


METHOD

Open Access



Using single-nucleus RNA-sequencing to interrogate transcriptomic profiles of archived human pancreatic islets

Giorgio Basile^{1†}, Sevim Kahraman^{1†}, Ercument Dirice^{1,2†}, Hui Pan³, Jonathan M. Dreyfuss³ and Rohit N. Kulkarni^{1,4,5*} 

Abstract

Background: Human pancreatic islets are a central focus of research in metabolic studies. Transcriptomics is frequently used to interrogate alterations in cultured human islet cells using single-cell RNA-sequencing (scRNA-seq). We introduce single-nucleus RNA-sequencing (snRNA-seq) as an alternative approach for investigating transplanted human islets.

Methods: The Nuclei EZ protocol was used to obtain nuclear preparations from fresh and frozen human islet cells. Such preparations were first used to generate snRNA-seq datasets and compared to scRNA-seq output obtained from cells from the same donor. Finally, we employed snRNA-seq to obtain the transcriptomic profile of archived human islets grafted in immunodeficient animals.

Results: We observed virtually complete concordance in identifying cell types and gene proportions as well as a strong association of global and islet cell type gene signatures between scRNA-seq and snRNA-seq applied to fresh and frozen cultured or transplanted human islet samples.

Conclusions: We propose snRNA-seq as a reliable strategy to probe transcriptomic profiles of freshly harvested or frozen sources of transplanted human islet cells especially when scRNA-seq is not ideal.

Keywords: Transplanted human islets, Single-nucleus RNA-sequencing, Single-cell RNA-sequencing, Human β -cells, Frozen samples, Archived tissue, Diabetes

Background

Type 1 diabetes (T1D) and type 2 diabetes (T2D) are both characterized by a progressive reduction of functional mass of insulin-producing β -cells [1, 2]. Therefore, restoring physiological numbers of endogenous β -cells, improving β -cell functionality, or generating insulin-producing β -like cells derived from stem cells for

transplantation are promising strategies to resolve diabetes in patients. While the identification of factors able to stimulate β -cell proliferation [3–11] and the improvement of differentiation protocols to generate functional β -like cells [12–17] continue to evolve, gaining insights into dynamic changes in the global transcriptome of β -cells, especially following manipulation in an in vivo environment (e.g., in transplanted islets in humanized mouse models) is worth exploring.

One challenge when investigating the biology of pancreatic islets is the presence of at least five different hormone-secreting endocrine cell types and non-endocrine cells such as endothelial cells, glia, fibroblasts,

* Correspondence: Rohit.Kulkarni@joslin.harvard.edu

[†]Giorgio Basile, Sevim Kahraman and Ercument Dirice contributed equally to this work.

¹Section of Islet Cell and Regenerative Biology, Joslin Diabetes Center and Harvard Medical School, Boston, MA 02215, USA

⁴Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA
Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

pericytes, and immune cells (tissue-resident macrophages, mast cells, B cells, cytotoxic T cells) [18]. A mixture of such diverse cell types can hinder precise identification of cell-specific transcriptomes or exclusive biological signals when bulk tissue analyses are performed. To circumvent this limitation, over the past few years, multiple groups have utilized single-cell RNA-seq (scRNA-seq) methodologies on islet cells isolated from mouse [19, 20] or human pancreas [21–24]. These studies have focused on dissecting the transcriptomic signature of endocrine cell types across different ages [25] and to define differentially expressed genes in type 2 diabetic β -cells [26–29]. However, single-cell preparations are also known to have limitations, including the need to harvest live cells [30] which may inadvertently induce stress responses [31, 32]. These aspects gain relevance when identifying disease-related transcriptomic signatures in tissues obtained from multiple human donors at different time points that require immediate, although usually varied, processing and individual analyses.

To overcome these limitations [31], single-nucleus RNA-sequencing (snRNA-seq) has been employed in studies on tissues composed of diverse cell types, including the brain [33, 34], kidney [35–37], heart [38], skeletal muscle [39, 40], stria vascularis [41], retina [42], liver [43, 44], lung [45], or white [46, 47] and brown adipose tissue [48, 49] obtained from mice or human donors. Here, we report a side-by-side comparison between snRNA-seq and scRNA-seq to validate the robustness of the former as an alternative sequencing strategy. We propose that snRNA-seq is a reliable approach to interrogate the transcriptomic profiles of archived frozen tissues that, to our knowledge, has not been applied previously to pancreatic islets or engrafted tissues.

Methods

Animal studies

Healthy female 8–12-week-old non-obese diabetic (NOD)/severe combined immunodeficiency (SCID)- γ (NSG) mice were used as human islet transplant recipients. We used 4 animals in total, where a single animal received human islet preparation from a single donor. Animals were housed in the Animal Care Facilities at Joslin Diabetes Center on a 12-h light/12-h dark cycle with water and food ad libitum. Studies and protocols were approved by the Institutional Animal Care and Use Committee of the Joslin Diabetes Center (IACUC #05-01).

