Cytotherapy 000 (2020) 1–9

Contents lists available at ScienceDirect

CYTOTHERAPY

journal homepage: www.isct-cytotherapy.org

Full-length article

Transcriptional dynamics of induced pluripotent stem cell differentiation into β cells reveals full endodermal commitment and homology with human islets

Silvia Pellegrini[1,**], Raniero Chimienti[1,**], Giulia Maria Scotti[2], Francesca Giannese[2], Dejan Lazarevic[2], Fabio Manenti[1], Gaia Poggi[1], Marta Tiffany Lombardo[1], Alessandro Cospito[1], Rita Nano[1], Lorenzo Piemonti[1,3], Valeria Sordi[1,*]

[1] *Diabetes Research Institute, IRCCS San Raffaele Hospital, Milan, Italy*
[2] *Centre of Omics Sciences, IRCCS San Raffaele Hospital, Milan, Italy*
[3] *Vita-Salute San Raffaele University, Milan, Italy*

ARTICLE INFO

Article History:
Received 24 July 2020
Accepted 27 October 2020
Available online xxx

Key Words:
β-cell differentiation
endocrine cells
induced pluripotent stem cells
pseudotime
single-cell RNA sequencing

ABSTRACT

Background aims: Induced pluripotent stem cells (iPSCs) have the capacity to generate β cells *in vitro*, but the differentiation is incomplete and generates a variable percentage of off-target cells. Single-cell RNA sequencing offers the possibility of characterizing the transcriptional dynamics throughout differentiation and determining the identity of the final differentiation product.
Methods: Single-cell transcriptomics data were obtained from four stages across differentiation of iPSCs into β cells and from human donor islets.
Results: Clustering analysis revealed that iPSCs undertake a full endoderm commitment, and the obtained endocrine pancreatic cells have high homology with mature islets. The iPSC-derived β cells were devoid of pluripotent residual cells, and the differentiation was pancreas-specific, as it did not generate ectodermal or mesodermal cells. Pseudotime trajectory identified a dichotomic endocrine/non-endocrine cell fate and distinct subgroups in the endocrine branch.
Conclusions: Future efforts to produce β cells from iPSCs must aim not only to improve the resulting endocrine cell but also to avoid differentiation into non-pancreatic endoderm cells.

Introduction

Human pluripotent stem cells (PSCs) have the capacity to generate β cells *in vitro* with different degrees of maturation [1,2]. Protocols consistently yield insulin-positive cells and, although improving over time, are not able to generate a pure population of mature β cells, resulting in substantial and uncharacterized heterogeneity within the final cell population, which includes a variable percentage (40–80%) of off-target cells [3]. The comprehensive characterization of the transcriptional dynamics throughout differentiation is a crucial step toward dissecting molecular regulators and mechanisms governing commitment and maturation of PSCs into fully differentiated cells. Additionally, defining the temporal changes in pluripotent and endocrine (EN) pancreas-related genes may provide indications about the safety and functional efficacy of these cells. Single-cell transcriptomics can generate new insights into cellular variability and subpopulation structures and also offers the possibility of inferring temporal gene expression kinetics by ordering individual cells by their pseudotimes based on their transcriptomic profiles, followed by identifying genes that display differential expression over time [4–6].

The single-cell perspective on PSCs has already helped to understand gene regulation and regulatory networks during exit from pluripotency and cell fate determination [7]. Studies of cardiac and liver differentiation have demonstrated that longitudinal application of single-cell RNA sequencing (scRNA-seq) permits tracking of transcriptional changes throughout differentiation at single-cell resolution [8–10]. In the past 5 years, over a dozen publications have reported results from the application of scRNA-seq to pancreatic islets [11–17], including islets from patients with type 2 diabetes and children [18–22]. A pseudotime reconstruction framework has also been applied to study the maturation process of β and α cells in the mouse pancreas [15,16,22]. An important outcome of these studies is the definition of gene signatures for all EN cell types. In addition,

* Correspondence: Valeria Sordi, PhD, Diabetes Research Institute, IRCCS San Raffaele Hospital, Via Olgettina 60, 20132, Milan, Italy.
E-mail address: sordi.valeria@hsr.it (V. Sordi).
** These authors contributed equally to this work.

https://doi.org/10.1016/j.jcyt.2020.10.004

cellular states along developmental processes, such as differentiation of stem cells into pancreatic β cells, can be profiled with transcriptomics data from individual EN cells. To date, four studies have sequenced PSC-derived populations across various stages of β-cell differentiation [3,23–25].

Given the intrinsic high variability of PSC lines and the different protocols, further studies that focus the analysis on the different stages of differentiation and dissect transcriptional dynamics, keeping the human islet as a reference point, are needed. In this study, the authors sought to use scRNA-seq to dissect the cellular heterogeneity and transcriptional dynamics of β-cell differentiation from human induced PSCs (iPSCs) and compare them with human islets from cadaveric donors. Single-cell transcriptomics data were obtained from cells at four time points across the differentiation protocol. Cell-to-cell variability and sample heterogeneity were characterized at each time point using a discrete cluster indicator for cell type and state and a continuous latent variable approach based on inference of differentiation trajectories.

Methods

The iPSC cell line and human islets

The human iPSC cell line CGTRCiB10 was obtained from Cell and Gene Therapy Catapult, London, UK. Cells were maintained in Essential 8 (Gibco) on 0.5 μg/cm^2 vitronectin recombinant human protein (Gibco), dissociated with 0.5 mM ethylenediaminetetraacetic acid (Ambion) for passages.

Human pancreatic islet (HI) preparations were isolated from heart-beating cadaveric organ donors as previously described [26] in the Pancreatic Islet Processing Unit of the Diabetes Research Institute at IRCCS San Raffaele Hospital in Milan, Italy. The use of human specimens (islet preparations discarded from clinical use) was approved by the institutional review board under the European Consortium for Islet Transplantation human islet distribution program supported by the Juvenile Diabetes Research Foundation (2-RSC-2019-724-I-X). Islets were hand-picked and their purity assessed as percentages of EN clusters positive to dithizone staining (>90%).

Differentiation of iPSCs into pancreatic cells

Human iPSCs were differentiated into insulin-producing cells following a seven-stage protocol established for PSCs [27,28], adding 2 μM iBET-151 (Selleckchem) at day 9 to day 13 [29], with a final step of maturation from day 18 to day 25 using CMRL 1066 (Mediatech), 10% fetal bovine serum (Lonza), 1% penicillin/streptomycin (Lonza), 1% L-glutamine (Lonza), 10 μM ALK5 inhibitor II, 1 μM L-3,3',5-triiodothyronine and 10 mM nicotinamide (Sigma). Details of the protocol are reported in the supplementary materials.

Gene expression analysis

Total RNA of cells at four different stages—undifferentiated iPSCs (day 0), pancreatic endoderm (PE) cells (day 14, stage five), EN cells (day 18, stage six), iPSC-derived β (iβ) cells (day 24, stage seven)—was extracted with the mirVana isolation kit (Ambion) and quantified by spectrophotometry with Epoch (BioTek) and Gen5 software (BioTek). After Turbo DNAse (Invitrogen) treatment, 2 μg of RNA was reverse transcribed with SuperScript IV reverse transcriptase (Invitrogen). Predesigned TaqMan gene expression assays (Applied Biosystems) (see supplementary Table 1) were used for gene expression study in a 7900HT real-time polymerase chain reaction system (Applied Biosystems). Gene expression levels were normalized using glyceraldehyde 3-phosphate dehydrogenase ($2^{-\Delta Ct}$ method) and are reported as fold change over undifferentiated iPSCs. Non-parametric test (Kruskal-Wallis with Dunn's multiple comparison test) was used to compare groups, and a two-tailed P value less than 0.05 or 0.01 was considered significant. Analysis of data was performed using Prism 5 software (GraphPad).

