

# Recapitulating endocrine cell clustering in culture promotes maturation of human stem-cell-derived $\beta$ cells

Gopika G. Nair<sup>1</sup>, Jennifer S. Liu<sup>1</sup>, Holger A. Russ<sup>1,4</sup>, Stella Tran<sup>1,5</sup>, Michael S. Saxton<sup>1</sup>, Richard Chen<sup>1</sup>, Charity Juang<sup>1</sup>, Mei-lan Li<sup>1</sup>, Vinh Q. Nguyen<sup>2</sup>, Simone Giacometti<sup>1</sup>, Sapna Puri<sup>1</sup>, Yuan Xing<sup>3</sup>, Yong Wang<sup>3</sup>, Gregory L. Szot<sup>2</sup>, Jose Oberholzer<sup>3</sup>, Anil Bhushan<sup>1</sup> and Matthias Hebrok<sup>1\*</sup>

**Despite advances in the differentiation of insulin-producing cells from human embryonic stem cells, the generation of mature functional  $\beta$  cells in vitro has remained elusive. To accomplish this goal, we have developed cell culture conditions to closely mimic events occurring during pancreatic islet organogenesis and  $\beta$  cell maturation. In particular, we have focused on recapitulating endocrine cell clustering by isolating and reagggregating immature  $\beta$ -like cells to form islet-sized enriched  $\beta$ -clusters (eBCs). eBCs display physiological properties analogous to primary human  $\beta$  cells, including robust dynamic insulin secretion, increased calcium signalling in response to secretagogues, and improved mitochondrial energization. Notably, endocrine cell clustering induces metabolic maturation by driving mitochondrial oxidative respiration, a process central to stimulus–secretion coupling in mature  $\beta$  cells. eBCs display glucose-stimulated insulin secretion as early as three days after transplantation in mice. In summary, replicating aspects of endocrine cell clustering permits the generation of stem-cell-derived  $\beta$  cells that resemble their endogenous counterparts.**

Pancreatic  $\beta$  cells are highly specialized nutrient sensors that efficiently regulate blood glucose levels, and their destruction and/or dysfunction causes type 1 and type 2 diabetes. Current therapy involves exogenous insulin administration that cannot fully replicate the rigorous glycaemic control provided by endogenous  $\beta$  cells. Islet transplantation can serve as an effective intervention for restoring normoglycaemia, but the demand for islets far outstrips the supply<sup>1</sup>. Considering the unlimited potential of human pluripotent stem cells (hPSCs) for self-renewal, generation of functional  $\beta$  cells from hPSCs has emerged as an attractive alternative. Although recent reports<sup>2–4</sup>, including studies from our own group, describe  $\beta$  and  $\beta$ -like cell formation from hPSCs, these cells possess limited functionality. And while a minority of  $\beta$ -like cells display calcium responses to glucose, these are slower and smaller in amplitude compared to adult islets<sup>3</sup>.  $\beta$ -like cells also do not terminate calcium influx when glucose is withdrawn, and fail to rapidly secrete insulin in dynamic perfusion assays, indicating an absent or delayed first-phase insulin secretion<sup>3</sup>. Dynamic insulin secretion tests demonstrating appropriate temporal responses to glucose stimulation were not reported for other hPSC-derived insulin-secreting cells<sup>3,4</sup>. Unlike human islets that function immediately, these hPSC-derived insulin-secreting cells respond to in vivo glucose challenges only 2–6 weeks after transplantation<sup>2–4</sup>. Despite progress in generating insulin-producing cells, developing conditions conducive for the maturation of hPSC-derived  $\beta$  cells in vitro without genetic modifications has been difficult<sup>5</sup>.

$\beta$  cell maturation occurs gradually during rodent<sup>6–9</sup> and human<sup>10,11</sup> postnatal development. The process is characterized by acquisition of robust glucose-stimulated insulin secretion (GSIS) at

the correct physiological set point to prevent hypo- and hyperglycaemia<sup>6,9</sup>. Dramatic changes in the  $\beta$  cell have been associated with maturation, including both enhanced expression of transcription factors such as *Neurod*, *Nkx6.1*, *Pdx1*, *Pax6*, *Mafb* and *Mafa*<sup>7,12–15</sup> and repression of ‘disallowed’ genes such as *Hk1*, *Ldha* and *Mct1*<sup>9,16</sup>. A metabolic switch from anaerobic glycolysis to mitochondrial oxidative phosphorylation in maturing cells is key to enabling and enhancing  $\beta$  cell function. For example, ectopic expression of a mitochondrial regulator, ERR $\gamma$ , led to enhanced function of hPSC-derived  $\beta$ -like cells<sup>5</sup>. Although evidence indicates changes in gene expression between young and adult human  $\beta$  cells<sup>10</sup>, the precise temporal dynamics of human  $\beta$  cell maturation remain unclear.

In addition to the intracellular processes described above, significant cellular reorganization involving endocrine cell clustering, compaction and formation of definitive islets proceed well into postnatal development and coincide with maturation. Studies of pancreas organogenesis in vivo have shown that endocrine hormone-expressing cells first appear scattered in the pancreatic epithelium, subsequently delaminating and aggregating into islet-like clusters where they acquire functional maturity<sup>11,17–19</sup>. In contrast, in vitro hPSC differentiation protocols produce a mixed population of PDX1<sup>+</sup>/NKX6.1<sup>+</sup> pancreatic progenitors intermingled with differentiating endocrine cells that are immature, possibly because delamination and clustering events occurring during endogenous islet formation and maturation are not incorporated.

Here, we aimed to mimic endogenous endocrine cell clustering by reagggregating immature  $\beta$ -like cells generated from hPSCs into enriched  $\beta$ -clusters (eBCs). eBCs display enhanced functional properties in vitro and in vivo, analogous to human islet  $\beta$  cells.

<sup>1</sup>Diabetes Center, University of California San Francisco, San Francisco, CA, USA. <sup>2</sup>Department of Surgery, University of California San Francisco, San Francisco, CA, USA. <sup>3</sup>Department of Surgery/Division of Transplantation, University of Virginia, Charlottesville, VA, USA. <sup>4</sup>Present address: Barbara Davis Center for Diabetes, University of Colorado, School of Medicine, Aurora, CO, USA. <sup>5</sup>Present address: Lawrence Berkeley National Laboratory, University of California-Berkeley, Berkeley, CA, USA. \*e-mail: [mhebrok@diabetes.ucsf.edu](mailto:mhebrok@diabetes.ucsf.edu)

Coalescence of immature  $\beta$ -like cells induces metabolic maturation and drives mitochondrial oxidative respiration, processes reported to occur in rodent models. The resulting mature  $\beta$  cells closely resemble primary adult  $\beta$  cells by transcriptome, physiological and functional analyses. Our findings suggest endocrine cell clustering is a vital step in hPSC-derived  $\beta$ -like cell maturation in cell culture.

## Results

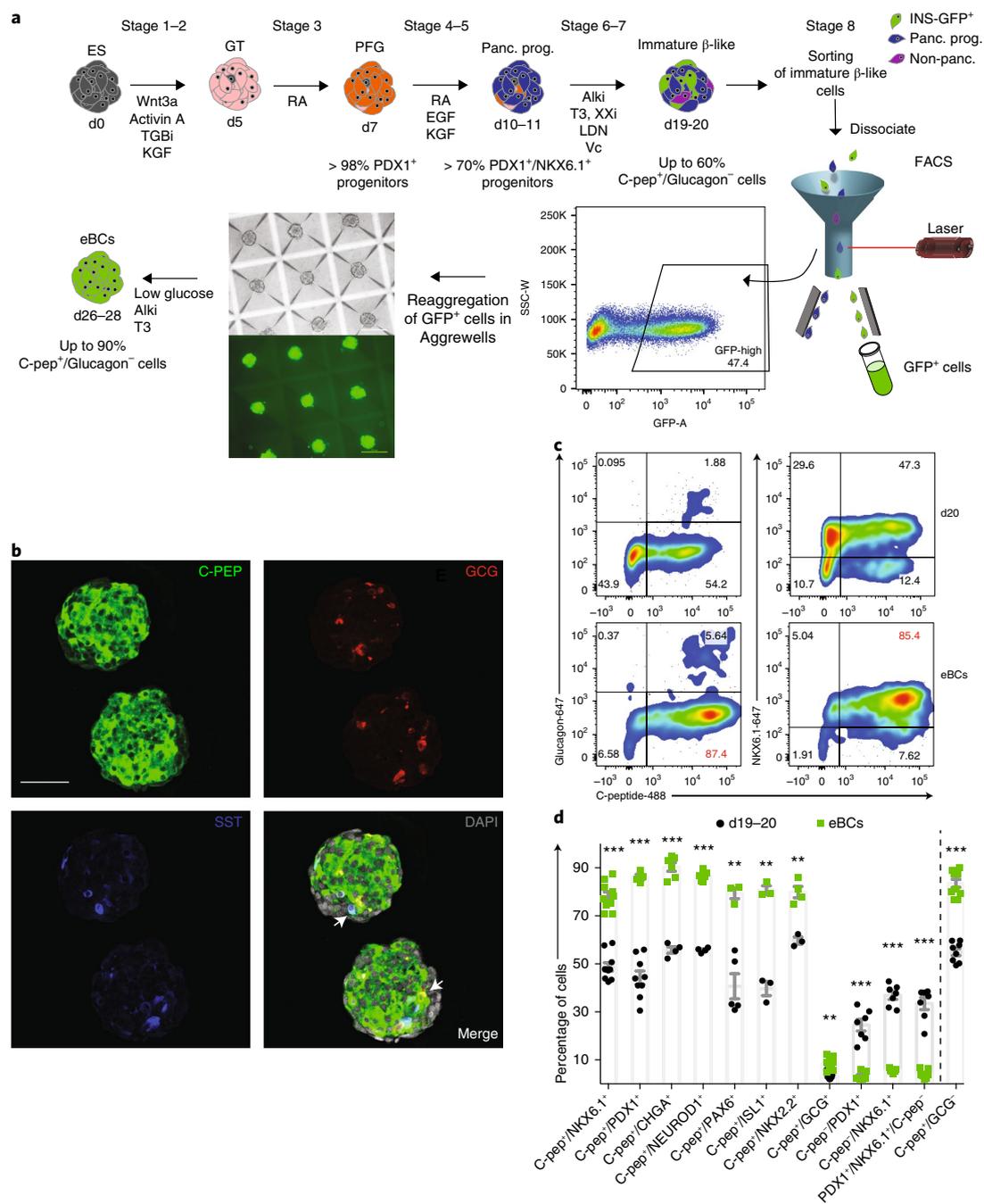
**Recapitulating endocrine cell clustering promotes formation of islet-like structures in vitro.** As described previously, current hPSC-to- $\beta$ -cell differentiation protocols generate a mixed population of cells at various developmental stages, including cells committed to the endocrine lineage, PDX1<sup>+</sup>/NKX6.1<sup>+</sup> pancreatic progenitors and potentially rare undifferentiated stem cells. These protocols do not recapitulate the critical process of endocrine cell clustering that occurs during islet formation in vivo, which is a prerequisite for functional maturation<sup>17,18,20</sup>. We sought to mimic this process in vitro by reaggregating hPSC-derived  $\beta$ -like cells into clusters. Using the INS<sup>GFP/W</sup> reporter cell line, we first optimized our previously published differentiation protocol<sup>4</sup> to produce sufficient numbers of monohormonal C-peptide<sup>+</sup> cells in 19–20 days (d19–20) (mean  $\pm$  s.e.m. = 54.8  $\pm$  1.5%) (Fig. 1a). A large proportion of the cells are co-positive for C-peptide and key  $\beta$  cell markers (Supplementary Table 1, Fig. 1d and Supplementary Fig. 1a). However, these monohormonal C-peptide<sup>+</sup> cells are not fully functional (Fig. 2a) and therefore represent immature  $\beta$ -like cells. On comparison of multiple functional aspects as well as gene expression profiles, these d19–20 cells appear to correspond closely to the end stage cells reported in current protocols<sup>2–4</sup>. Next, we isolated live INS<sup>GFP/W</sup>  $\beta$ -like cells (~45–50% of d20 cells) by fluorescence-activated cell sorting (FACS) and reaggregated them into clusters in pre-patterned chambers of Aggrewell-400 plates at a density of 1,000 cells per cluster to mimic the average number of  $\beta$  cells in one human islet equivalent<sup>21</sup> (Fig. 1a). The GFP-sorted cells self-organized into ~100  $\mu$ m-sized clusters after two days (Fig. 1a), and were cultured for an additional five to seven days (d26–28). Given the high percentage of C-peptide<sup>+</sup> cells in these clusters, we henceforth refer to them as eBCs.

eBCs were exclusively endocrine. 99% of cells expressed chromogranin A (Supplementary Fig. 1a) and all cells stained for synaptophysin (Supplementary Fig. 2, bottom). Duct and acinar cells marked by SOX9 (which also marks pancreas progenitor cells; Supplementary Fig. 2, middle), CK19 or amylase (Supplementary Fig. 2, bottom), were absent. Within the endocrine subtypes, eBCs mainly consisted of monohormonal C-peptide<sup>+</sup> cells, with rare cells co-expressing C-peptide and glucagon or somatostatin (Fig. 1b). Quantitative assessment via flow cytometry revealed a significant enrichment of C-peptide<sup>+</sup> and C-peptide<sup>+</sup>/NKX6.1<sup>+</sup> double-positive cells in the eBCs compared to d20 clusters (Fig. 1c). 85% (mean  $\pm$  s.e.m. = 78.1  $\pm$  1.8%,  $n$  = 10 independent experiments) of eBC cells co-expressed C-peptide and NKX6.1, the transcription factor essential for maintaining the functional state of mature  $\beta$  cells<sup>22</sup>. Nuclear staining of NKX6.1 and PDX1 was observed in C-peptide<sup>+</sup> cells (Supplementary Fig. 2, top). Approximately 90% (mean  $\pm$  s.e.m. = 83.5  $\pm$  1.7%,  $n$  = 10 independent experiments) of the cells were C-peptide<sup>+</sup>/glucagon<sup>-</sup>, ~4–12% (mean  $\pm$  s.e.m. = 8.6  $\pm$  0.9%) were C-peptide<sup>+</sup>/glucagon<sup>+</sup> (Fig. 1c) and ~8% were C-peptide<sup>+</sup>/somatostatin<sup>+</sup> (Supplementary Fig. 1a). A time-course analysis of endocrine differentiation from pancreatic progenitors (d11) to eBCs (d27) by flow cytometry demonstrated the progressive appearance of C-peptide<sup>+</sup> cells and increasing co-expression of endocrine/ $\beta$ -cell-specific markers of maturation, such as the transcription factors NEUROD1<sup>12</sup>, ISL1<sup>23</sup> and PAX6<sup>13</sup>, and the dense core granule component CHGA<sup>24</sup> (Supplementary Fig. 1a). We observed a significant increase in the intensity of C-peptide expression in eBC  $\beta$  cells compared with d20  $\beta$ -like cells (Supplementary