Human islet studies

Human islets were obtained from 5 non-diabetic brain-dead donors. Islet preparations were generated by the Integrated Islet Distribution Program or Prodo laboratories according to the standard procedures [50] (Additional file 1: Table S1). All studies and protocols used

were approved by the Joslin Diabetes Center's Committee on Human Studies (CHS#5-05). Upon receipt, human islets were centrifuged, washed, transferred to Petri dishes (2000–3000 human islets equivalents [IEQs]/plate), and cultured with fresh Miami Media #1A (Cellgro) overnight at 37 °C and 5% CO₂. Human islets from one donor ($N = 1$) were processed to isolate nuclei and obtain single-cell preparations as a common source for the scRNA-seq and snRNA-seq procedures (detailed below), while human islets from 4 donors ($N = 4$) were used in the kidney capsule transplantation experiments (detailed below).

Human islet transplantations

On the day of the experiment, 1000 hand-picked IEQs from 4 separate human islet donors were transplanted under the kidney capsule of 8-to-12-week-old male NSG mice and the animals were followed for 4 weeks. At the end of the follow-up period, mice were sacrificed by cervical dislocation. Human islet grafts were rapidly dissected under the microscope, snap-frozen, and stored at -80 °C.

Single-nucleus isolation

Frozen or freshly cultured or engrafted human islet samples were transferred to Dounce tissue grinder tubes (D8938; Sigma) containing 0.5 ml ice-cold Nuclei EZ lysis buffer (NUC-101; Sigma) and homogenized with pestles A and B for 1 min each on ice. Samples were transferred to clean 15 ml tubes and homogenizers were rinsed with 1.5 ml buffer followed by 2 ml buffer and transferred to the same 15 ml tubes to obtain a final volume of 4 ml. The tubes were vortexed briefly at moderate speed and kept on ice for 5 min for cell lysis. To separate the nucleus and cytoplasm, the tubes were centrifuged at 500 \times g for 5 min at 4 °C. Supernatants containing cytoplasmic components were saved for later analyses. The pellet containing nuclei was resuspended in 0.5 ml cold buffer by vortexing briefly at moderate speed followed by the addition of 3.5 ml cold buffer. The nuclear suspension was mixed by vortexing briefly and set on ice for 5 min. The tubes were centrifuged at 500 \times g for 5 min, the supernatant was saved for later analysis, and the pellet was resuspended in suspension buffer (0.5 ml PBS containing 0.01% non-acetylated bovine serum albumin (BSA); Sigma and 0.1% RNase inhibitor; 2313A from Clontech). The nuclear suspension was pipetted ten times with a 1 ml tip, filtered through a 30- μ m pre-separation filter (130-041-407; Miltenyi Biotec); cell number and cell viability were determined by cell counter using 0.4% trypan blue stain. The average number of total nuclei obtained from one-half graft was approximately 8.5×10^5 nuclei (1.7×10^6 cells/ml) with 5–10 μ m size and $93.3 \pm 1.1\%$ dead cell rate ($n = 32$

samples across three independent experiments). The number of isolated nuclei was adjusted to 1000 nuclei/ μ l with suspension buffer and 10,000 nuclei were immediately used for the generation of Gel Beads In-Emulsion (GEMs) and barcoding. Leftover nuclei were saved for subsequent analysis.

Human islet dispersion

Human islets were dissociated into single cells by TrypLE (12604-013; Thermo). Briefly, 1 ml TrypLE was added to human islet pellets and incubated at 37 °C for 12 min by mixing the tube every 3–4 min. At the end of incubation, TrypLE was neutralized by adding 9 ml cold Dulbecco's modified Eagle's medium with high glucose (DMEM HG, MT 10-017-CV; Corning) containing 10% fetal bovine serum (FBS) (10437028; Gibco). Cells were centrifuged at 1200 rpm for 3 min at 4 °C, supernatant was removed, and cells were resuspended in 0.5 ml Dulbecco's phosphate-buffered saline (DPBS) (14190250; Gibco). The cell suspension was filtered through 30 μ m filter to remove aggregates and counted. A final concentration of 1×10^6 cells/ml cells was used for DAPI/Phase microscopy. For single-cell RNA-seq procedures, human islet cell suspensions were filtered using a 30- μ m filter to remove any aggregates, and dead cells were excluded using Dead Cell Removal Kit (Miltenyi Biotec). Cells were counted and 10,000 cells were used for the generation of GEMs.

DAPI/Phase microscopy

Fifty thousand dispersed human islet cells (50 μ l) or isolated nuclei from frozen grafts were mixed with 50 μ l of 4',6-diamidino-2-phenylindole (double-stranded DNA staining, DAPI) solution (D9564; Sigma) diluted 1000 times in DPBS. Stained cells (10 μ l) were loaded onto a Hemocytometer slide and imaged under a fluorescent microscope. Digital images were taken at 40 \times magnification with AXIO Imager A2 upright microscope equipped with X-Cite series 120Q light source, Axiocam 512 color camera. Bright field and fluorescent images were overlaid using ImageJ Software to determine complete cell lysis and nuclear integrity in isolated nuclear samples.

Immunohistochemistry

Paraffin-embedded human islet graft sections were immunostained using anti-insulin (1:400, ab7842; abcam), anti-glucagon (1: 10,000, ab92517; abcam), and anti-somatostatin (1:500, ab30788; abcam) antibodies using previously described techniques [6, 51, 52]. Nuclei were labeled using DAPI. Images were acquired using a Zeiss LSM-710 Confocal Microscope and the Zen Black software (Carl Zeiss).