Cytofluorimetric analysis

Differentiated iPSCs were stained using Live/Dead stain (Invitrogen) to exclude dead cells from the analysis. Cells were permeabilized (Cytofix/Phosflow perm buffer III; BD Biosciences) and stained using the monoclonal antibodies 40/Oct-3 Alexa Fluor 647 anti-Oct3/4 (BD Biosciences), 658A5 Alexa Fluor 488 anti-PDX-1 (BD Biosciences), R11-560 PE anti-NKX6.1 (BD Biosciences) and T56-706 Alexa Fluor 647 anti-insulin (BD Biosciences). Analysis was carried out on a FACSCanto flow cytometer (BD Biosciences) and analyzed with FCS Express 6 Flow Research Edition (De Novo Software).

Immunofluorescence

For immunofluorescence, differentiated iPSCs were embedded in agarose, fixed with 4% paraformaldehyde, embedded in paraffin and cut into 3-μm sections. Cells were permeabilized for intracellular staining (2% bovine serum albumin [BSA], 5% fetal bovine serum, 0.2% Triton X-100 in phosphate-buffered saline) and stained using the antibodies listed in supplementary Table 2A,B. Images were acquired using a confocal UltraVIEW ERS microscope (PerkinElmer Life Sciences) and deconvolved with Huygens Professional 17.04 (Scientific Volume Imaging).

Hormone secretion

Hormone levels in iPSC culture supernatants were measured using the Bio-Plex Pro human diabetes kit (Bio-Rad). Samples were assayed according to the manufacturer's instructions. The plates were read on a Luminex xMAP instrument (Bio-Rad) and analyzed with Bio-Plex Manager 6.0 software (Bio-Rad).

Dynamic insulin secretion

A high-capacity automated perifusion system (BioRep Perifusion V2.0.0) was used to dynamically stimulate cell secretion. A low-pulsatility peristaltic pump was used to push 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid-buffered solution (125 mM sodium chloride, 5.9 mM potassium chloride [KCl], 2.56 mM calcium chloride, 1 mM magnesium chloride, 25 mM 4-[2-hydroxyethyl]-1-piperazineethanesulfonic acid, 0.1% BSA, pH 7.4) through a sample container harboring 50 iPSC clusters in Bio-Gel P-4 gel (Bio-Rad). Cells were stabilized with a slow flow perifusion rate (30 μL/min) with low-glucose solution (2 mM) for 60 min, followed by 20 min at 100 μL/min. A combined stimulus (11 mM glucose with or without 50 μM 3-isobutyl-1-methylxanthine) was then added for 20 min. Following a third step with 2 mM glucose for 20 min, cells were exposed to 30 mM KCl for 20 min and then to 2 mM glucose for another 20 min. The perifusates were collected every minute and stored at −20°C until insulin dosage by enzyme-linked immunosorbent assay kit (Mercodia).

Single-cell RNA sequencing

For scRNA-seq, the authors exploited the in-house Drop-seq platform, which combines microfluidic technology with nanoliter-sized droplets and allows the analysis of thousands of individual cells for parallel analysis [30]. Briefly, iPSCs at different time points and pancreatic islets were digested to single cells with trypsin-ethylenediaminetetraacetic acid, flushed through 0.40-μm filters and resuspended at 2×10^5 cells/mL in phosphate-buffered saline and 0.01% BSA. Single cells were encapsulated in oil droplets (Bio-Rad)

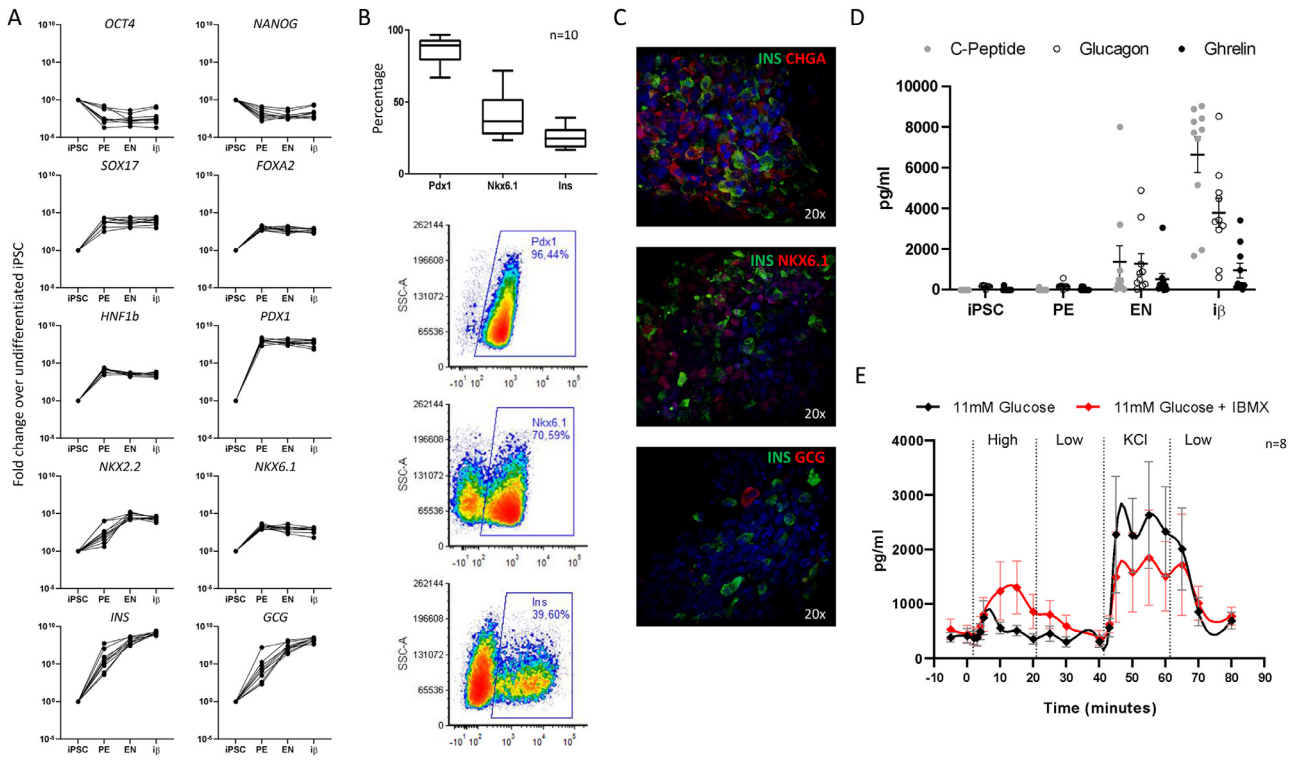

Figure 1. Characterization of iPSC-derived insulin-producing cells. (A) Gene expression analysis by TaqMan of markers of pluripotency (Oct-4 and NANOG), definitive EN (FOXA2 and SOX17), posterior foregut (HNF1b and PDX1), PE (NKX2.2 and NKX6.1) and EN cells (INS and GCG) in undifferentiated iPSC, PE, EN and terminally differentiated EN cells (iβ). Results are expressed as fold change over iPSCs. Each black line represents a single experiment (n = 10). (B) Top: box plot representing the percentage of Pdx1, Nkx6.1 and insulin-positive cells in terminally differentiated iβ cells measured by flow cytometry (n = 10). Bottom: dot plots of the percentage of Pdx1, Nkx.61 and insulin-positive cells of a representative experiment are reported. Gate delimitates positive events. (C) Protein expression analysis of iPSC-derived insulin-producing cells. Immunofluorescence of insulin (green) in iβ cells in co-staining with chromogranin A, NKX6.1 and glucagon (red). Nuclei stained in blue (DAPI). Magnification ×20. (D) Secretory capacity of iPSC-derived insulin-producing cells. Unstimulated levels of C-peptide, glucagon and ghrelin in pg/mL of the supernatant of iPSCs at different stages (PE, EN and iβ) of differentiation into insulin-producing cells (n = 10), measured by Luminex. (E) Dynamic insulin secretion by terminally differentiated iPSCs (n = 8) upon sequential stimulation with 0.5 mM glucose (basal), 11 mM glucose (stimulus) with (red line) or without (black line) IBMX, 0.5 mM glucose and 30 mM KCl. Dashed lines indicate the beginning of a stimulus. Insulin levels, measured by ELISA, are expressed in pg/mL (mean ± SEM). DAPI, 4′,6-diamidino-2-phenylindole; ELISA, enzyme-linked immunosorbent assay; IBMX, 3-isobutyl-1-methylxanthine; SSC.iβ, side scatter.

with lysis buffer and barcoded primer beads (ChemGenes Corporation). After collection, droplets were broken and beads subjected to reverse transcription, followed by library preparation using Nextera XT (Illumina) to obtain 2000 cells per experiment. Libraries were sequenced on the Illumina platform, with a coverage of 100,000 reads for each cell.