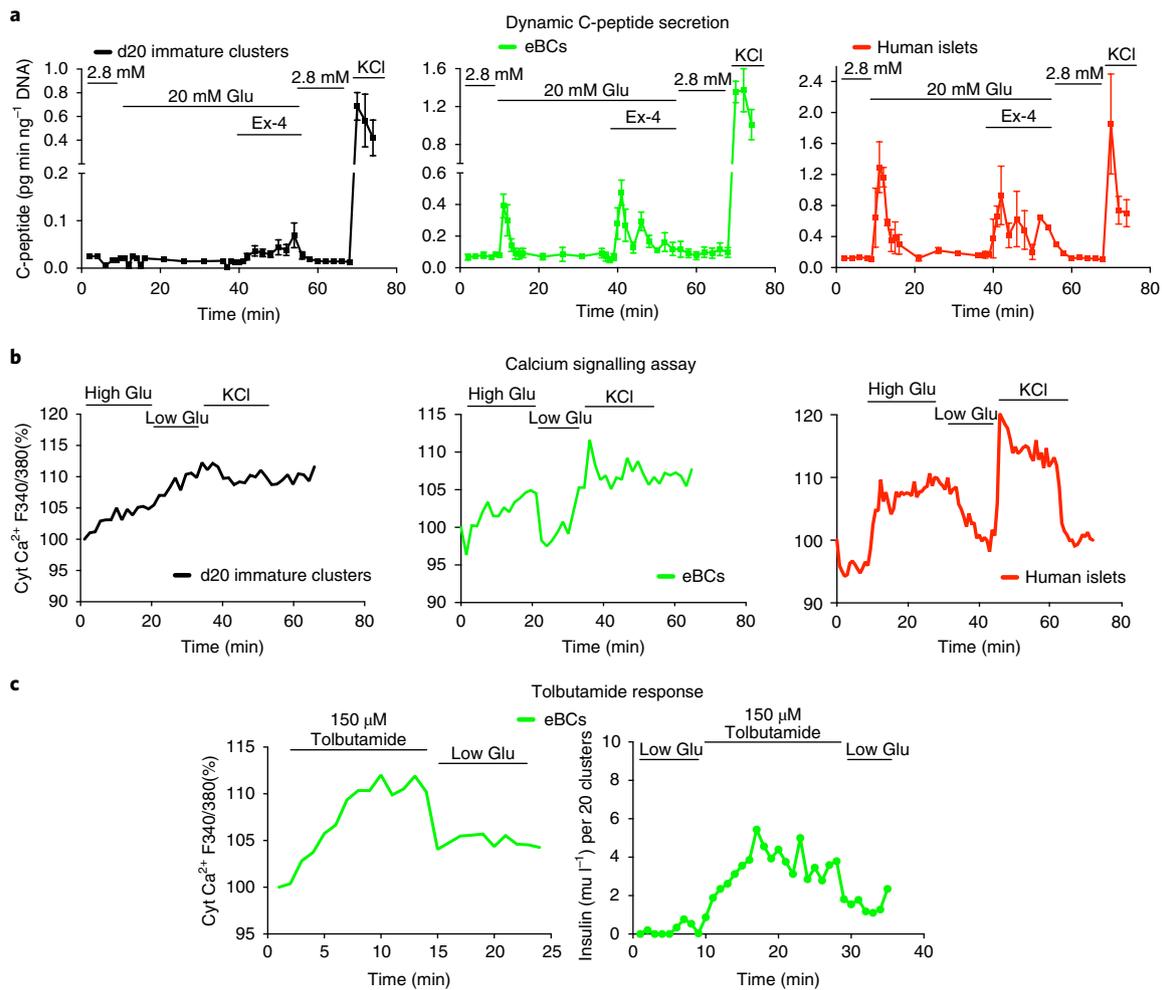
Fig. 1a; quantified in Supplementary Fig. 1b). Similar patterns were also observed in the intensities of PDX1, ISL1, PAX6 and CHGA (Supplementary Fig. 1a). The robustness of our protocol is indicated by the consistent generation of eBCs enriched for C-peptide and key  $\beta$  cell markers over ten batches of differentiation (Fig. 1d and Supplementary Table 1). eBCs were depleted of undifferentiated pancreatic epithelial progenitors that may give rise to tumours in vivo (PDX1<sup>+</sup>/NKX6.1<sup>+</sup>/C-peptide<sup>-</sup>) (Fig. 1d, Supplementary Table 1 and Supplementary Fig. 1a). Our data indicate that reaggregation of immature  $\beta$ -like cells followed by extended culture results in islet-like endocrine clusters with enhanced expression of  $\beta$  cell markers.

**eBCs display physiological properties similar to human islets in vitro.** Current hPSC-derived  $\beta$  cells<sup>2–4</sup> do not respond well to dynamic glucose challenges. To determine whether the enhanced expression of  $\beta$  cell markers upon reaggregation correlates with mature function, we tested the dynamic response of eBCs to secretagogues such as glucose, KCl and GLP-1 in a perfusion system. A rapid and marked first-phase response to glucose was observed similar to human islets, albeit at somewhat lower levels (Fig. 2a). In contrast, d20 clusters did not respond to glucose, confirming they are immature  $\beta$ -like cells. The second-phase response in eBCs was not sustained. Similarly to human islets, eBCs also responded to the GLP-1 receptor agonist exendin-4 in the presence of high glucose (Fig. 2a). Area under the curve (AUC) analyses revealed responses of eBCs as significantly higher than that of d20 clusters (Supplementary Fig. 3a). Insulin secretion and transcription are less than that observed with human islets (Supplementary Figs. 3a and 4a), a defect that is likely due to the loss of one insulin allele that is replaced by GFP in the INS<sup>GFP/W</sup> line. To discern whether this improved functional property is due to endocrine cell clustering or extended culture of eBCs, we used d27 non-enriched clusters (NECs) that were reaggregated in the same manner as eBCs but without GFP sorting. NECs responded to KCl but not to glucose, indicating that endocrine cell sorting and clustering confers mature  $\beta$  cell function to eBCs (Supplementary Fig. 3b). Maturation takes 7–8 days following endocrine cell clustering since eBCs are not yet functional two days after reaggregation (Supplementary Fig. 3c). Successive glucose stimulations resulted in distinct first-phase responses, indicating the presence of a readily releasable pool of granules primed for secretion in eBCs (Supplementary Fig. 3d). As expected from mature  $\beta$  cells, insulin secretion stops at low glucose concentrations.

To further assess the functional maturation status of eBCs, we investigated the cascade of molecular events preceding insulin secretion in detail. First, we examined the cytosolic calcium flux essential for stimulus–secretion coupling. d20 clusters, which contain progenitors and endocrine cells, displayed a slow rise in calcium flux upon stimulation with high glucose, which did not terminate after the reduction of glucose levels to baseline (black graph in Fig. 2b; see also Supplementary Fig. 3e). This observation is consistent with those previously reported in other  $\beta$ -like cells<sup>3</sup>. In contrast, all eBCs showed a significant and rapid increase in calcium influx following stimulation with high glucose and KCl, and the flux returned to baseline in many eBCs when the glucose concentration was lowered. eBCs thus had a profile closely resembling human islets (Fig. 2b and Supplementary Fig. 3e), suggesting appropriate activity of the voltage-dependent calcium channels. Less optimal responses in calcium flux observed in some eBCs are also commonly found in human islet preparations (Supplementary Fig. 3f). Next, we assessed the robustness of the K<sup>+</sup>-ATP channels that close in response to an increase in ATP/ADP ratio, a step preceding calcium influx. eBCs rapidly increased calcium flux and secreted insulin upon tolbutamide treatment, a sulfonylurea that closes K<sup>+</sup>-ATP channels (Fig. 2c). eBCs thus display many physiological properties of mature adult  $\beta$  cells, including dynamic insulin secretion, calcium signalling and sulfonylurea response.



**Fig. 1 | Generation of islet-like eBCs. a**, A schematic representation of the directed differentiation of hESCs to pancreatic  $\beta$  cells from d0 to d26–28. The process recapitulates embryonic development and includes a strategy of reaggregation of immature  $\beta$ -like cells to mimic  $\beta$  cell coalescence during islet formation. Immature  $INS^{GFP+}$   $\beta$ -like cells are isolated by FACS from d19–20 spheres, and subsequently reaggregated into 100  $\mu$ m-sized eBCs in Aggrewells. Culture conditions are shown beneath each arrow. Bright-field and fluorescence images of eBCs in Aggrewells are shown; scale bars, 200  $\mu$ m. ES, embryonic stem cells (dark grey); GT, gut tube (pink); PFG, pancreatic foregut (orange); Panc. prog., pancreatic progenitors (blue);  $INS^{GFP+}$ , GFP-positive insulin-producing cells (green); Non-panc., non-pancreatic cells (purple). **b**, Immunofluorescent staining of d26–27 eBCs for C-peptide (C-PEP), glucagon (GCG), somatostatin (SST) and nuclei (DAPI), indicating that the majority of cells are C-peptide<sup>+</sup> and monohormonal. White arrows indicate C-peptide<sup>+</sup>/glucagon<sup>+</sup> or C-peptide<sup>+</sup>/somatostatin<sup>+</sup> cells. Scale bar, 100  $\mu$ m. Images are representative of four independent experiments. **c**, Quantitative analysis of d20 clusters and d26–27 eBCs by flow cytometry showing enrichment and confirming that the majority of the cells are C-peptide<sup>+</sup>/NKX6.1<sup>+</sup>/glucagon<sup>-</sup> in eBCs. Flow experiments were repeated 10 times independently with similar results. **d**, Flow cytometric quantification of co-expression of C-peptide and various key  $\beta$  cell markers such as NKX6.1, PDX1, CHGA, NEUROD1, PAX6, ISL1 and NKX2.2, and C-peptide<sup>+</sup>/glucagon<sup>-</sup> cells before (d19–20, black circles: C-pep<sup>+</sup>/NKX6.1<sup>+</sup>,  $n=10$ ; C-pep<sup>+</sup>/PDX1<sup>+</sup>,  $n=9$ ; C-pep<sup>+</sup>/CHGA<sup>+</sup>, C-pep<sup>+</sup>/NEUROD1<sup>+</sup>,  $n=4$ ; C-pep<sup>+</sup>/PAX6<sup>+</sup>,  $n=6$ ; C-pep<sup>+</sup>/ISL1<sup>+</sup>, C-pep<sup>+</sup>/NKX2.2<sup>+</sup>,  $n=3$ ; C-pep<sup>+</sup>/GCG<sup>+</sup>, C-pep<sup>-</sup>/PDX1<sup>+</sup>, C-pep<sup>-</sup>/NKX6.1<sup>+</sup>, PDX1<sup>+</sup>/NKX6.1<sup>+</sup>/C-pep<sup>-</sup>,  $n=7$ ; C-pep<sup>+</sup>/GCG<sup>-</sup>,  $n=8$  biological samples) and after sorting and reaggregation (eBCs, green squares: C-pep<sup>+</sup>/NKX6.1<sup>+</sup>, C-pep<sup>+</sup>/GCG<sup>-</sup>,  $n=10$ ; C-pep<sup>+</sup>/PDX1<sup>+</sup>, C-pep<sup>+</sup>/NKX2.2<sup>+</sup>,  $n=4$ ; C-pep<sup>+</sup>/CHGA<sup>+</sup>, C-pep<sup>+</sup>/NEUROD1<sup>+</sup>,  $n=6$ ; C-pep<sup>+</sup>/PAX6<sup>+</sup>, C-pep<sup>+</sup>/ISL1<sup>+</sup>,  $n=3$ ; C-pep<sup>+</sup>/GCG<sup>+</sup>,  $n=8$ ; C-pep<sup>-</sup>/PDX1<sup>+</sup>, C-pep<sup>-</sup>/NKX6.1<sup>+</sup>, PDX1<sup>+</sup>/NKX6.1<sup>+</sup>/C-pep<sup>-</sup>,  $n=7$  biological samples). Data are presented as mean  $\pm$  s.e.m. See Supplementary Table 6 for source data. **\*\*** $P < 0.01$ , **\*\*\*** $P < 0.001$  determined using the Holm–Sidak method for multiple  $t$  two-sided tests, with  $\alpha = 0.05$ . See the Methods section for exact  $P$  values.