Western blotting

Nuclei or cell pellets were lysed in RIPA buffer (pH 7.4) containing 100 mM NaF, 50 mM Hepes, 150 mM NaCl, 10% Glycerol, 1.2% Triton X, 1 mM MgCl₂, 1 mM EDTA, 1 mM Na₃VO₄, protease inhibitor cocktail (P8340; Sigma), phosphatase inhibitor 2 (P5726; Sigma), and phosphatase inhibitor 3 (P0044; Sigma). The supernatants collected after the first and second centrifuge steps of nuclear isolation were used after adding protease and phosphatase inhibitors. Total protein concentration was determined by Pierce BCA Protein Assay Kit (23225; Thermo). Lysates (50 μ g protein) were run in 8% SDS-PAGE and transferred to PVDF membrane (Millipore). Membranes were blocked for 10 min at room temperature with 5% milk and were incubated overnight at 4 °C with antibodies against Lamin A/C (1:1000, 4777; Cell Signaling Technology) or GAPDH (1:1000, 5174; Cell Signaling Technology). After three washes (10 min), the membranes were incubated for 1 h at RT with antibodies against rabbit IgG-HRP conjugate (1:1000, 170-6515; Bio-Rad) or mouse IgG-HRP conjugate (1:1000, 170-6516, Bio-Rad). After three 10-min washes, signals were visualized via Pierce ECL Western blotting substrate (PI32106; Thermo).

RNA extraction and analysis

Cells were lysed in RLT buffer and RNA was extracted using the Qiagen RNeasy kit according to the manufacturer's instructions. RNA concentrations were measured by Nanodrop (Thermo). RNA integrity was determined by using Agilent RNA 6000 Nano Kit (5067-1511; Agilent) according to the manufacturer's instructions (Joslin Genomics Core).

Single-nucleus and single-cell RNA-sequencing procedures

Gel Bead In-Emulsion (GEMs) were generated using the Chromium 3' Single Cell Library Kit (v2, 10X Genomics, CA) according to the manufacturer's instructions and adapting the adjustments for the cDNA and libraries amplification steps, as recommended by 10X genomics for snRNA-seq procedures. Briefly, 10,000 cells or nuclei were combined with Single Cell Master Mix and encapsulated into the barcoded Gel Beads through the Chromium™ Controller. After GEM-RT incubation, cDNA samples were recovered, purified, and amplified through a cDNA Amplification Reaction using a 14-cycle setting. Quality controls on the undiluted amplified cDNA samples were performed using a High Sensitivity DNA Kit (Agilent, CA) on a 2100 BioAnalyzer (Agilent, CA) platform. Libraries were then constructed following Fragmentation and Adaptor Ligation. Sample Index PCR was performed adjusting the reactions at 15 cycles. Finally, purified libraries were run on 2100 BioAnalyzer (Agilent,

CA) using a High Sensitivity DNA Kit (Agilent, CA) to evaluate the quality of the ~400 bp fragments.

Next generation sequencing

Single-nucleus and single-cell libraries were sent for sequencing at GeneWiz or NovoGene laboratories. Samples were run in independent lanes on a HiSeq 4000 platform (Illumina, CA), using a coverage of 500,000 pair-ended reads targeted per cell.

Single-nucleus and single-cell RNA-seq data analysis

Both the single-nucleus (snRNA-seq) and single-cell (scRNA-seq) RNA-sequencing datasets were produced using cultured human islet samples obtained from the same donor. In particular, we generated single-cell and single-nucleus libraries using 4 technical replicates for each method. Gene counts were obtained by Cell Ranger (10x Genomics, CA) using the human reference genome. To eliminate empty droplets and technical artifacts, we applied Cell Bender [53]. Cell Bender concludes that droplets are empty if the probability that they are empty > 50%, and then estimates ambient RNA among non-empty droplets. We normalized each sample's data using *sctransform* [54], part of the Seurat toolkit, and detected and removed doublets with DoubletFinder [55]. To analyze the complete dataset, we combined the 4 samples, including genes that are detected in at least 3 cells and including cells where at least 200 genes are detected. From the combined dataset, we filtered out nuclei that have more than 20% mitochondrial unique molecular identifiers (UMI).

For the engrafted snRNA-seq, gene counts were generated by Cell Ranger (10x Genomics, CA) using the human-mouse joint reference genome. To eliminate empty droplets and technical artifacts, we applied Cell Bender [53]. We normalized each sample's data using *sctransform* [54], part of the Seurat toolkit, and detected and removed doublets with DoubletFinder [55]. To analyze the complete dataset, we combined the 4 samples, including genes that are detected in at least 3 cells and including cells where at least 200 genes are detected. From the combined dataset, we filtered out nuclei that have more than 20% mitochondrial UMI and more than 25% mouse UMI and removed all mouse genes in the remaining nuclei. For each of them (i.e., the engrafted islet snRNA-seq, the scRNA-seq, or the snRNA-seq from cultured human islets), we normalized the dataset of combined samples with *sctransform* [54], part of the Seurat toolkit; and then using Seurat, we identified clusters and marker genes per cluster and plotted the data as Uniform Manifold Approximation and Projection (UMAP), heatmaps, and violin plots. For all datasets, genes were considered expressed in a cell or nucleus if they had at least one UMI. We evaluated gene

expression levels between datasets per cell type by evaluating the linear regression of one dataset's mean natural log of counts per gene against the other datasets for the same gene and reporting the R^2 , which is the square of the Pearson correlation coefficient, and the p value.