Single-cell RNA sequencing data analysis

Raw reads were processed using Macosko-Nemesh [30] (https://github.com/broadinstitute/Drop-seq/releases) and UMI-tools (https://github.com/CGATOxford/UMI-tools) pipelines. Alignment of the reads to the reference human genome alignment step was based on STAR 2.5.3a. The reference used was the human genome hg38, annotated according to Gencode 27. Gene–cell count matrices were imported in R environment 3.4.1 and analyzed with Seurat 2.3.1 [31]. Single-cell pseudotime trajectory was constructed with Monocle 2.5.4. Details of the analysis are reported in the supplementary materials.

Data availability

All raw and processed sequencing data generated in this study have been submitted to the National Center for Biotechnology Information Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/) under accession number GSE149613.

Results

Induced PSCs differentiate into functional insulin-secreting cells

Human iPSCs were differentiated into pancreatic β cells using a slightly modified version of the authors' recently published protocol [28]. The analysis of gene expression on bulk cells (Figure 1A) revealed that pluripotency genes *OCT4* and *NANOG* were downregulated, whereas genes of pancreatic development, such as *SOX17, FOXA2, HNF1b, PDX1, NKX2.2* and *NKX6.1*, were significantly upregulated during differentiation. Moreover, pancreatic hormones *INS* and *GCG* became highly expressed in iβ cells compared with iPSCs. Protein expression of PDX1, NKX6.1 and insulin was analyzed at the final stages of differentiation. The iβ cells were positive by flow cytometry for PDX1 (mean ± standard error of the mean [SEM], 85.6 ± 3.1%, n = 10), NKX6.1 (mean ± SEM, 41.1 ± 5.5%, n = 10) and insulin (mean ± SEM, 25.6 ± 2.5%, n = 10) (Figure 1B). The expression of β-cell markers in iPSC-derived insulin-positive cells was also confirmed by immunofluorescence. Insulin-positive iβ cells co-expressed the β-cell-specific transcription factor NKX6.1 and the neuroendocrine marker chromogranin A, and few cells were glucagon-positive (Figure 1C; also see supplementary Figure 1).

The secretion of C-peptide, glucagon and ghrelin in supernatants was measured in unstimulated cells during differentiation stages. All three hormones were absent during the early stages, and cells started to release hormones from the EN stage, reaching 6650 ± 870 pg/mL of C-peptide, 3785 ± 720 pg/mL of glucagon and 955 ± 370 pg/mL of

ghrelin at the end of the differentiation (Figure 1D). iβ cells were then challenged with glucose and KCl in a dynamic perifusion assay to assess their insulin secretory profile, and they promptly responded to glucose stimulus (599.3 ± 103.7 pg/mL of insulin 10 min after stimulus), particularly when it was enhanced with the phosphodiesterase inhibitor 3-isobutyl-1-methylxanthine (1239.2 ± 533.7 pg/mL). A depolarizing stimulus with KCl resulted in an insulin secretory response in iβ cells that was higher in the cells previously stimulated with only glucose (Figure 1E). Taken together, these data demonstrate that iPSCs have the capacity to differentiate into bona fide β cells that are able to secrete insulin in response to glucose.

Longitudinal scRNA-seq analysis reveals the emergence of committed endodermal and pancreatic lineages from PSCs

To analyze the heterogeneity of iPSC-derived cells and follow the transition dynamics from pluripotent to pancreatic EN cell, the authors performed single-cell transcriptomics experiments at late time points of differentiation: PE (day 14, stage five), EN (day 18, stage six) and iβ (day 25, stage seven). With regard to the PE stage, canonical correlation analysis (CCA) and clustering analysis of 1784 cells and 2107 genes, with 15 correlation components and resolution = 0.6 (see supplementary Table 3), resulted in six distinct clusters (Figure 2A). Expressed top genes build a specific signature for each cluster. Gene ontology analysis with the cell types function of Enrichr identified predominant clusters zero and three (59.2% and 7.7% of total cells, respectively) as β cells and cluster four (4.2% of total cells) as endoderm cells, whereas clusters one and two (14.2% and 11.7% of total cells, respectively) were cells comprised of mixed identities, with some genes leading back to the fibroblast cell type and a small group of cells (cluster five, 3.0% of total cells) expressing genes associated with the midbrain (Figure 2A). In addition, analysis of a list of selected genes related to pluripotency, pancreatic development and β-cell function revealed that cell preparations at this stage contained a high proportion of precursor cells (expressing *HES1* and *SOX9*), but no cells expressed the pluripotency marker *POU5F1*. At the same time, PE cells began expressing genes of pancreas commitment (such as *FOXA2, NEUROD1* and *PDX1*) and a small cluster of cells expressed β-cell markers like *NKX6.1, CHGA* and *INSULIN* (Figure 2B).

With regard to the EN stage, CCA and clustering analysis (correlation components = 18, resolution = 0.6) (see supplementary Table 3) were performed with 2628 cells and 1583 genes and resulted in five distinct clusters (Figure 2C). Gene ontology analysis with Enrichr ascribed clusters one and four to EN cells (β and α cells), cluster two to endodermal cells (genes of small intestine, stomach and liver) and cluster three to a contaminant population of fibroblasts. Cluster zero, which represented 45.8% of the cell preparation, was assigned to kidney, but in the first 60 genes, which were significantly upregulated in comparison to the other clusters, the authors identified genes expressed by pancreatic cells (*CXCL14, NPC2, SPARC, KRT18, KRT8, NRF2, PCLAF*) and genes expressed during pancreatic development (*CLDN6, MEIS1, BIRC5, RARRES2*), suggesting that cells in this cluster were committed to pancreas (Figure 2C). Looking at the expression of gene markers of pluripotency, β-cell development and function in comparison to the previous stage, it was evident that the cell preparation at the EN stage did not contain pluripotent residual cells and had a limited percentage of cells expressing stem cell markers, mainly represented in clusters zero and two, and that cells with pancreatic commitment (*FOXA2, PDX1*) were increasing and concentrating in the cluster that had a clear EN cell fate, as shown by the expression of *NEUROD1, NKX2.2, CHGA* and *INSULIN* (Figure 2D). The cluster with a gene signature attributable to midbrain, present at the PE stage, had disappeared.

With regard to the iPSC-derived β/iβ-cell stage, CCA and clustering analysis were performed with 2631 cells and 1853 genes and resulted in six distinct clusters (Figure 3A; also see supplementary Table 3). The analysis of gene ontology revealed the presence of as many as four clusters out of six that were assigned to pancreatic islets