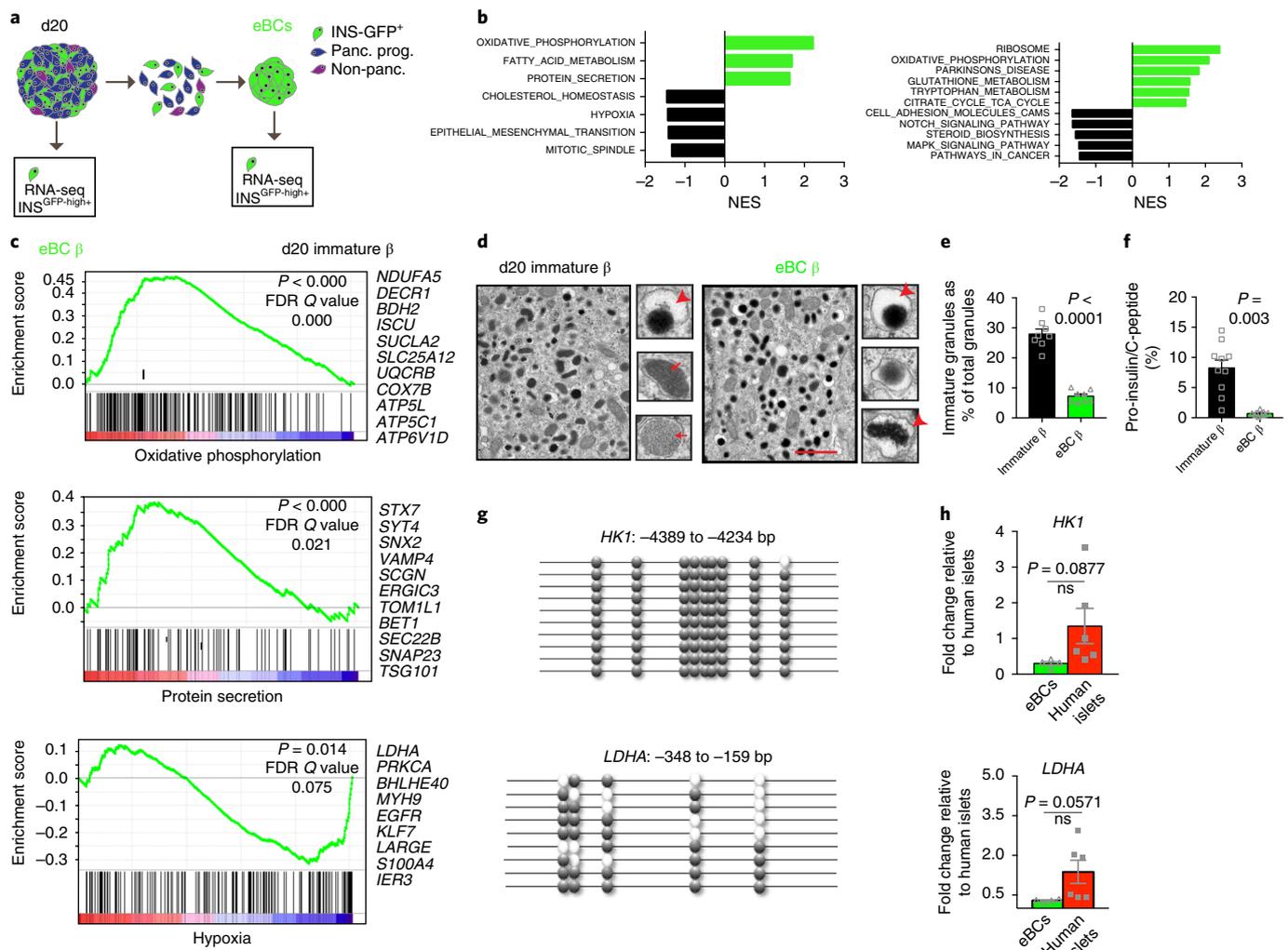


**Fig. 2 | eBCs exhibit functional characteristics similar to human islets in vitro.** **a**, Dynamic secretion of C-peptide in response to stimulation with 20 mM glucose, 10 nM exendin-4 (Ex-4) and 30 mM KCl in an in vitro perfusion assay with a starting basal glucose concentration of 2.8 mM.  $n = 3$  independent samples. Data are presented as mean  $\pm$  s.e.m. **b**, Cytosolic calcium signalling in response to alternating high (20 mM) and low (2.8 mM) glucose followed by KCl (30 mM) stimulation as measured by Fura-2/AM fluorescence emission intensity. Plots are population measurements from individual whole clusters (not pre-selected single cells). **c**, Calcium signalling and insulin secretion response of eBCs to tolbutamide, a sulfonylurea drug that blocks ATP-sensitive  $K^+$  channels. Calcium signalling analyses were performed with 5 independent samples of d20 clusters, 6 independent samples of eBCs and 6 independent islet preparations (Supplementary Fig. 3e,f). See Supplementary Table 6 for source data.

**Transcriptome-wide analysis of eBCs reveals hallmarks of function and maturation.** In vitro physiological analyses implied that clustering of  $\beta$  cells confers improved functional properties. To determine the gene expression changes induced in  $\beta$ -like cells following clustering, we performed qPCR analysis of sorted  $\beta$ -like cells ( $INS^{GFP-high+}$ ) from d20 clusters and eBCs (Supplementary Fig. 4a). Expression of canonical  $\beta$  cell markers such as *PDX1* and *NKX6.1* did not change. However, levels of *NEUROD1*, *PAX6*, *MAFB* and *SLC30A8*, implicated in maturation and maintenance of  $\beta$  cell identity and function<sup>10–13,15</sup>, were significantly higher in eBCs than in d20  $INS^{GFP-high+}$  cells. Expression of the endocrine progenitor marker *NEUROG3* was lower in eBCs than in d20  $INS^{GFP-high+}$  cells, further validating the immature nature of d20  $\beta$ -like cells. Compared to human islets, eBCs expressed similar levels of  $\beta$  cell regulators (*PDX1*, *NKX6.1*, *NKX2.2*, *PAX6* and *MAFB*), progenitor markers (*NEUROG3* and *SOX9*) and insulin synthesis and secretion machinery (*INS*, *SLC30A8*, *KCNJ11*, *GLIS3*, *PCSK2* and *GLUT-1*) (Supplementary Fig. 4a–c). The expression levels of some genes encoding the insulin secretion machinery (*ABCC8*, *PCSK1*, *KCNK3*) vary widely in human islet preparations, but were not statistically significantly different from the levels in eBCs. Given that

*MAFA* transcripts are not elevated in human islets until puberty<sup>10</sup>, it is not surprising that levels of this transcription factor were lower in eBCs than in adult human islets. Of note, *MAFB* levels are higher than *MAFA* levels in juvenile human  $\beta$  cells and eBCs (Supplementary Fig. 4a). Genes reported to be expressed at lower levels in previous protocols, including *PAX6*, *KCNK3*, *GLIS3* and *PCSK2*, are expressed similarly in eBCs and human islets/ $\beta$  cells (Supplementary Fig. 4). We observed equivalent levels of *GLUT1* (the predominate glucose transporter in human insulin-producing cells) in both eBCs and human  $\beta$  cells.

To further elucidate why eBCs display significantly better functional properties than d20 clusters, we conducted genome-wide transcriptome analysis of  $INS^{GFP-high+}$  cells sorted from eBCs and d20 clusters. Gene set enrichment analysis (GSEA) revealed that gene sets significantly upregulated in eBCs include oxidative phosphorylation, protein secretion, ribosomal constituents and various metabolic pathways such as the TCA cycle (Fig. 3a–c), all of which have been implicated in  $\beta$  cell maturation in mice<sup>5,25,26</sup>. Leading edge genes in the GSEA upregulated in eBCs are known to regulate aspects of  $\beta$  cell function, including components of the mitochondrial malate-aspartate shuttle (*SLC25A12/ARALARI*), ATP synthase subunits



**Fig. 3 | Enrichment of  $\beta$  cell function/maturation related genes and repression of disallowed genes in eBCs.** **a**, A schematic illustration depicting isolation of  $\text{INS}^{\text{GFP-high}+}$  cells from d20 clusters and eBCs for RNA sequencing (RNA-seq). **b**, Selected pathways enriched in eBC  $\text{INS}^{\text{GFP-high}+}$  cells (green) and d20  $\text{INS}^{\text{GFP-high}+}$  cells (black) by GSEA performed with ‘Hallmark’ (left) and ‘KEGG’ (right) gene sets from the Molecular Signatures Database (MSigDB). NES, normalized enrichment score. A full list of pathways is shown in Supplementary Table 5. **c**, Enrichment plots with genes in the leading edge of each pathway indicated on the right side of the respective trace. Note that only  $\text{INS}^{\text{GFP-high}+}$  cells were analysed for both the d20 immature and eBC populations. GSEA uses Kolmogorov–Smirnov statistics to calculate  $P$  values. **d**, Transmission electron micrograph of  $\beta$  cells present in d20 immature clusters and  $\beta$  cells found in eBCs. Higher magnification images (panels on the right of each main image) show representative types of individual granules present in the samples. Arrowheads, mature granules; arrows, immature granules. Scale bar,  $1\mu\text{m}$ . Images are representative from two independent experiments. **e**, Quantification of immature and mature granules using the metric described in **d** and expressed as immature granules as a percentage of total granules.  $n=8$  sections were analysed in each sample. Data are expressed as mean  $\pm$  s.e.m.  $P < 0.0001$  determined by two-tailed unpaired  $t$ -test with Welch’s correction. **f**, Percentage of pro-insulin to C-peptide in d20 immature clusters and eBCs. Immature d20 clusters,  $n=10$  independent samples; eBCs,  $n=9$  independent samples. Data are expressed as mean  $\pm$  s.e.m.  $P = 0.0003$  determined by two-tailed Welch’s  $t$  test. **g**, Bisulfite sequencing analysis for the disallowed genes *HK1* and *LDHA* at the indicated loci in eBCs. Each parallel line is an independent clone. Filled circles represent fully methylated and open circles represent hypomethylated CpGs. These regions are mostly methylated at both loci in eBCs. **h**, Gene expression of *HK1* and *LDHA* in eBCs relative to human islets. eBCs,  $n=4$  independent samples; human islets,  $n=6$  independent samples. Data are expressed as mean  $\pm$  s.e.m.  $P$  values determined by two-tailed Welch’s  $t$ -test. ns, not significant. See Supplementary Table 6 for source data.

(*ATP5L*, *ATP5C1*, *ATP6V1D*), SNARE protein complexes and syntaxins critical for insulin secretion (*SNAP23*, *VAMP4*), a syntaxin critical for insulin secretion (*SNAP23*, *VAMP4*), a syntaxin critical for insulin secretion (*SYT4*) and calcium sensors (*SCGN*) (Fig. 3c and Supplementary Table 5). These genes have not previously been implicated in the maturation of stem-cell-derived  $\beta$  cells<sup>2,3</sup>. In contrast, gene sets enriched in neonatal mouse and human fetal  $\beta$  cells<sup>28,29</sup> were upregulated in d20/immature  $\beta$ -like cells, including mitotic pathways, steroid biosynthesis, cell adhesion, hypoxia, Notch signalling and mitogen-activated protein kinase (MAPK) (Fig. 3a–c).

Since protein secretion and ribosome gene sets were upregulated in  $\text{INS}^{\text{GFP-high}+}$  cells of eBCs, we investigated the structure and composition of the main cargo of  $\beta$  cells, the insulin granules. The eBC  $\beta$  cells contained granules with the typical electron dense core and electron light halo (Fig. 3d) and had significantly lower percentages of immature granules (marked by diffuse electron density) than d20  $\beta$ -like cells (Fig. 3e). The ratio of fully processed insulin versus pro-insulin was significantly higher in eBC  $\beta$  cells compared to d20  $\beta$ -like cells (Fig. 3f). These results suggest the presence of mature granules in eBC  $\beta$  cells containing fully processed insulin.

Emerging evidence indicates that  $\beta$  cell maturation also requires the reduction of 'disallowed' genes that interfere with glucose sensing. Since DNA methylation has been implicated in repression of gene expression<sup>9</sup>, we analysed CpG-rich regions within loci of the critical glucose-secretion decoupling genes *HK1* and *LDHA*. Bisulfite sequencing of differentially methylated regions of *HK1* and *LDHA* in eBCs showed hypermethylation similar to the pattern observed in human islets in previous reports (Fig. 3g)<sup>9</sup>. Analysis of the transcript levels of these genes in eBCs and human islets by qPCR confirmed their low expression (Fig. 3h). Furthermore, RNA-seq on *INS*<sup>GFP-high+</sup> cells isolated from immature d20 clusters, NECs and eBCs revealed higher expression of key disallowed genes (*HK2*, *LDHA* and *MCT1*) in d20  $\beta$ -like cells compared to eBC  $\beta$  cells (Supplementary Fig. 5). These findings suggest that a metabolic switch from anaerobic to aerobic glycolysis, reported during  $\beta$  cell maturation, has occurred in eBC  $\beta$  cells.

**$\beta$  cells residing in a highly enriched endocrine niche are distinct from  $\beta$  cells in a progenitor-rich niche.** Our data so far indicate that  $\beta$  cells in eBCs are more mature than those in d20 clusters. To determine if endocrine cell coalescence drives maturation, and to investigate the changes induced by this process, we analysed RNA-seq data from *INS*<sup>GFP-high+</sup> cells isolated from d20 clusters, NECs and eBCs as shown in Fig. 4a. Heat map and hierarchical clustering of the three types of  $\beta$  cell populations showed that  $\beta$  cells from d20 and d27 NEC populations cluster together and are distinct from  $\beta$  cells found in d27 eBCs (Fig. 4b). 1,068 genes were differentially expressed between  $\beta$  cells from d27 eBCs and d20 immature clusters, and 937 genes were differentially expressed between  $\beta$  cells from d27 eBCs and d27 NECs (fold change (FC)  $\geq 2$ ; false discovery rate (FDR) = 0.05). We conducted GSEA with gene sets related to OXPHOS, the top hit in the 'Hallmark' collection (Fig. 3b,c). Several gene sets, including 'electron transport chain (ETC)', 'mitochondrial respiratory complex 1 biogenesis', 'TCA cycle' and 'mitochondrial ATP synthesis coupled proton transport', were significantly upregulated in  $\beta$  cells of eBCs compared with  $\beta$  cells of both d20 immature clusters and d27 NECs (Fig. 4c). A heat map of the leading edge ETC genes enriched in eBC  $\beta$  cells is shown in Fig. 4d.

Finally, we compared the transcriptome of  $\beta$  cells from eBCs to  $\beta$  cells isolated from adult human islets using antibodies reported previously<sup>30</sup>. A scatter plot of all protein coding genes shows a high degree of correlation between the two types of  $\beta$  cells (Pearson correlation coefficient = 0.86 and  $P$  value  $< 2.2 \times 10^{-16}$ ) (Fig. 4e), highlighting that  $\beta$  cells within eBCs are nearly equivalent to primary  $\beta$  cells. There were 4,156 genes differentially expressed between d20  $\beta$ -like cells and islet  $\beta$  cells, and 4,280 between NEC  $\beta$  cells and islet  $\beta$  cells (FC  $\geq 2$ ; FDR = 0.05). In contrast, only 3,313 genes out of a total of ~20,000 were differentially expressed between eBC  $\beta$  cells and islet  $\beta$  cells under stringent conditions (FC  $\geq 2$ ; FDR = 0.05; Fig. 4f) and just 1,410 with relaxed exclusion criteria (FC  $\geq 5$ ). The presence of non- $\beta$  cells in eBCs (bihormonal insulin/glucagon and insulin/somatostatin co-positive cells; Supplementary Fig. 1a) account for part of the differences in gene expression between eBCs and islet  $\beta$  cells. GSEA of the pathways enriched in Fig. 4 yielded lower NES scores and higher FDR Q values when the comparison was made between islet  $\beta$  cells and eBC than between islet  $\beta$  cells and NEC/d20  $\beta$ -like cells (Supplementary Table 6), further supporting that eBC  $\beta$  cells are more mature than NEC/d20  $\beta$ -like cells. Thus, by multiple criteria, reaggregation of immature *INS*<sup>GFP+</sup> cells followed by extended culture results in mature functional  $\beta$  cells closely resembling human islet  $\beta$  cells.