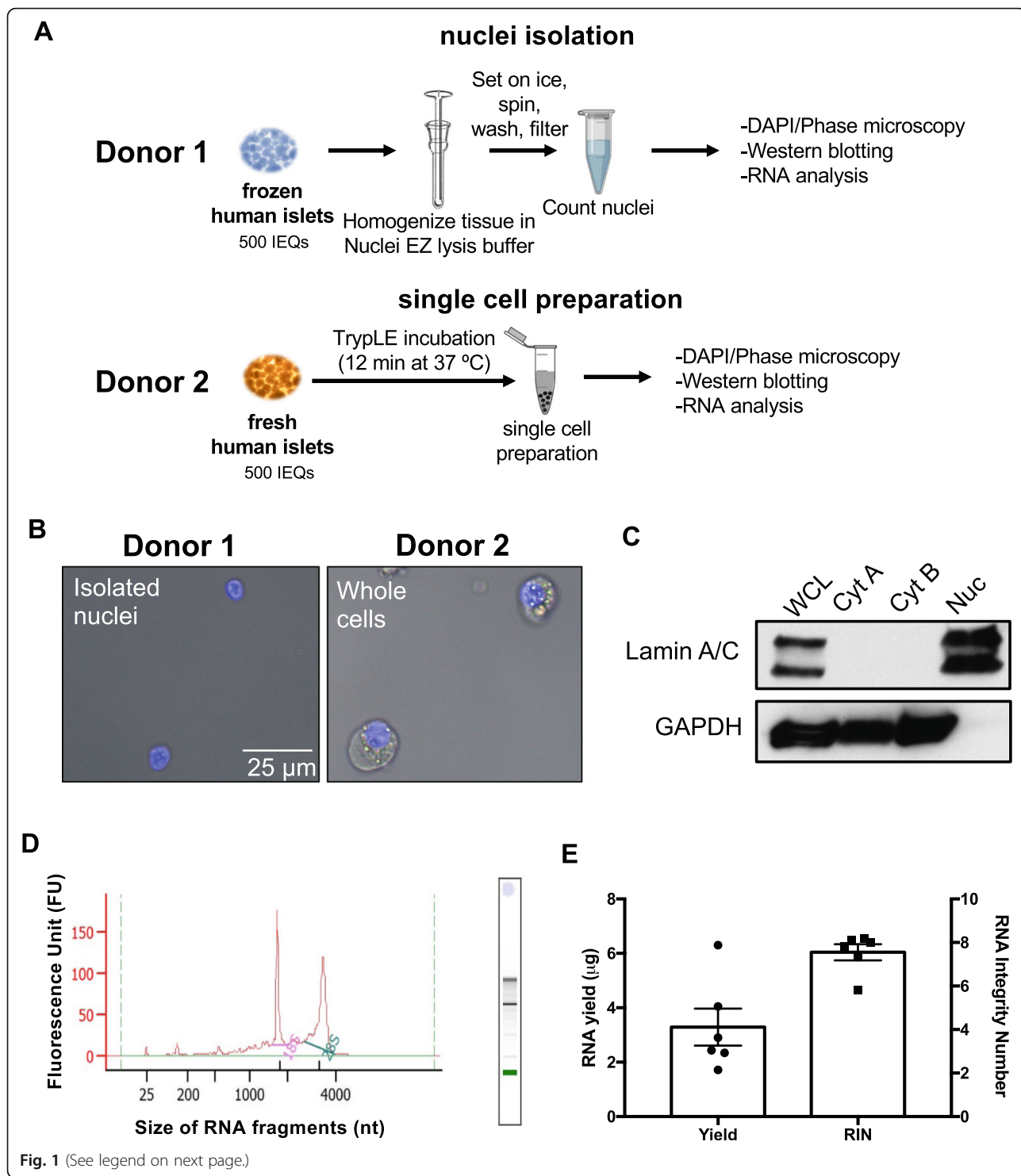
Integration with published scRNA-seq datasets

The ability of Seurat [56] to integrate single-cell datasets was demonstrated on publicly-available human pancreatic islet scRNA-seq datasets spanning 27 donors, four laboratories, and five technologies: The datasets that have been used in this article have been deposited in the public Genomic Spatial Event Database (GSE) and in the ArrayExpress Database at the EMBL-EBI. For InDrop, accession number GSE84133 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE84133>) [21]; for CelSeq2, accession number GSE85241 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE85241>) [23]; for SMART-Seq2, accession number E-MTAB-5061 (<https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-5061/>) [26]; for Fluidigm C1, accession number GSE86469 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE86469>) [27]; and for CelSeq, accession number GSE81076 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE81076>) [57] were used. We followed Seurat instructions to normalize these datasets with *sctransform* [54], identify integration anchors (we used dimensionality 25, consistent with the dimensionality we applied in Seurat analysis of our datasets), and construct our reference with Seurat function *IntegrateData*. We then used Seurat to project this universal scRNA-seq dataset, named "reference," onto our datasets (i.e., the engrafted islet snRNA-seq and the scRNA-seq and snRNA-seq from cultured human islets) to harmonize the data.