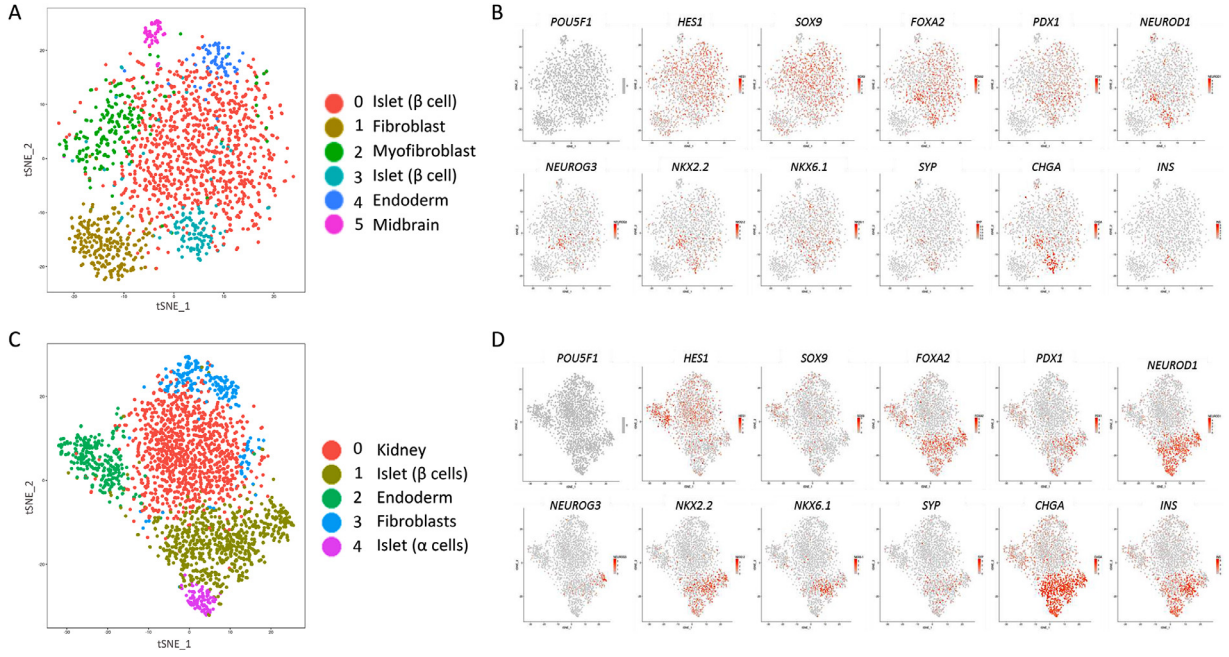

Figure 2. The scRNA-seq analysis of human iPSC-derived cells at PE and EN cell differentiation stages. (A) The t-SNE plot of scRNA-seq of cells at the PE stage. Colors denote different clusters of cells with similar feature expression patterns. On the right, cluster attribution inferred by enrichment analysis performed with Enrichr on DEGs with $P < 0.01$ of the six clusters. (B) The t-SNE plots colored according to the expression levels of specific features (feature plots) of PE cells. Here the authors considered genes involved in pancreatic differentiation. Intensity of red color indicates the normalized level (scaled UMI) of gene expression. (C) The t-SNE plot of scRNA-seq of cells at the EN stage. On the right, cluster attribution based on DEGs with $P < 0.01$ of the five clusters by gene ontology analysis with Enrichr. (D) Feature plots of genes involved in pancreatic differentiation in EN cells. Intensity of red color indicates the normalized level (scaled UMI) of gene expression. DEGs, differentially expressed genes; t-SNE, T-distributed stochastic neighbor embedding; UMI, unique molecular identifier.

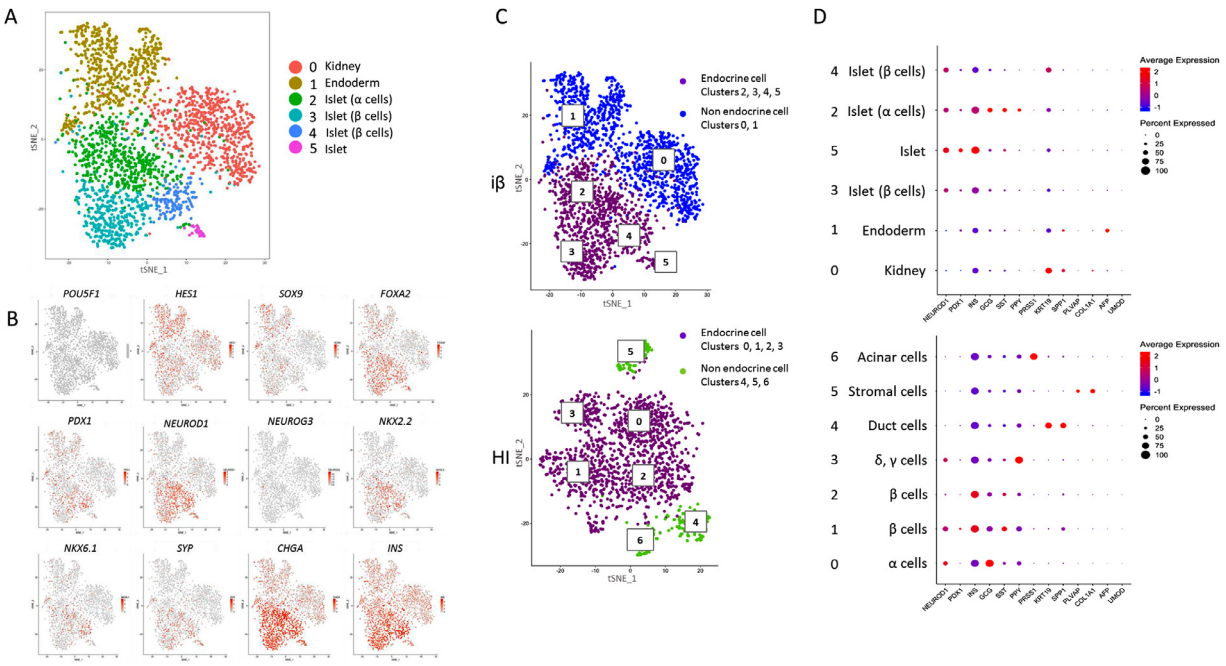

Figure 3. The scRNA-seq analysis of human iPSC-derived cells at the terminal differentiation stage of iβ cells and comparison with HI cells. (A) The t-SNE plot of scRNA-seq of cells at the iβ stage. On the right, cluster attribution based on DEGs with $P < 0.01$ of the six clusters by enrichment analysis with Enrichr. (B) Feature plots of genes involved in pancreatic differentiation in iβ cells. Intensity of red color indicates the normalized level (scaled UMI) of gene expression. (C) The t-SNE plot of scRNA-seq of iβ (upper panel) and HI (lower panel) cells, indicating the cells of the clusters related to EN cells (violet), non-EN cells of iβ cells (blue) and non-EN cells of HI cells (green). (D) On the left, cluster number and cluster identity based on DEGs with $P < 0.01$ across the six (for iβ cells) or seven (for HI cells) clusters by Enrichr enrichment analysis. On the right, balloon plot representing, for each cluster, the average expression of a set of gene markers of pancreatic progenitors (*NEUROD1, PDX1*), hormone-producing EN cells (*INS, GCG, SST, PPY*) and acinar (*PRSS1*), ductal (*SPP1, KRT19*), endothelial (*PLVAP*), stromal (*COL1A1*), liver (*AFP*) and kidney (*UMOD*) cells (from left to right). Dots show the percentage of cells expressing the gene (dot size) and the level of expression (scaling from blue to red). The cluster attributed to kidney does not express the kidney marker *UMOD* but strongly expresses ductal markers *SPP1* and *KRT19*. DEGs, differentially expressed genes; t-SNE, T-distributed stochastic neighbor embedding; UMI, unique molecular identifier.

or β and α cells, confirming the maturation of iPSCs in EN cells that produced hormones, although the authors also observed the persistence of cells expressing genes of progenitor cells (*HES1* and *SOX9*) (Figure 3B). Cluster zero (28.5%), at the EN stage, was attributed to kidney, but genes highly expressed in the pancreas were widely represented. The cluster identified as endoderm at the EN stage (with upregulated genes of small intestine, stomach and liver) persisted at this stage (cluster one, 25.5% of total cells), as evidenced by the fact that the process of differentiation was not complete and cells had gone in different directions.

The genes used as markers of the differentiation process in TaqMan bulk analysis (*OCT4, NANOG, SOX17, FOXA2, HNF1b, PDX1, NKX2.2, NKX6.1, INS* and *GCG*) (Figure 1A) confirmed the trend in expression at the single-cell level (see supplementary Figure 2). Overall, at the PE stage, the iPSC-derived cell population consisted mainly of pancreatic EN and posterior foregut-committed cells plus a significant proportion of fibroblasts and a small contribution of ectodermal cells. At the EN stage, cells were mainly EN progenitors and pancreatic EN and posterior foregut-committed cells plus a small proportion of fibroblasts. Most of the cells were pancreas or pancreas-committed, although they were at different stages of maturation. Finally, iβ cells did not contain pluripotent cells, nor cells of other lineages, only endodermal cells, mainly consisting of posterior foregut-committed and pancreatic EN cells.