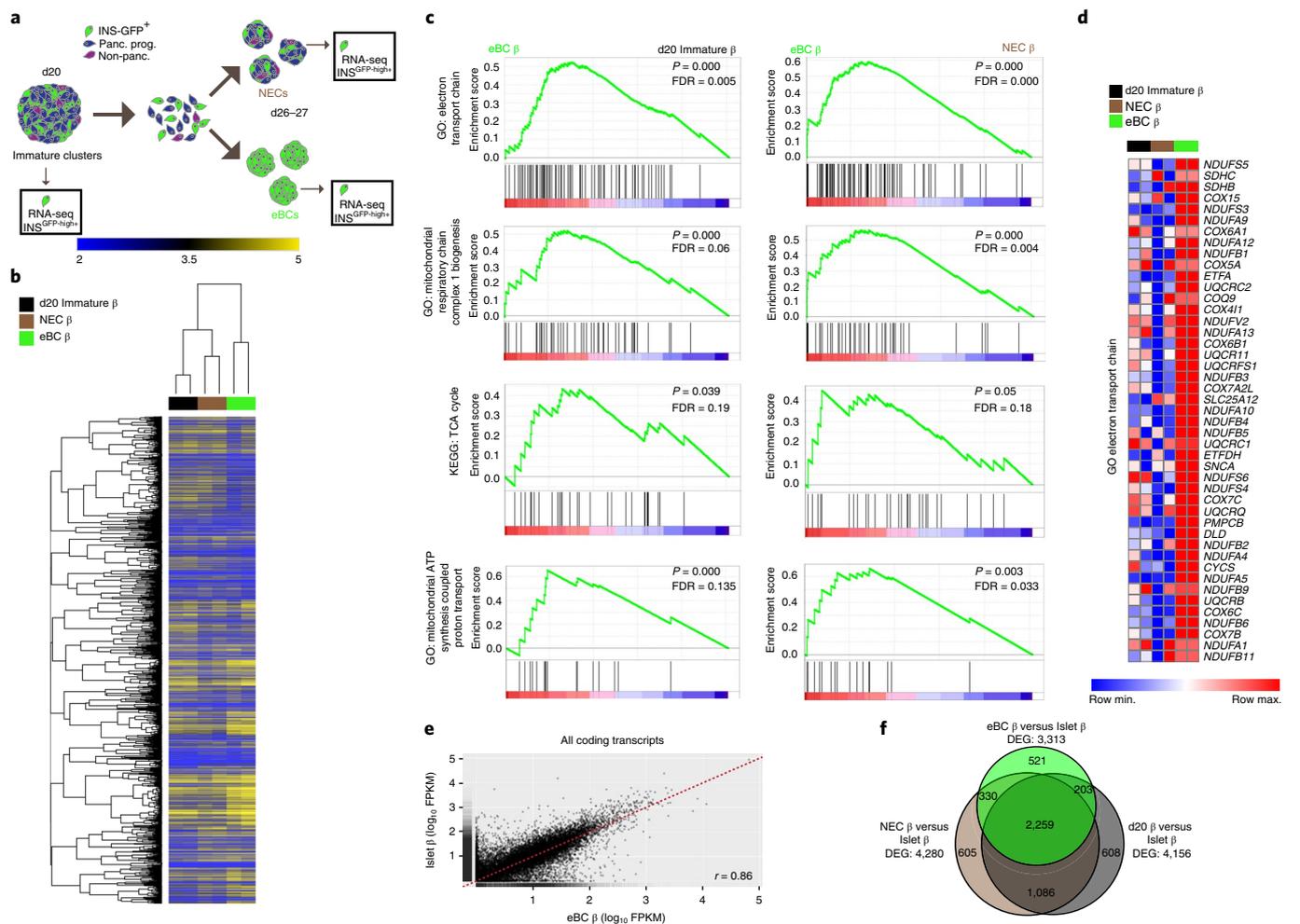
**$\beta$  cell clustering induces metabolic maturation of mitochondria.** RNA-seq results implicate activation of OXPHOS, electron transport chain and ATP production as potential mechanisms responsible for the maturation of eBCs. We assessed mitochondrial

respiratory function using a Seahorse XFe24 analyser. Both eBCs (Fig. 5a, centre) and human islets (Fig. 5a, bottom), but not d20 immature clusters (Fig. 5a, top), increased their oxygen consumption rate (OCR) upon stimulation with high glucose. The d20 clusters had almost no spare respiratory capacity (SRC), suggesting their mitochondria are immature. NEC clusters also did not show a significant glucose-induced rise in OCR or SRC (Supplementary Fig. 6a), indicating that  $\beta$  cell clustering and not extended culture is responsible for improved mitochondrial function. Basal OCR in eBCs and islets were similar and significantly lower than immature clusters when normalized to C-peptide content. Basal OCR levels of eBCs were higher than of islets when normalized to DNA (Fig. 5b). The capacity of eBCs, but not immature  $\beta$ -like cells or NECs, to raise OCR as efficiently as human islets in response to glucose suggests that they are metabolically mature and effectively engage mitochondrial oxidative phosphorylation for glucose oxidation.

Next, we interrogated mitochondrial energization, an event upstream of ATP production, by measuring the quenching of rhodamine-123 fluorescence during glucose stimulation. The rate of fluorescence decay is directly proportional to the membrane potential of mitochondria ( $\Delta\psi_m$ ). Membrane potentials of mitochondria in eBCs and human islets achieved similar levels, whereas those of d20 clusters were substantially lower on stimulation (Fig. 5c). This indicates the change in mitochondrial charge upon glucose stimulation is higher in eBCs than d20 clusters. We confirmed the validity of rhodamine-123 quenching experiments using FCCP and oligomycin (Supplementary Fig. 6b). To directly assess mitochondrial function at the single-cell level, we measured the membrane potential of mitochondria upon glucose stimulation in end-point analyses using MitoTracker Red CM-H<sub>2</sub>XRos, a dye that fluoresces only upon oxidation. Mitochondria in C-peptide<sup>+</sup> and C-peptide<sup>-</sup>/NKX6.1<sup>+</sup> double-positive cells in both eBCs and human islets are differentially charged in low and high glucose, a phenotype absent from the immature cells in d20 or d27 NECs (Fig. 5d). The specificity of MitoTracker Red CM-H<sub>2</sub>XRos binding to the mitochondria was confirmed by co-localization with a mitochondrial marker, TOMM20 (Supplementary Fig. 6c). These findings reveal that reaggregation of  $\beta$ -like cells promotes mitochondrial oxidative metabolism.

**$\beta$  cell clustering changes mitochondrial morphology and increases mitochondrial mass.** GSEA indicated significant enrichment of the components of the 'inner mitochondrial membrane complexes', the location of OXPHOS, in  $\beta$  cells following reaggregation (Fig. 6a). Examination of the ultrastructure of mitochondria in  $\beta$  cells of eBCs revealed highly folded (with a narrow gap between the two surfaces of the inner membrane) and stacked cristae, which indicate properly organized OXPHOS super-complexes necessary for efficient electron transport and ATP generation (Fig. 6b)<sup>31</sup>. By comparison, immature  $\beta$ -like cells possess mitochondria that appear less organized (Fig. 6b; high resolution in Supplementary Fig. 6d). The average inner-membrane length per mitochondrion in eBCs was significantly greater than in immature  $\beta$ -like cells (Fig. 6c). The enhanced mitochondrial function and ultrastructure also correlated with increased mitochondrial mass/number in eBCs (Fig. 6d–f). Mitochondrial mass, measured either in the whole population or only in the presumable  $\beta$  cells (C-peptide<sup>+</sup>/NKX6.1<sup>+</sup>), was higher in eBCs and human islets compared to d20 immature clusters or NECs (Fig. 6d,e). ERR $\gamma$ , a mitochondrial regulator recently reported to induce metabolic maturation<sup>5</sup>, was expressed at higher levels in eBC  $\beta$  cells than immature d20  $\beta$ -like cells and d27 NEC  $\beta$  cells (Fig. 6g).

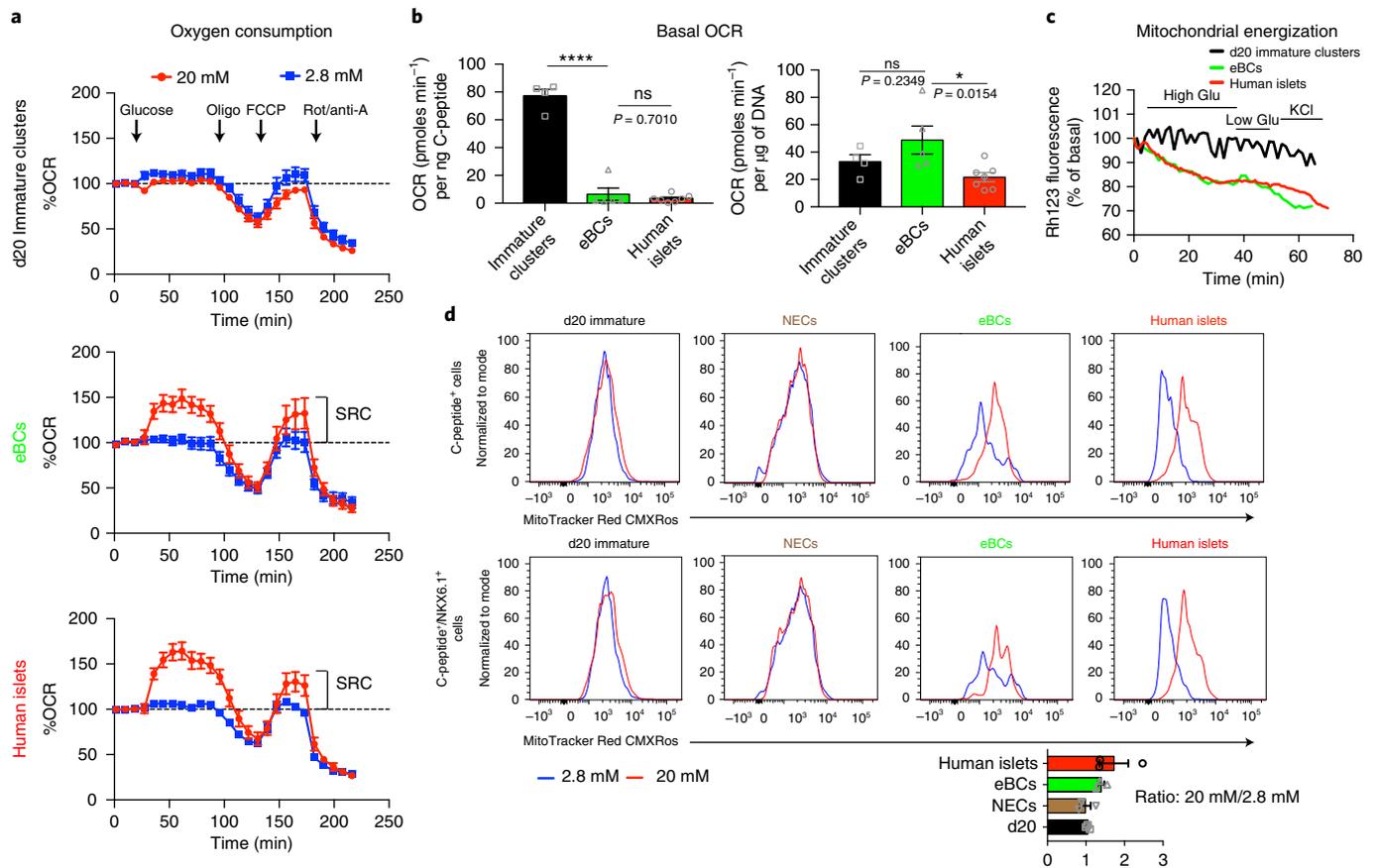
**Rapid onset of eBC functionality in vivo within days of transplantation.** A distinctive feature of human islets and mature  $\beta$  cells is their ability to function in vivo within days of transplantation. To test in vivo function, we transplanted 700 eBCs (~700,000 cells), of which ~80% are C-peptide<sup>+</sup>/NKX6.1<sup>+</sup> (~560,000  $\beta$  cells) under the



**Fig. 4 |  $\beta$  cells in a highly enriched endocrine niche of eBCs are distinct from immature  $\beta$  cells in a progenitor-rich niche.** **a**, A schematic illustration depicting the isolation of  $INS^{GFP-high+}$  cells via FACS from d20 immature clusters, d26–27 NECs and d26–27 eBCs for RNA-seq. Note that only  $INS^{GFP-high+}$  cells were used for transcriptome analysis. **b**, Heat map of differentially expressed genes between the three types of  $\beta$  cells. Hierarchical clustering indicates that d20 immature  $\beta$ -like cells (black) and NEC  $\beta$  cells (brown) are more closely related than  $\beta$  cells from eBCs (green) ( $n = 2$  independent samples). **c**, GSEA of the three types of  $\beta$  cells with gene sets from MSigDB biological process ontology (GO) reveals various sub-processes of oxidative phosphorylation (OXPHOS) that are highly enriched in  $\beta$  cells isolated from eBCs compared to  $\beta$ -like cells of d20 clusters (left) or NECs (right). GSEA uses Kolmogorov–Smirnov statistics to calculate  $P$  values. **d**, Heat map of GSEA leading edge genes in electron transport chain ontology. Sample FPKM values are normalized across every gene ( $n = 2$  independent samples). **e**, Scatter plot of  $\log_{10}$  FPKM values for all coding transcripts present in  $\beta$  cells of eBCs and  $\beta$  cells isolated from adult human islets. Pearson correlation coefficient ( $r$ ) = 0.86,  $P < 2.2 \times 10^{-16}$ . **f**, Venn diagram of differentially expressed genes between human islet  $\beta$  cells and d20  $\beta$ -like cells, NEC  $\beta$  cells and eBC  $\beta$  cells. eBC  $\beta$  cells are more closely related to human islet  $\beta$  cells by 967 genes (compared to NEC  $\beta$  cells) or 843 genes (compared to d20  $\beta$ -like cells). DEG, differentially expressed genes.