Results

Isolation of single nuclei from frozen human islets

To isolate nuclei from human islet preparations, we tested the Nuclei EZ lysis buffer-based protocol. We employed this isolation method because it was previously successfully used to isolate nuclei from frozen compact tissues, such as tumor tissues [58, 59]. In our hands, we validated that this protocol removes cytoplasmic content quickly and consistently. To test the efficacy of this protocol, we used a frozen sample of non-diabetic human islets (donor 1, Fig. 1A). Briefly, the archived human islet sample, consisting of 500 human islet equivalents (IEQs) was subjected to homogenization using the Nuclei EZ buffer. After a short incubation period, the cytoplasmic fraction was removed by centrifugation, and the pellet containing the nuclear fraction was washed to remove cytoplasmic contamination. The nuclei were filtered, counted, and tested in control assays including



(See figure on previous page.)

Fig. 1 Single nuclei isolation from frozen human islet samples. **A** Experimental workflow of isolation of nuclei from frozen cultured human islets (donor 1) and human islet dispersion into single cells (donor 2). **B** Isolated nuclei from frozen human islets (donor 1) and dissociated human islet cells from fresh intact islets (donor 2, whole cells) were imaged by DAPI/Phase microscopy. Left image: Intact nuclei stained in blue without cytoplasm surrounding the isolated nucleus. Right image: Nuclei stained in blue surrounded by cytoplasm of human islet cells. Scale bar is 25 μm . **C** Western blot analysis of human pancreatic islet cells (whole cell lysate; WCL, obtained from donor 2), supernatant collected after the first centrifuge step (cytosolic fraction A; Cyt A) and second centrifuge step (cytosolic fraction B; Cyt B) of the nuclear isolation protocol, and isolated nuclei (nuclear fraction; Nuc, obtained from Donor 1). LaminA/C is a nuclear marker; GAPDH is a cytoplasmic marker. **D** RNA was isolated from islet cell nuclei and analyzed by Bioanalyzer. Representative electropherogram showing two major peaks (18S and 28S rRNA) indicates minimal degradation of total RNA. The X-axis indicates the size of RNA fragments (nt: nucleotides) and the Y-axis represents the intensity of the fluorescence signal (FU, fluorescence unit). **E** Yield (μg , left Y-axis, dots) and RNA Integrity Number (RIN, right Y-axis, squares) of RNA isolated from six technical replicates across three independent experiments. Data are represented as mean \pm SEM

DAPI staining, protein, and RNA evaluation. To preclude confounding factors such as the quality of the isolation and the purity of the nuclear preparation we also obtained whole islet cell samples by dispersing a fresh sample of non-diabetic human islets, containing 500 IEQs (donor 2, Fig. 1A).

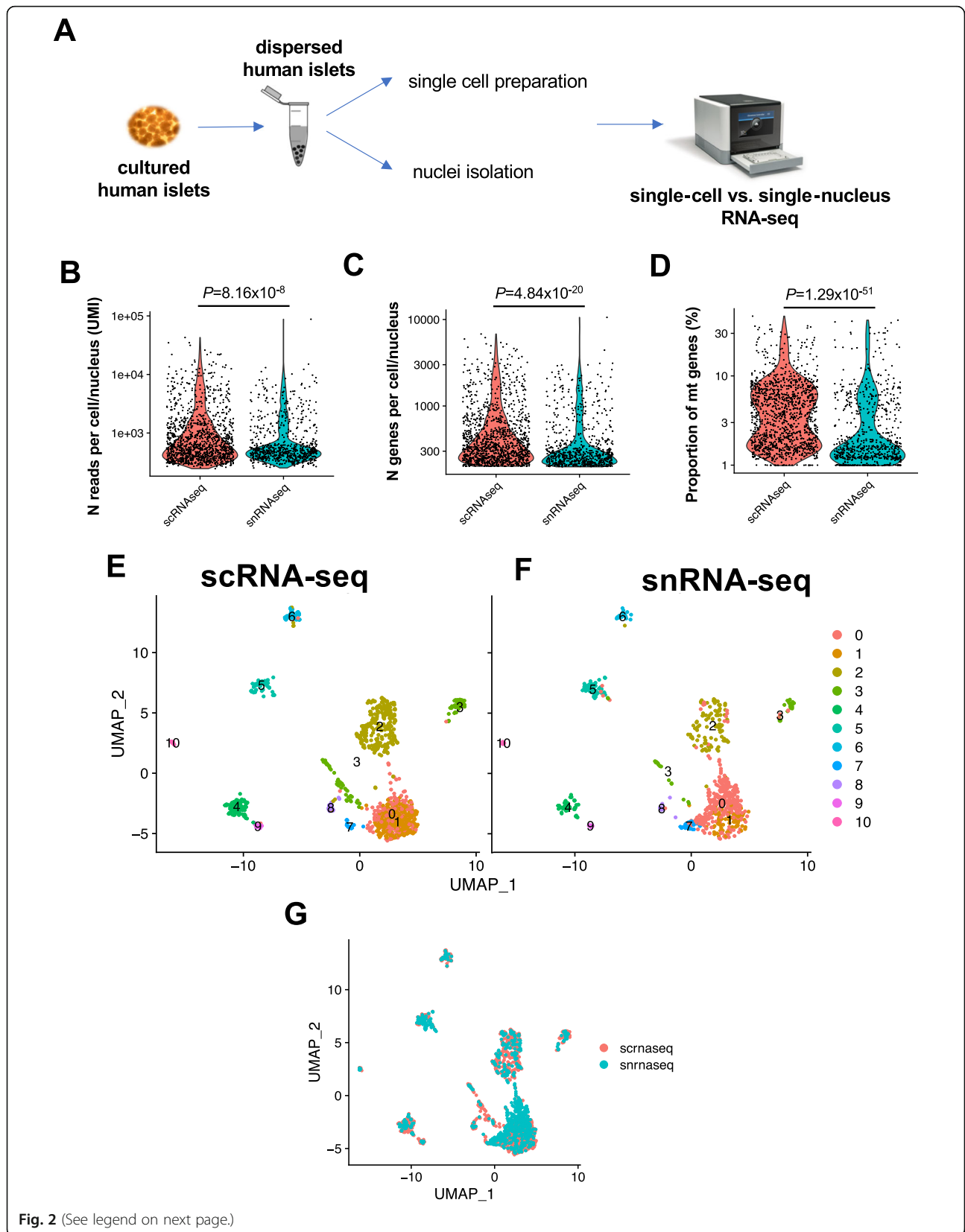
First, the quality of the isolation protocol was evaluated by visualizing the nuclei samples obtained from frozen human islets under phase-contrast microscopy and by assessing cell integrity in comparison to intact cell samples obtained by dissociating fresh human islets, collected by a separate donor, into single cells (whole cells), as previously described [60] (Fig. 1B). Phase-contrast images of isolated nuclei samples showed that the cytoplasm was depleted completely in all the cells, and the nucleus was intact (Fig. 1B). Whereas, as expected, the cytoplasm surrounding the DAPI-stained nucleus was easily distinguished in the dispersed single islet cells (Fig. 1B). Moreover, the lysis of the outer cell membranes in the samples processed for nuclei isolation was confirmed by assessing cytoplasmic (GAPDH) and nuclear (Lamin A/C) proteins in the samples collected during and after the isolation process (Fig. 1C). Western blot analysis of nuclei (Nuc), cytoplasmic fractions collected after the first (Cyt A) and the second (Cyt B) centrifuge steps, obtained by processing human islets from donor 1, and the whole cell lysate (WCL) sample, obtained by processing human islets from donor 2, showed that the nuclear isolation process efficiently removed the cytoplasm from nuclear samples, and preserved nuclear integrity, consistent with the DAPI staining (Fig. 1B). Indeed, we were able to isolate $8.32 \pm 1.6 \times 10^5$ nuclei from ~ 500 IEQs consistent with previous reports [61]. The amounts of nuclei were adequate to perform snRNA-seq—which required only 10,000 nuclei, as previously reported [48]—as well as for validation experiments such as qRT-PCR.