Comparison of iPSC-derived β cells and donor islets through scRNA-seq analysis reveals high level of homology

Transcriptomics of single cells can aid in understanding the distance between β cells produced *in vitro* with this protocol and mature β cells obtained from the digestion of the pancreas of organ donors. The authors performed single-cell sequencing on two preparations of hand-picked HIs from organ donors (1747 cells, 2605 genes) (see supplementary Figure 3; see supplementary Table 3) and compared HI with iPSC-derived β cells. Clustering analysis of HI and iβ cells clearly showed that there were three different families of clusters (Figure 3C). The first included cells with a transcriptome of pancreatic EN cells (clusters two, three, four and five of iβ cells and clusters zero, one, two and three of HI cells, which accounted for 46% and 80.5% of cells in the iβ and HI groups). A second family included pancreatic cells contaminating the islet preparation—that is, stromal and acinar cells (clusters four, five and six of HI cells, 19.5%)—which was not present in iβ cells. The third family included endoderm-committed cells and progenitors (clusters zero and one of iβ cells, 54%), which could not be found in HI cells. The authors then looked at the average expression and percentage of cells expressing gene markers of these families: *NEUROD1* and *PDX1* for progenitors; *INS, GCG, SST* and *PPY* for EN cells; *SPP1* and *KRT19* for ductal cells; *COL1A1* for stroma; *PLVAP* for endothelium; *PRSS1* for acinar cells; *AFP* for liver; and *UMOD* for kidney. This analysis confirmed that the authors' differentiation protocol generated not only mature β cells but also polyhormonal immature EN cells, which often maintain the expression of early transcriptional factors like *NEUROD1* and *PDX1*. In addition, non-EN clusters zero and one confirmed their pancreatic nature since they expressed insulin and ductal markers *KRT19* and *SPP1* (Figure 3D).

Since HI cells from organ donors and iPSC-derived β cells have a high degree of correlation (Figure 4A,B), the authors performed CCA and re-clustering analysis of the clusters in HI e iβ specifically recognizable as islet cells (clusters zero, one, two and three of HI cells and clusters two, three, four and five of iβ cells), which determined the formation of four clusters (zero, one, two, three), in which three accounted for >90% of the cells (zero, one and two) (Figure 4C,D). The three main clusters segregated different types of EN cells: β cells in cluster zero (*INS*-positive), α cells in cluster one (*GCG*-positive) and γ cells in cluster two (*PPY*-positive). The small cluster three was not attributable to any kind of islet cells or other contaminating recognizable cell types (Figure 4E).

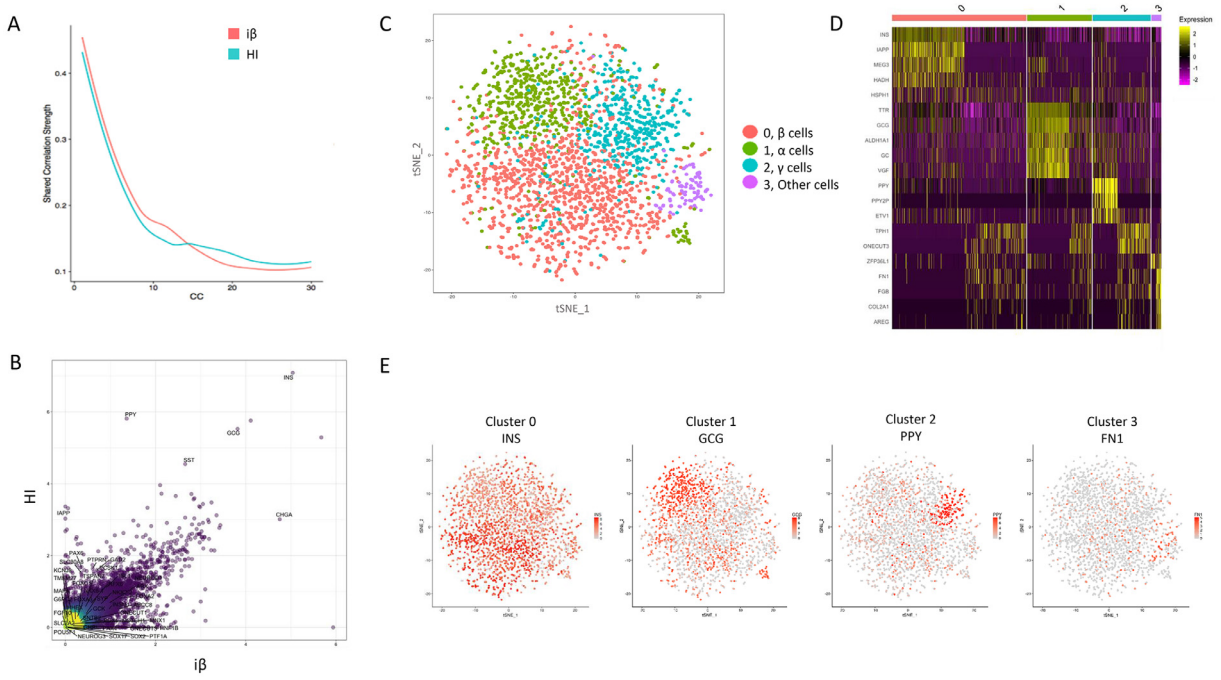

Figure 4. Correlation analysis of pancreatic EN cells of iβ and HI cells from organ donors. (A) Scree plot indicating the strength of shared correlation between iβ and HI cells explained by each CCA component. (B) Scatterplot representing the average expression in iβ and HI cells of commonly detected genes. *INS* is the most expressed gene in both populations. (C) The t-SNE plot of the merged preparations of EN cells of iβ and HI cells. Four clusters are shown in the tSNE plot. On the right, cluster attribution based on DEGs with $P <$ 0.01 of the four clusters by enrichment analysis with Enrichr. (D) Heatmap of the five top marker genes in the four clusters. (E) Feature plots of representative markers for each cluster (cluster zero, *INS* for Iβ cells, cluster one, *GCG* for α cells, cluster two, *PPY* for γ cells, cluster three, *FN1* for other cells). DEGs, differentially expressed genes; t-SNE, T-distributed stochastic neighbor embedding.

Cells in the T-distributed stochastic neighbor embedding clusters were similar in terms of expression of genes of the secretory machine, differentiation and stemness, suggesting that the differences among HI and iβ cells lay in the non-EN clusters. CCA analysis of HI cells and cells belonging to the β-cell clusters of iβ cells showed high similarity between the two sources, which equally contributed to the UMAP plot representation (see supplementary Figure 4).

Pseudotime trajectory identifies transcriptional signatures of EN commitment during the differentiation process

By using pseudotemporal analysis, the authors reconstructed the sequence of gene expression profiles and the progression from iPSC to the PE, EN and iβ stage, classifying cells without using known marker genes. The authors included in this analysis the single-cell transcriptomics data of iPSCs at day zero, before the beginning of the differentiation process, so as to use the pluripotency stage as the tree roots. The pseudotime trajectory assumed a tree-like structure where cells were distributed among branches. The authors found two main paths, enabling the allocation of three pseudotime-dependent progression states for iPSC-derived cells: pluripotent, EN and non-EN cells (Figure 5A). Upregulation of pancreatic EN markers (*INS, GCG, CHGA, CHGB, SST, PPY, ABCC8, SLC2A3, KCNJ6*) and transcriptional factors (*NEUROD1, FOXA2, ONECUT3, PDX1, NKX2.2, NKX6.1, ARX*) identified the EN branch. Similarly, increasing expression of known genes (*KRT19, KRT8, SPP1, SPARC, COL3A1, CLDN6, CXCL14, NPC2*) defined the non-EN branch (see supplementary Figure 5). The trend in the expression of some of the genes that identified the three progression states (pluripotency, EN and non-EN) in the two EN and non-EN branches during the pseudotime is represented in Figure 5B. Pseudotime analysis can also be used to discover which clusters identified during the differentiation stages contributed to the development of β cells: if we evaluate the positioning on the trajectory of the different clusters of PE and EN cells, we can clearly observe that clusters zero and three (islet) of the PE stage and, even more, clusters one and four (islet) of the EN stage were those that mostly contributed to the derivation of β cells (see supplementary Figure 6).