kidney capsule of non-diabetic NSG mice. eBCs secreted C-peptide after an acute glucose challenge as early as three days post-transplant. This function was maintained even 30 days after transplant (Fig. 7b). The graft recovered three days following transplant mainly consisted of C-peptide<sup>+</sup> cells with some C-peptide<sup>+</sup>/glucagon<sup>+</sup> or C-peptide<sup>+</sup>/somatostatin<sup>+</sup> co-positive cells (Fig. 7a, top). The corresponding H&E image illustrates intact, islet-like endocrine clusters, characterized by pale cytoplasm, and infiltrated with blood cells. The graft at day 48 was composed of clusters of monohormonal C-peptide<sup>+</sup>, glucagon<sup>+</sup> and somatostatin<sup>+</sup> endocrine cells organized in a manner similar to human islets. Although the majority of cells in the graft were C-peptide<sup>+</sup>, we found a significant number of glucagon<sup>+</sup> and a few somatostatin<sup>+</sup> single-positive cells (Fig. 7a, bottom). H&E analysis at this stage revealed a more extensive network of blood cells throughout the grafts. To show that eBC  $\beta$  cells are more functional at a per-cell basis compared to NEC or d20 immature  $\beta$ -like cells, clusters containing

approximately equivalent numbers of insulin-producing cells were transplanted from each population. While NECs and d20s were marked by the clear absence of robust glucose-stimulated insulin secretion, eBCs were functional rapidly following transplantation similar to human islets (Supplementary Fig. 7). We performed a glucose tolerance test on mice transplanted with human islets, eBCs, NECs and d20 clusters 30 days after transplantation. eBC-transplanted mice were significantly more glucose tolerant and also cleared glucose faster than mice that received either d20 clusters or NECs (Fig. 7c; note the significantly reduced AUC). We then tested whether eBCs provide resistance against streptozotocin (STZ)-induced diabetes. Transplantation of eBCs ameliorates hyperglycaemia whereas STZ-treated control mice became highly diabetic (blood glucose levels of 600 mg dl<sup>-1</sup>) at day 13. Blood glucose levels in eBC-treated mice rise to 600 mg dl<sup>-1</sup> upon removal of the graft (Nx), demonstrating the regulation of glucose levels by the human stem-cell-derived cells (Fig. 7d).

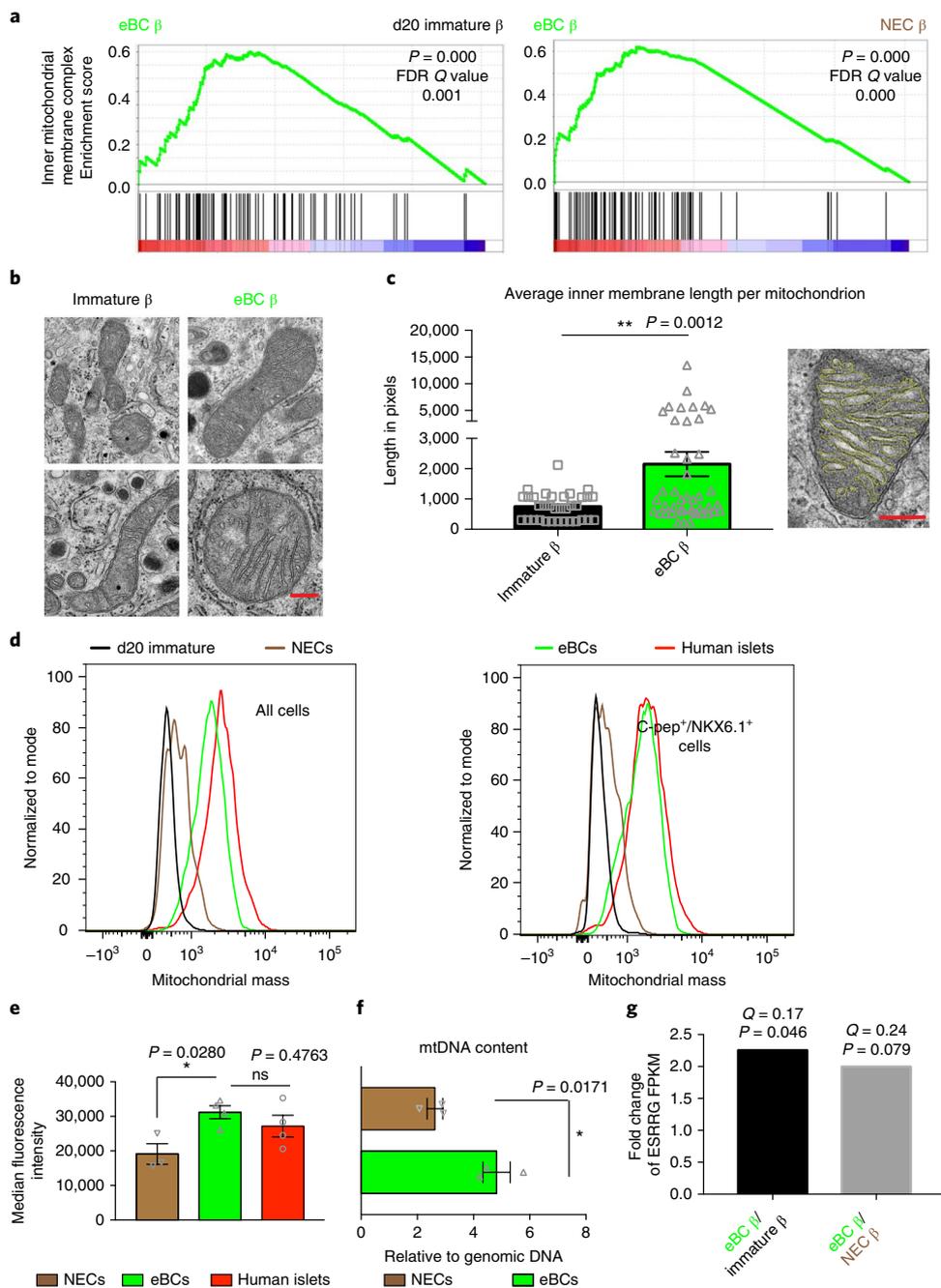


**Fig. 5 | eBCs possess functionally mature mitochondria.** **a–c**, Mitochondrial respiration assessed by a Cell Mito Stress Test using a Seahorse XFe24 Bioanalyzer. **a**, Oxygen consumption rate (OCR) was first measured under basal conditions (2.8 mM glucose) followed by sequential addition of either 2.8 mM (blue line) or 20 mM glucose (red line), 5  $\mu$ M oligomycin (oligo), 1  $\mu$ M FCCP and 5  $\mu$ M rotenone + antimycin A (rot/anti-M). eBCs show similar OCR profiles as human islets, whereas d20 immature clusters do not increase their OCR following glucose stimulation. d20 immature clusters: 2.8 mM glucose,  $n = 5$  independent samples; 20 mM glucose,  $n = 7$  independent samples. eBCs: 2.8 mM glucose,  $n = 5$  independent samples; 20 mM glucose,  $n = 6$  independent samples. Human islets: 2.8 mM glucose,  $n = 6$  independent samples; 20 mM glucose  $n = 7$  independent samples. Data are expressed as mean  $\pm$  s.e.m. **b**, Basal OCR of d20 immature clusters (black), eBCs (green) and human islets (red). Human islets,  $n = 7$  independent samples; eBCs,  $n = 5$  independent samples; d20 immature clusters,  $n = 4$  independent samples. Data are expressed as mean  $\pm$  s.e.m. \*\*\*\* $P = 0.0001$ .  $P$  values determined by one-way ANOVA with Dunnett test for multiple comparisons versus eBCs. ns, not significant. **c**, Dynamic mitochondrial energization as monitored by quenching of rhodamine-123 fluorescence. The rate of fluorescence decay/quenching of rhodamine-123 fluorescence is directly proportional to mitochondrial membrane potential. Representative experiment is shown. Mitochondrial membrane potential analyses were repeated with three distinct hESC  $\beta$  cell differentiations independently with similar results. Of note, the profile of mitochondrial potential changes on glucose stimulation is distinct between rodent and human  $\beta$  cells. **d**, Single-cell end-point analyses of mitochondrial potential in the indicated populations measured by flow cytometry after incubation with 2.8 mM glucose or 20 mM glucose and stained with the mitochondrial membrane potential indicator dye, MitoTracker Red CMXRos. Upper panel shows MitoTracker staining in C-peptide<sup>+</sup> cells and lower panel shows MitoTracker staining in C-peptide<sup>+</sup>/NKX6.1<sup>+</sup> double-positive cells. The experiment was repeated with three independent differentiations. Ratio of median fluorescence intensity at 20 mM and 2.8 mM glucose is shown (mean  $\pm$  s.e.m.). See Supplementary Table 6 for source data.

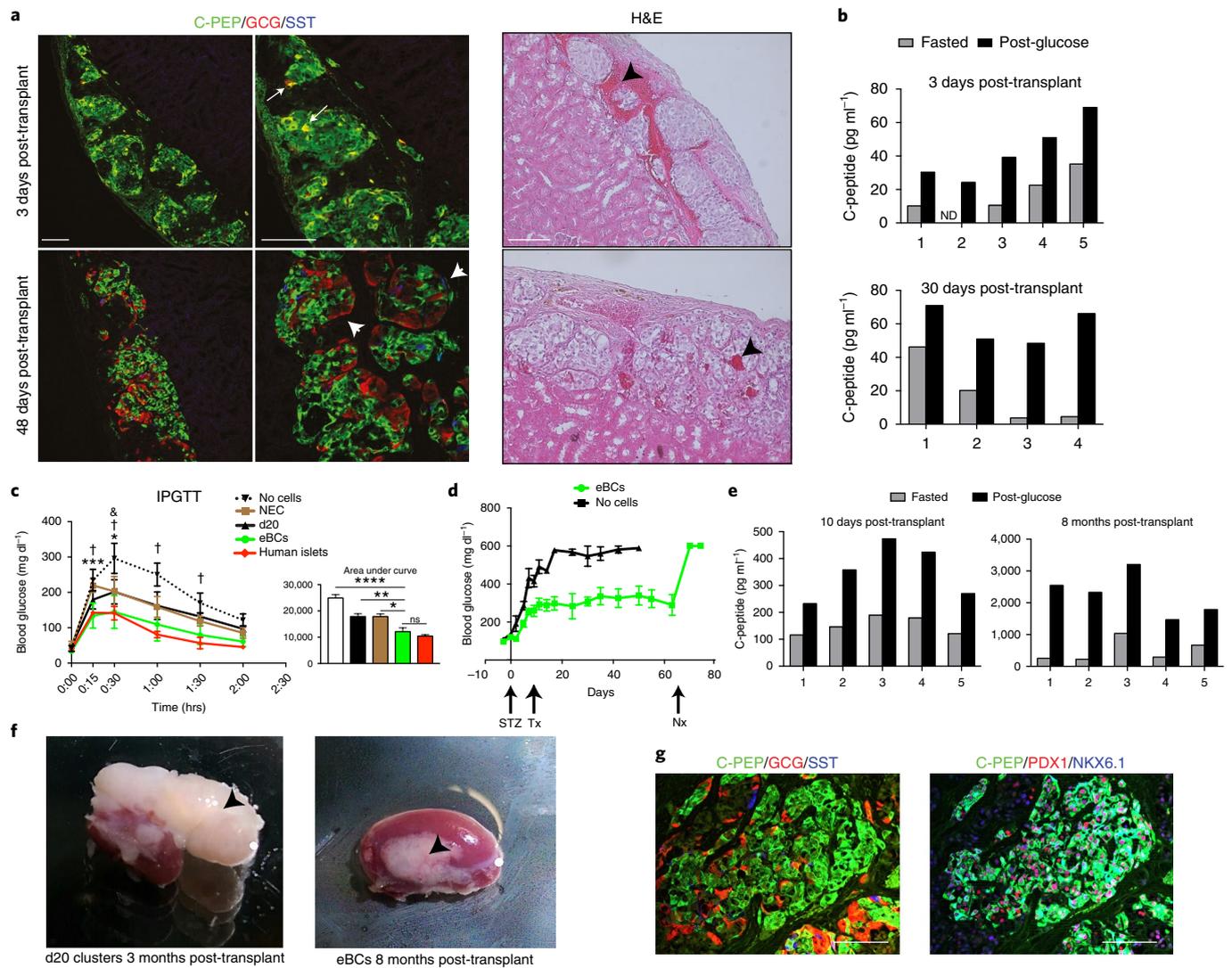
Finally we assessed the long-term efficacy and safety of eBC grafts. We transplanted 4,000 eBCs ( $\sim 4 \times 10^6$  cells) into non-diabetic NSG mice and followed their function over time. eBCs were still functional and secreted significantly higher levels of C-peptide 8 months later compared to the early period following transplant (Fig. 7e). Importantly, the graft appeared roughly the same size as the initial transplant even after 8 months in vivo (5/5 mice) (Fig. 7f). In contrast, grafts from mice that received equal number of d20 cells developed into large masses at 3 months post-transplant (7/10 mice) (Fig. 7f). The architecture of eBC grafts at 8 months was reminiscent of human islets with intermingled C-peptide<sup>+</sup> and glucagon<sup>+</sup> cells interspersed with few somatostatin<sup>+</sup> cells. The C-peptide<sup>+</sup> cells expressed both PDX1 and NKX6.1 (Fig. 7e). These data indicate that eBCs are immediately functional upon transplantation, and emphasize the efficacy and safety of eBCs on prolonged engraftment.