Total RNA was isolated from the nuclear samples, and the yield and integrity were evaluated using a Bioanalyzer [62] (Fig. 1D, E). We were able to isolate $3.3 \pm 0.7 \mu\text{g}$ RNA from isolated nuclear samples ($n = 6$; three independent experiments). The RNA yield was in the

expected range, considering that a mammalian cell contains 10–20 pg of total RNA of which 20–30% resides in the nucleus [63, 64]. RNA integrity evaluated by measuring two major peaks representing 18S and 28S rRNA revealed an average RNA integrity number (RIN) of 7.6 ± 0.4 which indicates high quality with minimal degradation [62]. Based on the efficacy of the tested protocol for obtaining pure high-quality nuclei from human islet samples, we decided to use this methodology for subsequent single-nucleus RNA-seq experiments.

Side-by-side comparison of scRNA-seq and snRNA-seq methods in cultured human islets

To compare the single-nucleus and single-cell RNA-sequencing procedures (snRNA-seq and scRNA-seq, respectively), we obtained human islets from a non-diabetic donor and divided them into two groups: (1) one to generate single-cell suspensions and (2) a second to isolate single nuclei using the protocol described above (Fig. 2A). We generated 4 technical replicates for each group of 10,000 cells or nuclei and loaded them into the 10X genomics Chromium Controller to obtain Gel-beads in Emulsion (GEMs). Following the 10X genomics standard protocol, we obtained single-cell and single-nuclei libraries that were then sequenced using next-generation sequencing (NGS). Following the application of the quality check filters, including removal of doublets and multiplets, we recovered 1277.2 ± 234.2 and 976.2 ± 82.0 (mean \pm SD) high-quality cells and nuclei per sample, respectively (Additional file 2: Table S2). Although the number of reads and genes sequenced per cell/nucleus was slightly higher in the scRNA-seq compared to the snRNA-seq method (Fig. 2B, C), the duplication rate, indicating the ratio between usable vs. sequenced reads, as a read-out of sequencing efficiency, were similar between the two transcriptomic procedures (Additional file 3: Figure S1C, Additional file 2: Table S2). As confirmation of the high purity of the single-nucleus preparations, the percentage of mitochondrial genes sequenced in the nucleus-containing droplets was lower than the cell-containing particles (2.7% and 4.2%, respectively) (Fig. 2D, Additional file 2: Table S2). We



(See figure on previous page.)