Pseudotime re-analysis of the EN branch identifies different subtypes of EN cells

The authors performed pseudotime analysis on iβ cells to highlight putative different cell fates (or intermediate states) inside the EN branch identified from previous pseudotime analysis, organizing cells into a new trajectory of evolutional nodes. The first node generated a sub-branch called A (Figure 6A), which contained the real β cells, expressing high insulin levels (Figure 6B), but also other typical β-cell markers such as *CPE, SCG5, SCG3, ISL1, ABCC8, SYT4, GAD2, MAFB, CHGA, CHGB, NKX6-1*. The cells that continued on the evolutionary path then met another divergence point, where they hesitated in tip B or continued along another branch, from which they emerged, in turn, at tip C and the other two most distal tips, D and E. In tip B, there were mainly polyhormonal pancreatic EN cells, probably blocked at an earlier stage of functional maturity, these cells in fact overexpressing *GCG, SST* and *PPY* genes and other pancreatic islet genes, such as *ARX, TPH1, DDC, NEFM, CD24, CPE, ONECUT3, SCG2* and *SCG3*. Cells in tip C, by contrast, were ε cells, which expressed high levels of ghrelin, some in co-expression with insulin, whereas others were exocrine pancreatic cells, as could be understood from the expression of genes like *GP2* and *HES1*. In tip D, there were also islet progenitor cells, as evidenced by the overexpression of *CD24*, a marker of pancreatic progenitors used for the selection of EN precursors during differentiation [32]. Finally, at the last branch point of the pseudotime, there were two tips, which could be combined in one group, as their division was mainly based on the differential expression of housekeeping and cell matrix genes. This group, called tip E, contained polyhormonal EN cells or, as reported in a recent article [3], alpha-like polyhormonal cells, as they overexpressed *GCG, SST, PPY* and *ARX*. The expression of the most significant genes in determining this pseudotime and their distribution

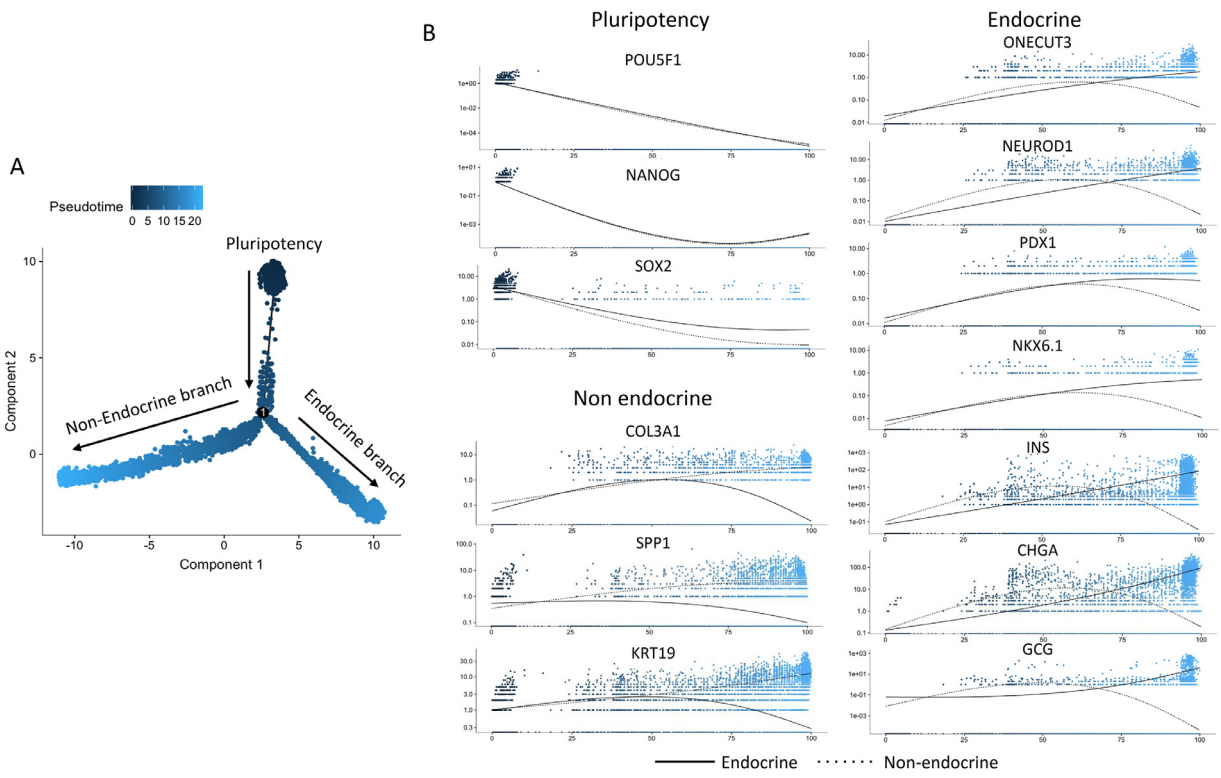

Figure 5. Pseudotime trajectories of iPSCs during differentiation into β cells. (A) Pseudotime trajectory of cells from iPSCs before differentiation and PE, EN and iβ cells during differentiation. From root state of pluripotency, cells along trajectory divided into two branches: EN and non-EN. (B) Pseudotime kinetics from root (undifferentiated iPSCs) to EN (continuous line) and non-EN (dashed line) cell fate. Three groups of genes are shown: those related to pluripotency (*POU5F1*, *NANOG*, *SOX2*), those related to non-EN cells (*COL3A1*, *SPP1*, *KRT19*) and those related to EN cells (*ONECUT3*, *NEUROD1*, *PDX1*, *NKX6.1*, *INS*, *CHGA*, *GCG*).

among the different sub-branch tips are shown in Figure 6C,D. Overall, the genes that defined the EN branch and its sub-branches were strongly correlated with each other and could be considered a network of interconnected genes.

The most distal tips of the EN tree were also characterized by the presence of an EN cell type that expressed *CHGA*, *TPH1*, *LMX1A*, *FEV*, *DDC* and *SLC18A1* and resembled enterochromaffin cells, recently described in stem cell-derived EN cells (see supplementary Figure 7) [3].