**Discussion**

While significant progress has been made towards in vitro generation of insulin-secreting cells from hPSCs<sup>2–4</sup>, these cells still appear immature ( $\beta$ -like) based on transcript and functional analyses, with features such as compromised dynamic GSIS, failure to terminate insulin secretion when glucose levels are lowered and impaired upstream calcium signalling. Here, we recapitulate aspects of islet formation in vivo, including clustering of newly formed  $\beta$  cells that had been omitted in previous differentiation protocols. Islet formation continues through postnatal development and this period is coincident with gradual functional  $\beta$  cell maturation, which broadly encompasses acquisition of glucose sensing, dense-core-granule biogenesis, stimulus-secretion coupling, and increased insulin output. Consistent with this, incorporating a reaggregation step permits maturation of in vitro generated hPSC-derived  $\beta$ -like cells into



**Fig. 6 | Endocrine cell clustering improves mitochondrial morphology and increases mitochondrial numbers.** **a**, GSEA traces demonstrating enrichment of the inner mitochondrial membrane complex signature in the cellular component ontology gene sets in  $INS^{GFP-high+}$  cells of eBC compared to those in d20 clusters (left) and NECs (right). GSEA uses Kolmogorov-Smirnov statistics to calculate  $P$  values. **b**, Representative transmission electron micrographs of mitochondria in d20 immature  $\beta$ -like cells and eBC  $\beta$  cells. The mitochondrial cristae are denser and more tightly folded in eBCs. Scale bar, 250 nm. Images are representative from two biologically independent samples. **c**, Inner mitochondrial membrane length was calculated using ImageJ software and represented as length in pixels in eBC  $\beta$  cells and d20 immature  $\beta$ -like cells. The image on the right shows an example mitochondrion in which inner mitochondrial membrane length was marked with the ImageJ tool for quantification (yellow lines). Scale bar, 250 nm. eBC  $\beta$  cells,  $n = 44$  mitochondria; immature  $\beta$  cells,  $n = 34$  mitochondria. Data are expressed as mean  $\pm$  s.e.m.  $**P < 0.01$  determined by two-tailed unpaired  $t$ -test with Welch's correction. **d**, Flow cytometry with Mitochondrial mass was used to assess mitochondrial mass in the indicated populations. Left: flow histogram of whole populations stained with Mitochondrial mass. Right: flow histogram of C-pep $^{+}$ /NKX6.1 $^{+}$  double-positive cells within each population stained with Mitochondrial mass. The experiment was repeated with three independent differentiations. **e**, Quantification of staining with Mitochondrial mass. eBCs,  $n = 4$  independent biological samples; NECs,  $n = 3$  independent biological samples; human islets,  $n = 4$  independent biological samples. Data are presented as mean fluorescence intensity  $\pm$  s.e.m.  $*P < 0.05$  determined by one-way ANOVA with Dunnett test for multiple comparisons versus eBCs. ns, not significant. **f**, Mitochondrial DNA (mtDNA) content assessed by the ratio of mtDNA/nuclear DNA by qPCR for the mitochondrial 16S rRNA gene and the nuclear  $\beta 2$  microglobulin gene.  $n = 3$  independent samples. Data are presented as mean  $\pm$  s.e.m.  $*P < 0.05$  determined by two-tailed unpaired  $t$ -test. See Supplementary Table 6 for source data. **g**, Levels of ERR $\gamma$  determined as FPKM in RNA-seq data from  $INS^{GFP-high+}$  from eBCs, NECs and d20 immature clusters. Data shown as fold change between eBC  $\beta$ /d20  $\beta$ -like and eBC  $\beta$ /NEC  $\beta$ .  $P$  and FDR-corrected  $Q$  values are indicated. Cuffdiff uses a theoretic metric derived from the Jensen-Shannon divergence to calculate  $P$  values<sup>41</sup>.



**Fig. 7 | eBCs are functional in vivo as early as 3 days after transplant, retain function long term and do not form tumours.** **a, b**, 700 eBCs ( $\sim 0.7 \times 10^6$  cells) were transplanted under the kidney capsule of non-diabetic male NSG mice. **a**, Grafts removed 3 days post-transplant (top) or 48 days post-transplant (bottom) were stained for C-peptide (C-PEP), glucagon (GCG) and somatostatin (SST) (left), or H&E (right). White arrows indicate GCG<sup>+</sup>/C-PEP<sup>+</sup> double-positive cells. White arrowheads mark monohormonal GCG<sup>+</sup> or SST<sup>+</sup> cells. Black arrowheads in the H&E images indicate blood cells. Scale bars, 100  $\mu$ m. Images are representative of five independent experiments. **b**, In vivo glucose challenge test at indicated days after transplant. Human C-peptide levels in the serum were measured after an overnight fast (grey) and again 30 min after an intraperitoneal glucose injection (black). The numbers on the x axes indicate individual animals. ND, not detected. **c**, The intraperitoneal glucose tolerance test (IPGTT) was performed 30 days after transplantation with eBCs (green,  $n = 5$  animals), NECs (brown,  $n = 4$  animals), d20 clusters (black,  $n = 5$  animals), human islets (red,  $n = 3$  animals) and control non-transplanted NSG mice (no cells, dotted line,  $n = 4$  animals). Statistics determined two-way repeated measures ANOVA with Sidak's multiple comparison tests. \* $P < 0.05$ , \*\*\* $P < 0.001$ , NEC versus eBCs; † $P < 0.001$ , no cells versus eBCs; ‡ $P < 0.05$ , d20 versus eBCs. Area under the curve was determined for each group. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\*\* $P < 0.0001$ , ns = not significant determined by one-way ANOVA with Dunnett test for multiple comparisons versus eBCs. Data presented as mean  $\pm$  s.e.m. **d**, Random fed glucose measurements taken from STZ-treated control NSG mice (no cells, black) or NSG mice transplanted with 6,000 eBCs (green).  $n = 3$  animals. One of the diabetic control non-transplanted mice expired on day 24. Data presented as mean  $\pm$  s.e.m. See Supplementary Table 6 for source data. **e**, 4,000 eBCs ( $\sim 4 \times 10^6$  cells) were transplanted under the kidney capsule of non-diabetic male NSG mice. Human C-peptide levels following fasting (grey) and 60 min after an intraperitoneal glucose bolus (black) were measured 10 days and 8 months after transplant. **f**, Images of kidneys from mice transplanted with  $4 \times 10^6$  cells each of d20 immature clusters 3 months post-transplant (left) or eBCs 8 months post-transplant (right). **g**, Immunofluorescent images of eBC grafts 8 months after transplant stained with C-peptide/glucagon/somatostatin (left) and C-peptide/PDX1/NKX6.1 (right). Images are representative of five independent experiments.

eBCs with enhanced functional properties, including (1) the stereotypical rapid and robust release of C-peptide synchronous with glucose concentration, (2) elevated calcium signalling on glucose stimulation that is ablated on stimulus removal, (3) highly sensitive K<sup>+</sup>-ATP channels that can be reversibly closed and opened, and (4) mitochondrial energization analogous to adult human islets. eBC

size (100  $\mu$ m) falls within the range of smaller human islets that have better function in vitro and improved outcome post transplantation<sup>32</sup>. eBCs function rapidly following transplantation, similarly to human islets, and generate islet-like grafts with intermingled  $\beta$ ,  $\alpha$  and  $\delta$  cells.  $\alpha$  cells can arise from insulin/glucagon double-positive progenitors during human embryonic development<sup>18,33</sup>.

Somatostatin-producing  $\delta$  cells also separate from a shared  $\beta/\delta$  cell lineage<sup>34</sup>. Thus, while sorting for GFP/insulin, we also capture a small population of cells that are dual-hormone-positive, the likely progenitors for the monohormonal glucagon- and somatostatin-positive cells that appear upon transplantation. eBC-grafts examined eight months post-transplant lack tumours that can arise from progenitors/uncommitted stem cells. The biggest long-term challenge of translating hPSC-derived cell types into therapeutic applications is to ensure efficient differentiation into a population of only the desired cell types and to guarantee absence of cells with tumorigenic potential. Our findings emphasize that enriched islet/ $\beta$  clusters can pose a safe and effective alternative to human islets for cell therapy.

Although the relevance of  $\beta$  cell contacts in human and mouse islet function has been reported<sup>35–37</sup>, the surprising outcome of clustering immature  $\beta$ -like cells in our study was the induction of mitochondrial metabolic maturation.  $\beta$  cells in eBCs were enriched in oxidative metabolic pathways, including OXPHOS, electron transport chain, TCA cycle and ATP biosynthesis, compared with  $\beta$  cells residing in progenitor-rich environments. We observed increased mitochondrial respiration, mitochondrial mass and differential membrane depolarization on glucose stimulation in  $\beta$  cells of eBCs. Mitochondrial cristae of these mature  $\beta$  cells were highly folded and stacked, a noteworthy observation given recent discoveries linking dynamics of cristae morphology to function of the OXPHOS system<sup>31</sup>. Our findings are in concert with the transition from glycolysis-centric to OXPHOS-centric metabolism posited during postnatal functional maturation<sup>5,8,25,26</sup>, which may in part be driven by clustering of  $\beta$  cells. Clustering resulted in enrichment of ERR $\gamma$ , a mitochondrial regulator of the metabolic transition, suggesting that reaggregation activates endogenous pathways driving increased mitochondrial activity. We suppose multiple pathways, in addition to ERR $\gamma$ , may modulate the changes observed upon reaggregation.

Correlation analysis from RNA-seq experiments unveiled that eBC  $\beta$  cells more closely resemble primary  $\beta$  cells ( $r=0.86$ ) than d20  $\beta$ -like cells or NEC  $\beta$  cells. Nonetheless, we found differences in expression of certain adult  $\beta$  cell markers, including *MAFA*, *UCN3* and *G6PC2*, between the two groups. These data are in line with an extended postnatal maturation period in humans lasting through childhood<sup>11</sup>, wherein ‘juvenile’  $\beta$  cells have high expression of *MAFB*<sup>10</sup> but not *MAFA* as observed in eBCs. These juvenile cells respond like adult islets in dynamic secretion assays except they release lower quantities of insulin<sup>10</sup>, findings similar to those we observed in eBCs. Unlike neonatal islets, infant islets (age: 1–3 years) secrete insulin in a manner similar to adult islets on stimulation with various secretagogues<sup>38–40</sup>, indicating that human islets are functionally mature by the age of one year. These results suggest that human  $\beta$  cells progress after birth to assume a ‘juvenile’ functional state, and the insulin secretory capacity of ‘juvenile’ human  $\beta$  cells continues to increase until puberty with enrichment of factors such as *MAFA*, *SIX2* and *SIX3*<sup>10,15</sup>. High expression of *MAFB* in young  $\beta$  cells suggest an indispensable role for this gene in attaining early function.

We demonstrate that clustering/reaggregation of immature  $\beta$ -like cells is a critical step in maturation and generation of fully functional  $\beta$  cells from hPSCs in vitro. Our data reveal that coalescence of immature  $\beta$ -like cells into endocrine-enriched niches induces metabolic maturation of mitochondria, thus enabling mature  $\beta$  cell function. We believe that our study delivers a cell type closely resembling human  $\beta$  cells, opening the opportunity to study  $\beta$  cell physiology and functional toxicology that was not previously possible with immature  $\beta$ -like cells.

### Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and

associated accession codes are available at <https://doi.org/10.1038/s41556-018-0271-4>.

Received: 28 June 2018; Accepted: 20 December 2018;  
Published online: 1 February 2019

### References

- Barton, F. B. et al. Improvement in outcomes of clinical islet transplantation: 1999–2010. *Diabetes Care* **35**, 1436–1445 (2012).
- Pagliuca, F. W. et al. Generation of functional human pancreatic  $\beta$  cells in vitro. *Cell* **159**, 428–439 (2014).
- Rezania, A. et al. Reversal of diabetes with insulin-producing cells derived in vitro from human pluripotent stem cells. *Nat. Biotech.* **32**, 1121–1133 (2014).
- Russ, H. A. et al. Controlled induction of human pancreatic progenitors produces functional  $\beta$ -like cells in vitro. *EMBO J.* **34**, 1759–1772 (2015).
- Yoshihara, E. et al. ERR $\gamma$  is required for the metabolic maturation of therapeutically functional glucose-responsive  $\beta$  cells. *Cell. Metab.* **23**, 622–634 (2016).
- Blum, B. et al. Functional  $\beta$ -cell maturation is marked by an increased glucose threshold and by expression of *urocortin 3*. *Nat. Biotech.* **30**, 261–264 (2012).
- Aguayo-Mazzucato, C. et al. *Mafa* expression enhances glucose-responsive insulin secretion in neonatal rat  $\beta$  cells. *Diabetologia* **54**, 583–593 (2011).
- Jermendy, A. et al. Rat neonatal  $\beta$  cells lack the specialised metabolic phenotype of mature  $\beta$  cells. *Diabetologia* **54**, 594–604 (2011).
- Dhawan, S. et al. DNA methylation directs functional maturation of pancreatic  $\beta$  cells. *J. Clin. Invest.* **125**, 2851–2860 (2015).
- Arda, H. E. et al. Age-dependent pancreatic gene regulation reveals mechanisms governing human  $\beta$  cell function. *Cell. Metab.* **23**, 909–920 (2016).
- Gregg, B. E. et al. Formation of a human  $\beta$ -cell population within pancreatic islets is set early in life. *J. Clin. Endocrinol. Metab.* **97**, 3197–3206 (2012).
- Gu, C. et al. Pancreatic  $\beta$  cells require neuro to achieve and maintain functional maturity. *Cell. Metab.* **11**, 298–310 (2010).
- Gosmain, Y. et al. Pax6 is crucial for  $\beta$ -cell function, insulin biosynthesis, and glucose-induced insulin secretion. *Mol. Endocrinol.* **26**, 696–709 (2012).
- Dai, C. et al. Islet-enriched gene expression and glucose-induced insulin secretion in human and mouse islets. *Diabetologia* **55**, 707–718 (2012).
- Scoville, D. W. et al. MLL3 and MLL4 methyltransferases bind to the *mafa* and *mafb* transcription factors to regulate islet  $\beta$ -cell function. *Diabetes* **64**, 3772–3783 (2015).
- Lemaire, K., Thorrez, L. & Schuit, F. Disallowed and allowed gene expression: two faces of mature islet beta cells. *Annu. Rev. Nutr.* **36**, 45–71 (2016).
- Nair, G. & Hebrok, M. Islet formation in mice and men: lessons for the generation of functional insulin-producing  $\beta$ -cells from human pluripotent stem cells. *Curr. Op. Genet. Dev.* **32**, 171–180 (2015).
- Jeon, J. et al. Endocrine cell clustering during human pancreas development. *J. Histochem. Cytochem.* **57**, 811–824 (2009).
- Rahier, J., Wallon, J. & Henquin, J. C. Cell populations in the endocrine pancreas of human neonates and infants. *Diabetologia* **20**, 540–546 (1981).
- Borden, P. et al. Sympathetic innervation during development is necessary for pancreatic islet architecture and functional maturation. *Cell Rep.* **4**, 287–301 (2013).
- Pisania, A. et al. Quantitative analysis of cell composition and purity of human pancreatic islet preparations. *Lab. Invest.* **90**, 1661–1675 (2010).
- Taylor, B. L., Liu, F.-F. & Sander, M. Nkx6.1 is essential for maintaining the functional state of pancreatic  $\beta$  cells. *Cell Rep.* **4**, 1262–1275 (2013).
- Du, A. et al. Islet-1 is required for the maturation, proliferation, and survival of the endocrine pancreas. *Diabetes* **58**, 2059–2069 (2009).
- Portela-Gomes, G. M. et al. The importance of chromogranin A in the development and function of endocrine pancreas. *Regul. Pept.* **151**, 19–25 (2008).
- Jacovetti, C. et al. Postnatal  $\beta$ -cell maturation is associated with islet-specific microRNA changes induced by nutrient shifts at weaning. *Nat. Commun.* **6**, 8084 (2015).
- Stolovich-Rain, M. et al. Weaning triggers a maturation step of pancreatic  $\beta$  cells. *Dev. Cell.* **32**, 535–545 (2015).
- Huang, C. et al. Synaptotagmin 4 regulates pancreatic  $\beta$  cell maturation by modulating the Ca<sup>2+</sup> sensitivity of insulin secretion vesicles. *Dev. Cell.* **45**, 347–361.e5 (2018).
- Qiu, W.-L. et al. Deciphering pancreatic islet  $\beta$  cell and  $\alpha$  cell maturation pathways and characteristic features at the single-cell level. *Cell. Metab.* **25**, 1194–1205 (2017).
- Hrvatin, S. et al. Differentiated human stem cells resemble fetal, not adult,  $\beta$  cells. *Proc. Natl Acad. Sci. USA* **111**, 3038–3043 (2014).
- Dorrell, C. et al. Transcriptomes of the major human pancreatic cell types. *Diabetologia* **54**, 2832 (2011).