Fig. 2 Side-by-side comparison of single-cell vs. single-nucleus RNA-sequencing in freshly cultured human islets. **A** Experimental design of the workflow for generating single-cell RNA-seq (scRNA-seq) and single-nucleus RNA-seq (snRNA-seq) datasets of cultured human islets. **B–D** Violin plots representing **B** number of reads (UMI), **C** number of genes, and **D** proportion (expressed in %) of mitochondrial genes per cell/nucleus in scRNA-seq (orange violins) and snRNA-seq (blue violins) datasets. Data are expressed in logarithmic scale (Y-axis). Statistical significance was tested using a Wilcoxon rank-sum test. **E, F** UMAP plots representing the clustering of the high-quality **E** cells and **F** nuclei in the **E** scRNA-seq and the **F** snRNA-seq datasets. Y-axis of plot in **F** is like the plot in **E**. **G** Global integrated UMAP plot representing the distribution of high-quality cells (from scRNA-seq, orange dots) and nuclei (from snRNA-seq, blue dots)

obtained estimates of ambient RNA within the cell-/nucleus-containing droplets by applying Cell Bender [53], with averages of 5.87 ± 1.76 % and 1.42 ± 0.72 % (mean \pm SD) among the four single-cell and four single-nucleus library replicates, respectively (Additional file 2: Table S2, Additional file 3: Figure S1A,B).

We clustered the data with Seurat, shown as UMAP plots of recovered cells (Fig. 2E) and nuclei (Fig. 2F). In particular, we were able to identify 10 clusters of cells and nuclei. Cells and nuclei expressing high levels of *INS* gene (natural log of counts > 3) were distributed in clusters #3, while *GCG*-enriched cells/nuclei were distributed in clusters #1 (Fig. 2E, F; Additional file 3: Figure S1D-G). Interestingly, cluster #7 was enriched in low *INS*-expressing nuclei and *GCG*-expressing nuclei, but not cells (Fig. 2E, F; Additional file 3: Figure S1D,E). Specimens expressing high levels of *SST* (natural log of counts > 3) were also found in clusters #3, while cells and nuclei expressing high levels of *PPY* (natural log of counts > 1.5) were distributed in clusters #1 and #3 (Fig. 2E, F; Additional file 3: Figure S1H-K). In addition, cluster #0 in both the scRNA-seq and snRNA-seq sets comprised of cells and nuclei expressing low levels of all the endocrine cell marker genes (Fig. 2E, F; Additional file 3: Figure S1D-K). These findings suggest that the snRNA-seq protocol allows the identification of endocrine marker gene-expressing nuclei in a manner that is similar to scRNA-seq methods.

Comparison of gene expression signatures of human islet cells using scRNA-seq and snRNA-seq

To determine whether snRNA-seq would represent a reliable transcriptomic method to identify human islet cell types, we compared our snRNA-seq and scRNA-seq datasets with the publicly available scRNA-seq datasets. We chose five published scRNA-seq datasets of human pancreatic islets spanning 27 donors, four laboratories, and five technologies for harmonizing with Seurat [21,23,26,27,57]. We used this harmonized dataset as a “reference” (Fig. 3A) upon which to integrate our data, allowing for correspondence of cells and clusters. Indeed, when we compared the cultured human islet scRNA-seq and snRNA-seq datasets we had generated to the reference, we were able to identify all the major islet endocrine (i.e., α -cells, β -cells, PP-cells, δ -cells) and non-endocrine

(endothelial cells, stellate cells, Schwann cells, acinar cells, and ductal cells) cell types previously reported using established scRNA-seq methodologies, suggesting that snRNA-seq protocols did not hinder the characterization of any islet cell type, including the low abundance PP-cell population (Fig. 3A, Additional file 4: Table S3). We then undertook one-to-one comparisons of global and cell type-specific gene expression profiles between (a) the scRNA-seq and the reference datasets and (b) between the scRNA-seq and the snRNA-seq datasets we had generated. We observed an almost total overlap (99.9 %) of the genes detected between the reference (10,497 genes) and our scRNA-seq (11,694 genes) and between the scRNA-seq and the snRNA-seq (11,692 genes) datasets (Fig. 3B, C). We tested the association of gene expression levels between datasets within the major islet cell types (α -cells, β -cells, PP-cells, and δ -cells) and observed positive correlations within α -cells, β -cells, PP-cells, and δ -cells between scRNA-seq and the reference (Fig. 3D). In addition, such correlations were driven by cell type-specific genes, such as *GCG* for α -cells, *INS* for β -cells, and so forth, an aspect that was absent in all the other correlations between a given islet cell type from the scRNA-seq dataset and a different one from the reference output (Additional file 3: Figure S2A). Notably, the positive correlations of the gene signatures were also evident within α -cells, β -cells, PP-cells, and δ -cells in the scRNA-seq and snRNA-seq dataset comparisons (Fig. 3E).