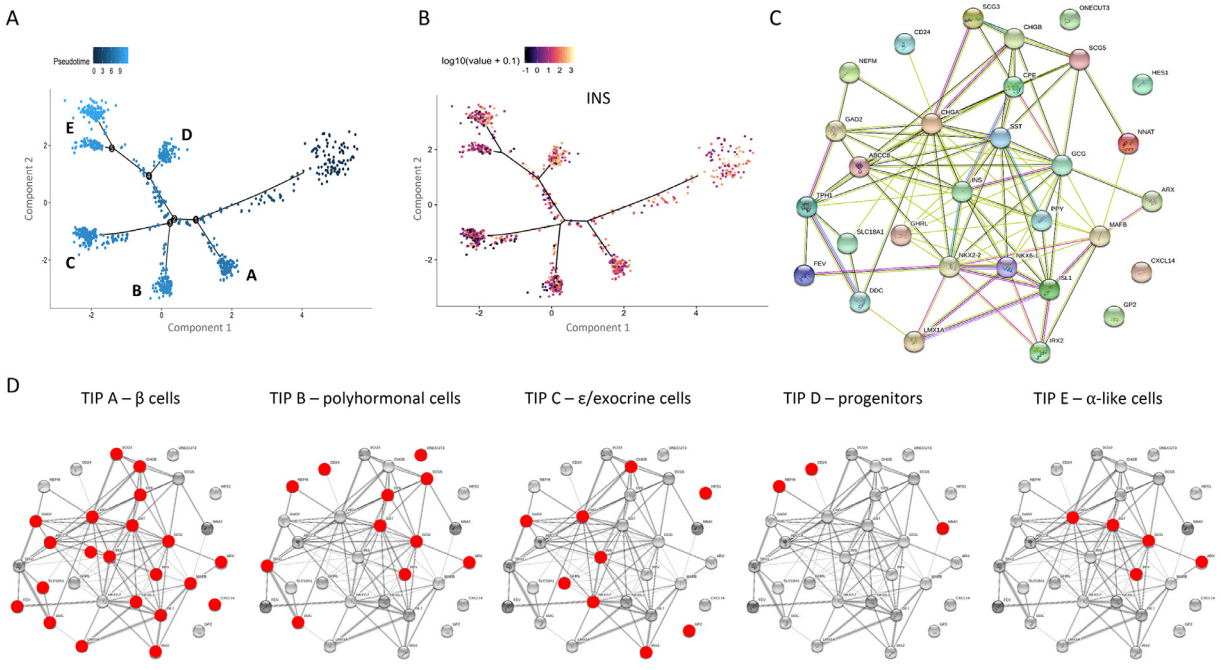

Figure 6. Pseudotime trajectories of iβ cells. (A) Pseudotime trajectory ordering cells at the last stage of differentiation. Branch tips are nominated from A to E. (B) Pseudotime kinetics of *INS* gene. Cells are colored by marker expression level (from black to yellow). (C) Functional enrichment analyses of the most significant genes used to construct single-cell trajectory using STRING 11.0. (D) Interaction maps of the most significant genes leading pseudotime progression. Red dots indicate the DEGs for each tip. DEGs, differentially expressed genes.

In conclusion, this pseudotime analysis underlines that, even after undertaking a path toward becoming EN cells, differentiated iPSCs can become mature β cells but also assume specific traits of other islet or immature EN cells.

Discussion

In this study, the authors have used a droplet-based scRNA-seq platform to map the transcriptional changes throughout the differentiation of iPSCs into pancreatic β cells, analyzing the late stages of differentiation, comparing the final product with human islets from organ donors and building a pseudotime progression based on the cell transcriptome of cells during differentiation. Although the differentiation protocol used has already demonstrated its efficacy with other iPSC lines [27,28] and in various publications, and despite the fact that the changes introduced here have led to further improvements, the transformation from iPSCs to β cells remains an incomplete process, both in quantitative and qualitative terms. Therefore, a percentage of off-target cells persists, and analysis of the different populations that alternate in the transition from iPSC to β cell remains crucial.

From the scRNA-seq analysis of iPSCs during differentiation emerges the finding that cells consistently undertake a differentiation in the EN lineage. In fact, only at the PE stage is it possible to trace a small population with an identity attributable to the ectoderm; this disappears in the next steps. The presence of a component of cells referable to midbrain has never been reported, but it is possible that this result is determined by a limitation of the authors' analysis. Indeed, to attribute an identity to the clusters, the list of the top 100 genes overexpressed in a cluster compared with all the others was submitted to the gene enrichment analysis of the Enrichr software, which hardly discriminates between different subtypes of pancreatic progenitors. For this reason, in this study, the authors integrated Enrichr results with an analysis made from known genes and hypotheses aroused by the previous literature. In addition to the small midbrain population present only at the PE stage, the rest of the cells appeared to have EN commitment, and part of them already showed a marked pancreatic EN cell fate. At the EN stage, committed EN cells increased, although a significant presence of cells with a transcriptome attributable to non-PE cells remained. At the iβ stage, all cells had a gene expression that led back to the endoderm, a significant part had become pancreatic islet cells and another substantial part had generic non-PE markers, including some liver markers. In addition, it is likely that the subpopulation of cells identified as kidney by Enrichr software actually contained a significant proportion of markers ascribable to pancreatic ductal cells. Even the work of Krentz *et al.* [25], although conducted with embryonic PSCs, identified off-target cells at an advanced stage of differentiation attributable to ductal and liver cells.

A strength of the authors' work is that we have been able to compare iPSC-derived β cells with cells of two human islet preparations obtained from the digestion of donor pancreas. The scRNA-seq of human islets gave results in line with previous publications, in particular by highlighting the presence, in addition to hormone-secreting EN cells, of residual pancreatic components such as ductal, stromal and acinar cells [11–15]. From the direct comparison with iPSC-derived β cells, it emerged that the EN component had a high degree of homology with EN cells of the islets, whereas iβ cells showed a delay or incorrect differentiation into non-EN clusters.

Here the authors performed the first pseudotime analysis of human iPSCs during differentiation into β cells. Pseudotime trajectory analysis clearly evidenced the dichotomic emergence of two distinct populations expressing either EN or non-EN markers, providing a gene signature of cells that undertook the "correct" or "incorrect" path. However, it should be emphasized that the cells identified as non-EN were actually pancreatic cells (mainly expressing ductal markers). These genes could be used as new targets of drugs/factors/small molecules aimed at directing differentiation. The authors also performed a re-analysis of cells that were classified in the EN branch. EN cells were further distributed in subgroups, among which the authors could identify mature β, α, δ, exocrine and polyhormonal immature EN cells. Furthermore, recent studies of scRNA-seq in PSCs report the appearance, during EN pancreatic differentiation, of genes associated with the production of serotonin (*FEV, TPH1, DDC*) that lead to the formation of enterochromaffin cells [3,24,25]. Moreover, in the authors' study, from the pseudotime analysis of differentiated cells in the EN branch emerged the presence of three small subgroups of cells expressing *CHGA, DDC, FEV, LMX1A, SLC18A1* and *TPH1*, which constitute a gene signature of enterochromaffin cells. Recent evidence demonstrates that β cells may release serotonin—for instance, as a mechanism of control of glucagon secretion by α cells [33]. In addition, serotonin has been shown to regulate β-cell mass by stimulating β-cell proliferation, so it cannot be excluded that this pattern of expression represents cells that are functional for β-cell terminal differentiation. To corroborate this observation, we checked for the expression of marker genes of entrochromaffin cells in our samples of human islets from organ donors and found that they are expressed (data not shown) https://doi.org/10.2337/db19-0546.

Conclusions

The authors' analysis of iPSCs at different stages of transition from pluripotency to pancreatic EN cells clearly shows that the changes cells undergo *in vitro* are accompanied by profound transcriptional transformations. The ability to follow these transformations at the single-cell level revealed that, at the beginning, the bulk of cells quickly undertook a massive pancreas specification. Most of the cells differentiated into pancreatic EN cells even if many of them did not become fully mature β cells. In addition, during the earlier stages, cells of other embryonic germ layers were formed, such as ectoderm and fibroblastic cells, which then did not persist. What persisted until the last stages were the progenitor cells, or cells with a still uncertain fate, and cells that underwent incomplete differentiation toward other cell types of the endoderm. Finally, the terminally differentiated cell product did not contain contaminants of pluripotent residual cells or cells differentiated into ectoderm or mesoderm, which could represent an important drawback. These results suggest that future efforts to produce iPSC-derived β cells as an alternative to donor islets for transplantation must aim not only to improve the EN cells obtained but also, above all, to avoid differentiation into non-PE cells.

Funding

This work was supported by the European Commission (H2020 grant no. 681070), Fondazione Diabete Ricerca ONLUS (Società Italiana di Diabetologia-SID), Eli Lilly Italia (grant Sostegno alla ricerca sul diabete 2017), Ministero dell'Università e della ricerca (PRIN2015, code no. 2015373Z39_010) and SOStegno 70 Insieme ai ragazzi diabetici Associazione Onlus (project "Beta is better").

Declaration of Competing Interest

The authors have no commercial, proprietary or financial interest in the products or companies described in this article.

Author Contributions

Conception and design of the study: Silvia Pellegrini, Lorenzo Piemonti, Valeria Sordi. Acquisition of data: Silvia Pellegrini, Francesca Giannese, Dejan Lazarevic, Fabio Manenti, Gaia Poggi, Marta Tiffany Lombardo, Alessandro Cospito, Rita Nano. Analysis and interpretation

of data: Silvia Pellegrini, Raniero Chimienti, Giulia Maria Scotti, Valeria Sordi. Drafting or revising the manuscript: Silvia Pellegrini. Raniero Chimienti, Giulia Maria Scotti, Lorenzo Piemonti, Valeria Sordi. All authors have approved the final article.

