note, potential loss of sample material during this

technically challenging procedure resulting in high percentages of secreted hormones relative to total content.

Ratio calculation

Response to low glucose was calculated as the average insulin secretion read out for the initial 10 min low glucose incubation. The high glucose response was taken as the highest insulin secretion reached during the 20 min high glucose incubation. The KCl response was taken as the highest insulin secretion reached during the 5 min KCl incubation.

Additional references:

41. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single- cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol 2018;36:411–420

42. Trapnell C, Cacchiarelli D, Grimsby J, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotechnol 2014;32:381–386

43. La Manno G, Soldatov R, Zeisel A, et al. RNA velocity of single cells. Nature 2018;560:494–498

44. Raudvere U, Kolberg L, Kuzmin I, et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Res 2019;47:W191–W198

45. MacGowan GA, Du C, Glonty V, Suhan JP, Koretsky AP, Farkas DL. Rhod- 2 based measurements of intracellular calcium in the perfused mouse heart: cellular and subcellular localization and response to positive inotropy. J Biomed Opt 2001;6:23–30