31. Cogliati, S., Enriquez, J. A. & Scorrano, L. Mitochondrial cristae: where beauty meets functionality. *Trends. Biochem. Sci.* **41**, 261–2273 (2016).
32. Lehmann, R. et al. Superiority of small islets in human islet transplantation. *Diabetes* **56**, 594–603 (2007).
33. Riedel, M. J. et al. Immunohistochemical characterisation of cells co-producing insulin and glucagon in the developing human pancreas. *Diabetologia* **55**, 372–381 (2012).
34. Sosa-Pineda, B. et al. The *Pax4* gene is essential for differentiation of insulin-producing  $\beta$  cells in the mammalian pancreas. *Nature* **386**, 399 (1997).
35. Hauge-Evans, A. C. et al. Pancreatic beta-cell-to-beta-cell interactions are required for integrated responses to nutrient stimuli: enhanced  $\text{Ca}^{2+}$  and insulin secretory responses of MIN6 pseudoislets. *Diabetes* **48**, 1402 (1999).
36. Luther, M. J. et al. MIN6  $\beta$ -cell– $\beta$ -cell interactions influence insulin secretory responses to nutrients and non-nutrients. *Biochem. Biophys. Res. Commun.* **343**, 99–104 (2006).
37. Persaud, S. et al. Pseudoislets as primary islet replacements for research: report on a symposium at King's College London. *Islets* **2**, 236–239 (2010).
38. Henquin, J.-C. & Nenquin, M. Dynamics and regulation of insulin secretion in pancreatic islets from normal young children. *PLoS ONE* **11**, e0165961 (2016).
39. Hawdon, J. M. et al. The role of pancreatic insulin secretion in neonatal gluco-regulation. I. Healthy term and preterm infants. *Arch. Dis. Child* **68**, 274–279 (1993).
40. Kaye, R. et al. The response of blood glucose, ketones, and plasma nonesterified fatty acids to fasting and epinephrine injection in infants and children. *J. Pediatr.* **59**, 836–847 (1961).
41. Trapnell, C. et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* **7**, 562–578 (2012).

## Acknowledgements

We would like to thank the members of the Hebrok laboratory for advice and insightful comments during the preparation of the manuscript. G.G.N. was supported by a Kraft Family Fellowship and a JDRF postdoctoral fellowship (1-PNF-2016-320-S-B). Work in the laboratory of M.H. was supported by grants from the NIH (DK105831, DK108666). Imaging and flow cytometry experiments were supported by resources from the UCSF Diabetes and Endocrinology Research Center (DRC) and UCSF Flow Cytometry Core. Image acquisition was supported by the University of California, San Francisco Diabetes and Endocrinology Research Center (DRC) microscopy core P30 DK63720.

## Author contributions

Conception and study design: G.G.N. and M.H. Execution of experiments: G.G.N., J.S.L., H.A.R., S.T., M.S., R.C., C.J., M.L., Y.X., Y.W., V.Q.N. and G.S. Data analysis and interpretation: G.G.N., J.S.L., H.A.R., S.T., R.C., Y.X., Y.W., S.G., A.B. and M.H. Manuscript writing: G.G.N., J.S.L., S.P. and M.H. Final approval of manuscript: G.G.N., J.S.L., H.A.R., S.T., M.S., R.C., M.L., C.J., Y.X., Y.W., S.G., S.P., V.Q.N., G.S., J.O., A.B. and M.H.

## Competing interests

M.H. is affiliated with Semma Therapeutics (consultant, SAB member and stock holder) and Encellin Inc. (SAB member and stock holder). He holds stocks from Viacety Inc.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41556-018-0271-4>.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Correspondence and requests for materials** should be addressed to M.H.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019

## Methods

**Cell culture.** Mel1 INS<sup>GFP/W</sup> human embryonic stem cells were obtained from S. J. Micallef and E. G. Stanley (Monash Immunology and Stem Cell Laboratories, Australia). They were maintained and propagated on mouse embryonic fibroblasts (MEFs) in hESC media. The cells were passaged by enzymatic dissociation using TrypLE (Gibco). Mycoplasma contamination was regularly checked to ensure the integrity of the cells. To initiate differentiation, confluent hESC cultures were dissociated into single-cell suspensions using TrypLE, counted and seeded at  $5.5 \times 10^6$  cells per well in 6-well suspension plates in 5.5 ml hESC media supplemented with  $10 \text{ ng ml}^{-1}$  activin A (R&D Systems) and  $10 \text{ ng ml}^{-1}$  heregulinB (Peprotech). The plates were incubated at  $37^\circ\text{C}$  and 5%  $\text{CO}_2$  on an orbital shaker at 100 rpm to induce 3D sphere formation. After 24 hours, the spheres were collected in 50 ml falcon tubes, allowed to settle by gravity, washed once with PBS and resuspended in day 1 media. The resuspended spheres in day 1 media were distributed into fresh 6-well suspension plates with 5.5 ml of media per well. Thereafter, media was changed every day. Until day 3, spheres were fed by removing 5 ml of the media and replenishing with 5.5 ml. From days 4 to 20, 4.5 ml media was removed and 5 ml of fresh media was added. Media compositions are as follows: Day 1: RPMI (Gibco) containing 0.2% FBS, 1:5,000 ITS (Gibco),  $100 \text{ ng ml}^{-1}$  activin A and  $50 \text{ ng ml}^{-1}$  WNT3a (R&D Systems). Day 2: RPMI containing 0.2% FBS, 1:2,000 ITS and  $100 \text{ ng ml}^{-1}$  activin A. Day 3: RPMI containing 0.2% FBS, 1:1,000 ITS,  $2.5 \mu\text{M}$  TGF $\beta$ 1 IV (CalBioChem) and  $25 \text{ ng ml}^{-1}$  KGF (R&D Systems). Day 4–5: RPMI containing 0.4% FBS, 1:1,000 ITS and  $25 \text{ ng ml}^{-1}$  KGF. Day 6–7: DMEM (Gibco) with 25 mM glucose containing 1:100 B27 (Gibco) and 3 nM TTNPB (Sigma). Day 8: DMEM with 25 mM glucose containing 1:100 B27, 3 nM TTNPB and  $50 \text{ ng ml}^{-1}$  EGF (R&D Systems). Day 9–11: DMEM with 25 mM glucose containing 1:100 B27,  $50 \text{ ng ml}^{-1}$  EGF and  $50 \text{ ng ml}^{-1}$  KGF. Day 12–20: DMEM with 25 mM glucose containing 1:100 B27, 1:100 Glutamax (Gibco), 1:100 NEAA (Gibco),  $10 \mu\text{M}$  ALKi II (Axxora), 500 nM LDN-193189 (Stemgent),  $1 \mu\text{M}$  Xxi (Millipore),  $1 \mu\text{M}$  T3 (Sigma-Aldrich), 0.5 mM vitamin C, 1 mM N-acetyl cysteine (Sigma-Aldrich),  $10 \mu\text{M}$  zinc sulfate (Sigma-Aldrich) and  $10 \mu\text{g ml}^{-1}$  of heparin sulfate. Day 20: the spheres were collected, incubated with Accumax briefly and dissociated into single cells for flow cytometry. Live GFP<sup>high</sup> cells were sorted on Aria II at low flow rates and reaggregated in Aggrewell-400 (StemCell Technologies) at 1,000 cells per cluster in CMRL containing 1:100 B27 (or 10% FBS), 1:100 Glutamax (Gibco), 1:100 NEAA (Gibco),  $10 \mu\text{M}$  ALKi II (Axxora), 0.5 mM vitamin C,  $1 \mu\text{M}$  T3 (Sigma-Aldrich), 1 mM N-acetyl Cysteine (Sigma-Aldrich),  $10 \mu\text{M}$  zinc sulfate (Sigma-Aldrich) and  $10 \mu\text{g ml}^{-1}$  of heparin sulfate. Day 22–23: the reaggregated enhanced  $\beta$ -clusters (eBCs) were transferred from Aggrewells and placed on orbital shakers at 100 rpm, and further cultured for up to two weeks. Media was changed every third day following reaggregation. A step-by-step protocol describing the maintenance of undifferentiated hESCs, differentiation of hESCs to  $\beta$  cells and FACS sorting can be found in Protocol Exchange<sup>42</sup>.

Human islets used in the experiments were obtained from the UCSF Islet Core and Prodo laboratories. They were used for functional and RNA-seq analysis within 24–48 hours following isolation. Human islet preparations ranged in age from 32 to 62. Human islet donor information including age, BMI, HbA1c and gender is provided in Supplementary Table 4. The procurement and use of human islets used in the study was approved by the institutional biosafety committee at UCSF. The study is compliant with all relevant ethical regulations regarding research involving human participants, and informed consent was obtained by all participants at the isolation facility.

**Immunofluorescence staining.** Clusters were fixed with 4% PFA for 15 min at room temperature, washed with PBS, and stored at  $4^\circ\text{C}$  until processing for paraffin sectioning. Later, clusters were first embedded in 2% agar, followed by dehydration, paraffin embedding and sectioning at  $5 \mu\text{m}$  thickness. Sections were then stained according to a previously described protocol<sup>41</sup>. Primary antibodies are detailed in Supplementary Table 2. Secondary antibodies were conjugated AlexaFluor 488, 568 and 647 (Molecular probes) of appropriate species, and used at 1:500 dilution. Cells were counterstained with DAPI to mark nuclei. Slides were mounted with coverslips using Prolong Diamond Antifade reagent (Invitrogen). Images were generated using Leica SP5 confocal microscope or Zeiss apotome.

**Flow cytometry.** Clusters at indicated stages were dissociated, fixed, permeabilized and stained for various intracellular markers for analysis on LSRFortessa X20 Dual, as described previously<sup>4</sup>. Data were analysed with FlowJo software. Anti-glucagon and Anti-human C-peptide antibodies were conjugated in-house using the Molecular Probes Antibody Labeling Kits according to manufacturer's instructions. Antibody details are listed in Supplementary Table 2.

**Quantitative real-time PCR.** Human islets and hPSC-derived clusters were harvested at indicated stages of differentiation by homogenization in TRIzol (Invitrogen) or Buffer RLT (Qiagen). RNA was isolated and purified using RNeasy Mini/Micro kits (Qiagen). qPCR was performed using gene expression assays from Applied Biosystems. Assay IDs are given in Supplementary Table 3. The comparative threshold ( $\Delta\Delta\text{CT}$ ) method was used to quantify transcript abundance.

**Dynamic insulin secretion assay.** A perfusion system from Biorep technologies was used for dynamic secretion assays. 30–50 hPSC-derived clusters or human islets were placed on filters in plastic chambers that were maintained at  $37^\circ\text{C}$  in a temperature controlled chamber. Under temperature- and  $\text{CO}_2$ -controlled conditions, the clusters were perfused at  $100 \mu\text{l min}^{-1}$  with Krebs-Ringer buffer (KRB) using a peristaltic pump. After an initial 1–2 hour long preincubation in 2.8 mM glucose-KRB, alternating low (2.8 mM) and high (20 mM) glucose with/without  $10 \text{ nM}$  exendin-4 and 30 mM KCl were perfused through the system. Flow-through was collected over the course of the experiment, and C-peptide levels were measured using the STELLUX Chemi Human C-peptide ELISA kit (Alpco). After the experiment, clusters/islets were recovered from the chambers, and insulin and DNA content were measured by acid-ethanol extraction and Quant-iT PicoGreen dsDNA Assay Kit, respectively.

**Pro-insulin to C-peptide content analysis.** eBCs, NECs and d20 clusters were harvested and treated over night with acid-ethanol. The extract from each sample was neutralized with Tris-HCl, diluted and analysed with both human pro-insulin specific ELISA kit (STELLUX Chemi Human Total Proinsulin ELISA-Alpco) and human C-peptide specific ELISA kit (STELLUX Chemi Human C-peptide ELISA kit-Alpco).