Acknowledgments

The authors thank Paola Macchieraldo, Antonio Mincione, Elena Riva, Antonio Civita, Andrea Marchesi and Michele Mainardi for supporting the fundraising campaign Un brutto t1po. Part of this work was carried out in the Advanced Light and Electron Microscopy BioImaging Center, an advanced microscopy laboratory established by IRCCS Ospedale San Raffaele and Università Vita-Salute San Raffaele.

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.jcyt.2020.10.004.

References

[1] Harb G, Poh Y-C, Pagliuca F. Stem Cell-Derived Insulin-Producing β Cells to Treat Diabetes. Curr Transplant Reports 2017;4:202–10.
[2] Millman JR, Pagliuca FW. Autologous Pluripotent Stem Cell–Derived β-Like Cells for Diabetes Cellular Therapy. Diabetes 2017;66:1111–20.
[3] Veres A, Faust AL, Bushnell HL, Engquist EN, Kenty JHR, Harb G, et al. Charting cellular identity during human in vitro β-cell differentiation. Nature 2019;569:368–73.
[4] Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotechnol 2014;32:381–6.
[5] Wagner A, Regev A, Yosef N. Revealing the vectors of cellular identity with single-cell genomics. Nat Biotechnol 2016;34:1145–60.
[6] Saelens W, Cannoodt R, Todorov H, Saeys Y. A comparison of single-cell trajectory inference methods: towards more accurate and robust tools. Nat Biotechnol 2018:276907. https://doi.org/10.1038/s41587-019-0071-9.
[7] Natarajan KN, Teichmann SA, Kolodziejczyk AA. Single cell transcriptomics of pluripotent stem cells: reprogramming and differentiation. Curr Opin Genet Dev 2017;46:66–76.
[8] Camp JG, Sekine K, Gerber T, Loeffler-Wirth H, Binder H, Gac M, et al. Multilineage communication regulates human liver bud development from pluripotency. Nature 2017;546:533–8.
[9] Loh KM, Chen A, Koh PW, Deng TZ, Sinha R, Tsai JM, et al. Mapping the Pairwise Choices Leading from Pluripotency to Human Bone, Heart, and Other Mesoderm Cell Types. Cell 2016;166:451–67.
[10] McCracken IR, Taylor RS, Kok FO, de la Cuesta F, Dobie R, Henderson BEP, et al. Transcriptional dynamics of pluripotent stem cell-derived endothelial cell differentiation revealed by single-cell RNA sequencing. Eur Heart J 2020;41:1024–36. https://doi.org/10.1093/eurheartj/ehz351.
[11] Baron M, Veres A, Wolock SL, Faust AL, Gaujoux R, Vetere A, et al. A Single-Cell Transcriptomic Map of the Human and Mouse Pancreas Reveals Inter- and Intra-cell Population Structure. Cell Syst 2016;3. 346–60.e4.
[12] Li J, Klughammer J, Farlik M, Penz T, Spittler A, Barbieux C, et al. Single–cell transcriptomes reveal characteristic features of human pancreatic islet cell types. EMBO Rep 2016;17:178–87.
[13] Muraro MJ, Dharmadhikari G, Grün D, Groen N, Dielen T, Jansen E, et al. A Single-Cell Transcriptome Atlas of the Human Pancreas. Cell Syst 2016;3:385–94. .e3.
[14] Enge M, Arda HE, Mignardi M, Beausang J, Bottino R, Kim SK, et al. Single-Cell Analysis of Human Pancreas Reveals Transcriptional Signatures of Aging and Somatic Mutation Patterns. Cell 2017;171. 321–30.e14.
[15] Qiu WL, Zhang YW, Feng Y, Li LC, Yang L, Xu CR. Deciphering Pancreatic Islet β Cell and α Cell Maturation Pathways and Characteristic Features at the Single-Cell Level. Cell Metab 2017;25. 1194–205.e4.
[16] Xin Y, Gutierrez GD, Okamoto H, Kim J, Lee AH, Adler C, et al. Pseudotime ordering of single human B-cells reveals states of insulin production and unfolded protein response. Diabetes 2018;67:1783–94.
[17] Teo AKK, Lim CS, Cheow LF, Kin T, Shapiro JA, Kang NY, et al. Single-cell analyses of human islet cells reveal de-differentiation signatures. Cell Death Discov 2018;4:1–11.
[18] Segerstolpe Å, Palasantza A, Eliasson P, Andersson EM, Andréasson AC, Sun X, et al. Single-Cell Transcriptome Profiling of Human Pancreatic Islets in Health and Type 2 Diabetes. Cell Metab 2016;24:593–607.
[19] Wang YJ, Schug J, Won KJ, Liu C, Naji A, Avrahami D, et al. Single-cell transcriptomics of the human endocrine pancreas. Diabetes 2016;65:3028–38.
[20] Xin Y, Kim J, Okamoto H, Ni M, Wei Y, Adler C, et al. RNA Sequencing of Single Human Islet Cells Reveals Type 2 Diabetes Genes. Cell Metab 2016;24:608–15.
[21] Lawlor N, George J, Bolisetty M, Kursawe R, Sun L, Sivakamasundari V, et al. Single-cell transcriptomes identify human islet cell signatures and reveal cell-type-specific expression changes in type 2 diabetes. Genome Res 2017;27:208–22.
[22] Zeng C, Mulas F, Sui Y, Guan T, Miller N, Tan Y, et al. Pseudotemporal Ordering of Single Cells Reveals Metabolic Control of Postnatal β Cell Proliferation. Cell Metab 2017;25. 1160–75.e11.
[23] Ramond C, Beydag-Tasöz BS, Azad A, van de Bunt M, Petersen MBK, Beer NL, et al. Understanding human fetal pancreas development using subpopulation sorting, RNA sequencing and single-cell profiling. Development 2018;145:dev165480. https://doi.org/10.1242/dev.165480.
[24] Balboa D, Saarimäki-Vire J, Borshagovski D, Survila M, Lindholm P, Galli E, et al. Insulin mutations impair beta-cell development in a patient-derived iPSC model of neonatal diabetes. Elife 2018;7:e38519. https://doi.org/10.7554/eLife.38519.
[25] Krentz NAJ, Lee MYY, Xu EE, Sproul SLJ, Maslova A, Sasaki S, et al. Single-Cell Transcriptome Profiling of Mouse and hESC-Derived Pancreatic Progenitors. Stem Cell Reports 2018;11:1551–64.
[26] Ricordi C, Lacy PE, Finke EH, Olack BJ, Scharp DW. Automated method for isolation of human pancreatic islets. Diabetes 1988;37:413–20.
[27] Pagliuca FW, Millman JR, Gürtler M, Segel M, Van Dervort A, Ryu JH, et al. Generation of Functional Human Pancreatic β Cells *In Vitro*. Cell 2014;159:428–39.
[28] Pellegrini S, Manenti F, Chimienti R, Nano R, Ottoboni L, Ruffini F, et al. Differentiation of Sendai Virus-Reprogrammed iPSC into β Cells, Compared with Human Pancreatic Islets and Immortalized β Cell Line. Cell Transplant 2018;27:1548–60.
[29] Huijbregts L, Kjær Petersen MB, Berthault C, Hansson M, Aiello V, Rachdi L, et al. Bromodomain and Extra Terminal Proteins Inhibitors Promote Pancreatic Endocrine Cell Fate. Diabetes 2019;68:db180224.
[30] Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. Cell 2015;161:1202–14.
[31] Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol 2018;36:411–20.
[32] Jiang W, Sui X, Zhang D, Liu M, Ding M, Shi Y, et al. CD24: a novel surface marker for PDX1-positive pancreatic progenitors derived from human embryonic stem cells. Stem Cells 2011;29:609–17.
[33] Almaça J, Molina J, Menegaz D, Pronin AN, Tamayo A, Slepak V, et al. Human Beta Cells Produce and Release Serotonin to Inhibit Glucagon Secretion from Alpha Cells. Cell Rep 2016;17:3281–91.