**RNA-sequencing library preparation and data analysis.** Total RNA was isolated from INS<sup>GFP-high+</sup> cells FACS sorted from d27 eBCs, NECs and d20 immature clusters; and  $\beta$  cells sorted from adult human islets using RNeasy Micro kits (Qiagen). RNA isolated from sorted cells was further concentrated employing RNA Clean & Concentrator (Zymo research). Only RNA samples with RIN > 7 were used to generate libraries for deep sequencing. Total RNA libraries were prepared from eBC, NEC and d20 samples. rRNA was depleted by RiboGone (Clontech), and 325–350 bp sized strand-specific cDNA libraries were prepared using the SMARTer Stranded Total RNA Sample Prep Kit – Low Input Mammalian Kit (Clontech). mRNA libraries were prepared from sorted primary  $\beta$  cell samples using Kappa stranded mRNA-seq kit. All libraries were further sequenced on Illumina HiSeq instrument generating paired-end 100-base-pair reads. Read qualities were assessed by FASTQC tool on Galaxy. Reads were then mapped to the human genome assembly (hg38) using TopHat version 2.1.0<sup>44</sup> using the corresponding sample's mean inner distance between mate pairs, library-type fr-firststrand for strand determination leaving other parameters default. The tuxedo suite tools were used for further analysis: rRNA and tRNA reads were masked and only mRNA reads were quantified during transcript assembly using Cufflinks by using a mRNA gtf file for annotation. Differential expression analysis was performed using Cuffdiff. Seeking an unbiased approach to pathway analysis, we used Gene set enrichment analysis (GSEA) tool developed by Broad Institute<sup>43</sup> that identifies groups of coordinately regulated genes present in gene sets annotated in the Molecular Signatures Database (MSigDB). Only genes with counts > 1 FPKM in all replicates of any one condition were used for GSEA. The following parameters were used for the analysis: ranking metric: 't test', number of permutations: 1000, MSigDB collections: hallmark gene sets, curated gene sets such as KEGG, and GO gene sets such as biological process, molecular function and cellular component. The program R was used to plot scatter plots and calculate correlation coefficients. The online tool, Morpheus, from Broad Institute was used to generate heat maps and perform hierarchical clustering. Contaminating acinar/zymogen genes and histone genes that do not have polyadenylated mRNAs were filtered from the list of differentially expressed genes reported in Fig. 4f. We also used DESeq2, another bioinformatics pipeline, to confirm the RNA-seq results. The top up- and downregulated pathways on GSEA were similar from the results obtained through analysis via Tuxedo suite and DESeq2 (Supplementary Table 5).

**Bisulfite sequencing.** DNA samples were bisulfite converted and purified using the Epitect Plus kit (Qiagen). Bisulfite-treated DNA was used as template to amplify the regions of interest using the bisulfite-converted DNA primers previously described<sup>9</sup>. Annealing temperatures were  $64^\circ\text{C}$  for HK1 and  $61^\circ\text{C}$  for LDHA using KAPA HiFi Uracil + DNA polymerase and ReadyMix. PCR products were gel purified and used for TOPO-TA cloning (Invitrogen), followed by Sanger sequencing. Bisulfite sequencing data were aligned and analysed using the Sequencher software. In the resulting data, each line of the diagram represents one clone and 8–10 clones were analysed per sample. Filled circles represent methylated CpGs and open (white) circles represent unmethylated ones. Bisulfite-seq raw data for HK1 and LDHA loci are provided in Statistics Source Data (Supplementary Table 6).

**Electron microscopy.** Differentiated clusters were spun down and excess media was removed, followed by addition of ice-cold fixative (0.1 M sodium cacodylate solution containing 2% paraformaldehyde and 2.5% glutaraldehyde, 3M  $\text{CaCl}_2$ , final pH 7.4) and incubation on ice for 20 min. The clusters were further processed using standard transmission electron microscopy procedures and analysed by the Gladstone Electron Microscopy Core. High-resolution electron microscopy images were used to quantify insulin granules and length of the inner mitochondrial membrane. Mature insulin granules were identified with a characteristic electron dense core surrounded by a halo and immature granules as having a

non-compacted core within the vesicular membrane. The mitochondrial inner membrane length was calculated by first outlining the inner membranes in ImageJ software followed by using a perimeter quantification tool.

**In vitro metabolic flux analysis.** Mitochondrial oxygen consumption rate (OCR) and extracellular acidification rate (ECAR) were measured in real time using XFe24 extracellular flux analyser (Seahorse Bioscience). hPSC-derived clusters and human islets were first rinsed with pH adjusted (7.4) XF base media (sodium bicarbonate free) supplemented with 3 mM glucose. 20–30 clusters/islets were placed per well of an islet plate and each experiment had at least two replicates per group. This was followed by insertion of a mesh to prevent movement of clusters during the assay and incubation in a non-CO<sub>2</sub> incubator at 37°C for at least one hour. Three baseline measurements were taken, following which glucose (2.8 or 20 mM), oligomycin (5 µM), carbonyl cyanide-4- (trifluoromethoxy) phenylhydrazone (FCCP, 1 µM), and rotenone and antimycin A (5 µM) were injected sequentially. OCR and ECAR were measured at 37°C in real time throughout the assay period. OCR was normalized to average baseline measurement and expressed as percent change during the course of the experiment. Insulin and DNA content of each well were also determined by STELLUX Chemi Human C-peptide ELISA kit (Alpco) and Quant-iT PicoGreen dsDNA Assay Kit, respectively.

**Islet perfusion and imaging for calcium flux and mitochondrial depolarization analysis.** All imaging and perfusion experiments were conducted according to previously described methods<sup>44</sup>. In brief, hPSC-derived clusters and islets were incubated in 2 ml of Krebs buffer containing both 5 µM Fura-2/AM and 2.5 µM Rhodamine 123 fluorescent dyes (Molecular Probes, CA) for 30 minutes prior to loading the device. The samples were then loaded into the temperature equilibrated microfluidic device mounted on an inverted epifluorescence microscope (Leica DMI 4000B, location). Dual-wavelength Fura-2/AM was excited at 340 and 380 nm and fluorescence emission was detected at 510 nm. Intracellular Ca<sup>2+</sup> was expressed as a ratio of fluorescent emission intensity F340 / F380 (%). Rh123 was excited at 490 ± 10 nm, and emission was measured at 530 ± 10 nm. Fura-2 and Rh123 fluorescence emission spectra were filtered using a Fura-2/FITC polychroic beamsplitter and double band emission filter (Chroma Technology, Part number: 73100bs). These images were collected with a CCD (Retiga-SRV, Fast 1394, QImaging). SimplePCI software (Hamamatsu Corp. location) was used for image acquisition and analysis. Both fluorescence signals were expressed as 'change in percentage' after being normalized against basal intensity levels established before stimulation. 5µM Oligomycin and 1µM FCCP were used when indicated.

**Mitochondria experiments.** For estimation of mitochondrial mass, dissociated cell populations were incubated in warm media containing MitoID at a dilution of 1:10,000 for one hour at 37°C with gentle shaking. Following washes with PBS, they were fixed with 4%PFA, stained with anti-C-peptide and anti-NKX6.1 antibodies, and analysed on LSRFortessa X20 Dual (BD Biosciences). For estimation of mitochondrial membrane potential, MitoTracker Red CM-H<sub>2</sub>XRos (Thermo Fisher Scientific), a mitotracker dye that fluoresces only upon oxidation in live cells, was used. Dissociated cell populations were incubated with 500 nm MitoTracker Red CM-H<sub>2</sub>XRos in KRB containing either 2.8 or 20 mM glucose for one hour at 37°C with gentle shaking. Cells were fixed and further stained with antibodies against C-peptide and NKX6.1 before FACS analysis. For measuring mitochondrial DNA copy number, DNA was isolated from indicated cell populations using QIAamp DNA Micro Kit (Qiagen). qPCR was used to determine the ratio of mitochondrial mtDNA 16S rRNA gene to nuclear β-2-microglobulin (β2M) gene using SYBR Green (Roche) based detection. Human specific primers used for mtDNA 16S rRNA were: 5'-GCCTTCCCCCGTAAATGATA-3' and 5'-TTATGCGATTACCGGGCTCT-3'; and for β2M were: 5'-TGCTGTCTCCATGTTGATGTATCT-3' and 5'TCTCTGCTCCCCACCTCTAAGT-3'.

**Mouse studies.** NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ mice (NSG) were obtained from Jackson Laboratories and bred in our facility. Male mice between the age group of 10–20 weeks were used in this study and were maintained according to protocols approved by the University of California, San Francisco, Institutional Animal Care and Use Committee. This study is compliant with all relevant ethical regulations regarding animal research. Mice were anaesthetized with isoflurane and transplanted with cells under the kidney capsule. One cohort of mice was

transplanted with 700 eBCs (~700,000 cells), a second cohort received 4,000 eBCs (~4 × 10<sup>6</sup> cells), a third cohort received 6,000 eBCs (~6 × 10<sup>6</sup> cells) and a fourth received ~2,000 eBCs (2 × 10<sup>6</sup> cells) or 4 × 10<sup>6</sup> cells each of d20 clusters or NECs or 500 IEQ human islets. For in vivo glucose challenge experiments, transplanted mice were fasted overnight, and serum before and 30 minutes following an intraperitoneal injection of 2 g kg<sup>-1</sup> D-glucose solution was collected by sub-mandibular bleeding. Kidneys bearing grafts were removed at indicated time points for immunofluorescent and H&E staining. To determine the ability of the grafts to protect against diabetes, mice were injected with multiple low doses of STZ (35 mg kg<sup>-1</sup> d<sup>-1</sup> for 5 days) followed by transplantation of half the cohort with eBCs, and were monitored for hyperglycaemia. One control untransplanted mouse died on day 24. A survival nephrectomy was performed to remove the grafts on day 68. Random fed blood glucose levels of the mice were monitored throughout the course of the experiment. No statistical method was employed to determine samples size, analysis was not blinded and mice were not randomized.

**Statistics and reproducibility.** Statistical tests performed for specific data sets are described in the corresponding figure legends. In brief, under the assumption of normal distribution, two-tailed unpaired *t*-tests (Student's *t*-test) were used if standard deviation (SD) was equal or two-tailed unpaired *t*-tests with Welch's correction were used if SD was unequal for pairwise comparison in Figs. 3e,f,h and 6c,f and Supplementary Figs. 1b, 6 and 7. One-way ANOVA with Tukey test for multiple comparisons was used to calculate significance in Supplementary Fig. 3. One-way ANOVA with Dunnett test for multiple comparisons was used to calculate significance in Figs. 5b and 6e and Supplementary Fig. 3 and for area under the curve in Fig. 7d. Two-way repeated measures ANOVA with Sidak's multiple comparison tests were used to determine *P* values in IPGTT in Fig. 7d. Significance in Fig. 1d was calculated using the Holm-Sidak method for multiple *t*-tests. *P* values were as follows: C-pep<sup>+</sup>/NKX6.1<sup>+</sup>, 4.45683 × 10<sup>-3</sup>; C-pep<sup>+</sup>/NKX6.1<sup>+</sup>, 8.41475 × 10<sup>-9</sup>; C-pep<sup>+</sup>/GCG<sup>+</sup>, 1.04554 × 10<sup>-8</sup>; C-pep<sup>+</sup>/NEUROD1<sup>+</sup>, 3.79435 × 10<sup>-8</sup>; PDX1<sup>+</sup>/NKX6.1<sup>+</sup>/C-pep<sup>+</sup>, 7.28718 × 10<sup>-7</sup>; C-pep<sup>+</sup>/CHGA<sup>+</sup>, 5.051 × 10<sup>-6</sup>; C-pep<sup>+</sup>/PDX1<sup>+</sup>, 6.52052 × 10<sup>-6</sup>; C-pep<sup>-</sup>/PDX1<sup>+</sup>, 1.07891 × 10<sup>-5</sup>; C-pep<sup>+</sup>/ISL1<sup>+</sup>, 0.001036039; C-pep<sup>+</sup>/GCG<sup>+</sup>, 0.001060223; C-pep<sup>+</sup>/NKX2.2<sup>+</sup>, 0.002352639; C-pep<sup>+</sup>/PAX6<sup>+</sup>, 0.002352639. All statistical tests were performed in GraphPad Prism Software v7. Statistical methods were not used to determine sample size. Experiments in the study were repeated independently at least three times except experiments in Figs. 3d, 4d, 6b and Supplementary Figs. 3c,d, 5 and 6d were repeated twice independently and the experiment in Supplementary Fig. 6b was performed once. Our collaborators were blinded to the treatment groups in calcium signalling, mitochondrial energization and DNA methylation experiments.

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

Materials and protocols that were used in this study are available from the corresponding author upon request. RNA-seq data that support the findings of this study have been deposited in the Gene Expression Omnibus (GEO) under accession code GSE10979. Source data for Figs. 1–3 and 5–7 and Supplementary Figs. 1, 3, 4 and 6 have been provided in Supplementary Table 6. All other data supporting the findings of this study are available from the corresponding author on reasonable request.

## References

- Nair, G. G., Holger, R. & Hebrok, M. Rapid generation of mature functional pancreatic islet-beta cells from human pluripotent stem cells. *Prot. Exchange* <https://doi.org/10.1038/protex.2018.140> (2019).
- Subramanian, A. et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences* **102**, 15545–15550 (2005).
- Adewola, A. F. et al. Microfluidic perfusion and imaging device for multi-parametric islet function assessment. *Biomed. Microdevices*. **12**, 409–417 (2010).

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated
- Clearly defined error bars  
*State explicitly what error bars represent (e.g. SD, SE, CI)*

*Our web collection on [statistics for biologists](#) may be useful.*

### Software and code

Policy information about [availability of computer code](#)

Data collection

SimplePCI software (Hamamatsu Corp. location) was used for image acquisition and analysis for the islet perfusion and calcium/Rhodamine-123 imaging experiments

Data analysis

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

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

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

## Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/authors/policies/ReportingSummary-flat.pdf](https://nature.com/authors/policies/ReportingSummary-flat.pdf)

## Life sciences study design

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

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

## Reporting for specific materials, systems and methods

### Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Unique biological materials
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants

### Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Unique biological materials

Policy information about [availability of materials](#)

Obtaining unique materials      The human ES-cell line, Me11 INSGFP/W used in the study is available from Prof. Ed Stanley at Murdoch Children's Research Institute. We are happy to provide an aliquot of the INS-GFP cells cultured in our

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

## Antibodies

### Antibodies used

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

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

### Validation

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

## Eukaryotic cell lines

### Policy information about [cell lines](#)

#### Cell line source(s)

Me11 INSGFP/W human embryonic stem cells were used in the study and were obtained from S. J. Micallef and E. G. Stanley

#### Authentication

Cell line authentication was not done.

#### Mycoplasma contamination

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

#### Commonly misidentified lines (See [ICLAC](#) register)

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

## Animals and other organisms

### Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

#### Laboratory animals

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

#### Wild animals

The study did not involve wild animals

#### Field-collected samples

Field samples were not collected for the study

## Human research participants

Policy information about [studies involving human research participants](#)

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

## Flow Cytometry

### Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

### Methodology

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

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