We then analyzed the fractional overlap of increasing N number of top genes ($N = 100, 200, 500,$ and 1000 top genes) within cell types between the scRNA-seq and the reference and between the snRNA-seq and the scRNA-seq datasets (Fig. 3F). In particular, a significant proportion ($\sim 72\%$) of the top 100 transcripts was shared by the major islet cell types (α -cells, β -cells, and δ -cells) between the reference and the scRNA-seq datasets (Fig. 3F, Additional file 3: Figure S2B-E). Similar levels of fractional overlap of an increasing number of top genes were also observed in the scRNA-seq vs. snRNA-seq comparison (Fig. 3F). Curiously, the δ -cell groups displayed the highest overlap, reaching 95% in the top 100 and 85% in the top 1000 cell type genes (Fig. 3F). These results support the notion that snRNA-seq identifies genes expressed in the least as well as the most abundant

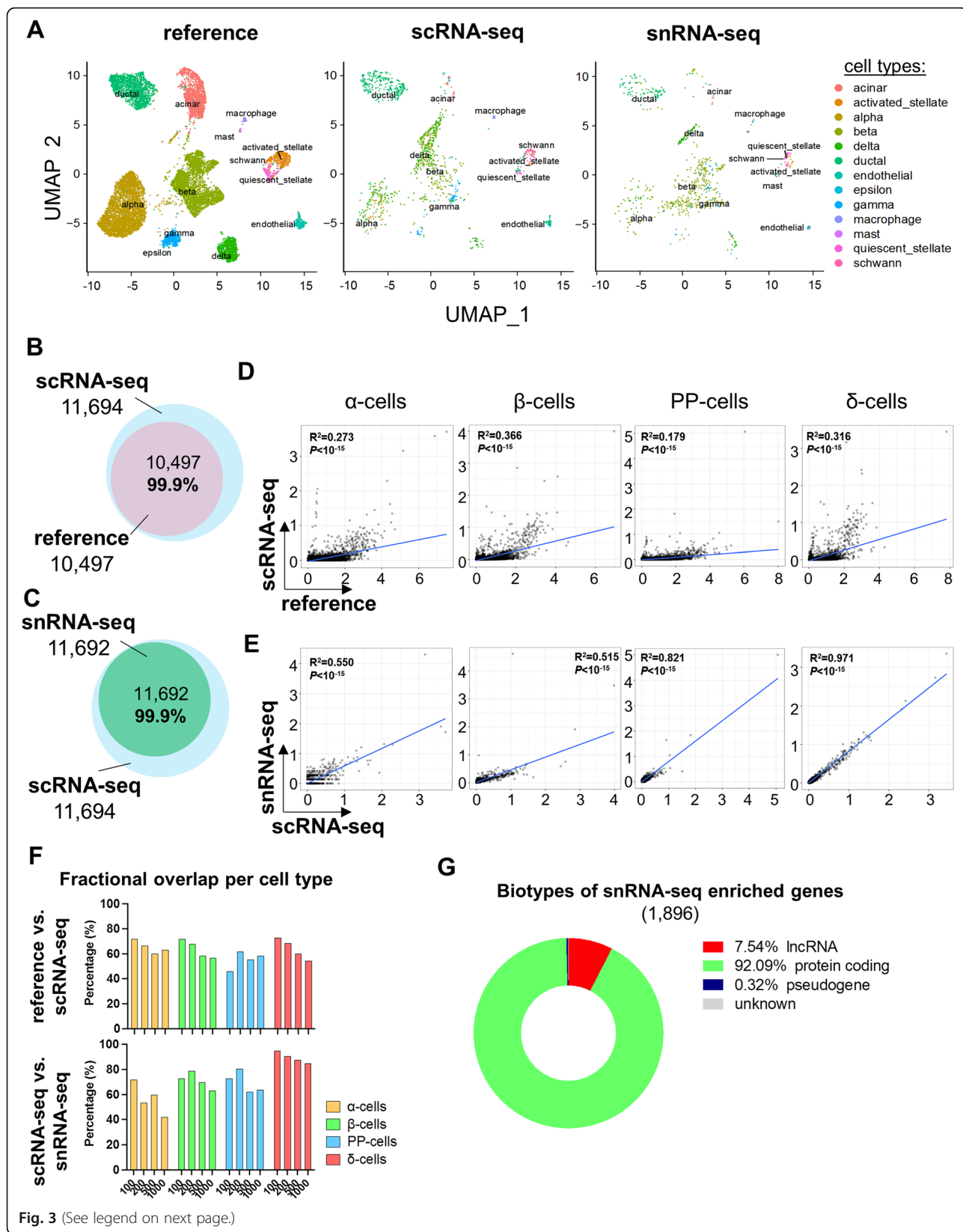


Fig. 3 (See legend on next page.)

(See figure on previous page.)

Fig. 3 Comparison between snRNA-seq and scRNA-seq datasets following harmonization on reference dataset. **A** Global UMAP plots and cell type prediction in the scRNA-seq (middle panel) and snRNA-seq (right panel) following harmonization on the reference dataset generated by integrating five published scRNA-seq datasets of cultured human islets [21, 23, 26, 27, 57] (left panel). Y-axis is similar in all three panels. **B** Venn diagram of detected genes in the reference (pink circle) and in our scRNA-seq dataset (blue circle). **C** Venn diagram of detected genes in our scRNA-seq (blue circle) and in our snRNA-seq dataset (green circle). **D, E** Scatter plots of harmonized cell type-specific gene expression in α -cells, β -cells, PP-cells, or δ -cells **D** between the reference and the scRNA-seq datasets and **E** between the scRNA-seq and the snRNA-seq datasets. X-axis and Y-axis represent the expression levels in natural log of counts in the indicated datasets. The blue line in each plot represents the regression line, whose fit is indicated by the R^2 value (the square of the Pearson correlation coefficient). The P and R values are provided for each correlation. **F** Fractional overlap expressed in percentages (%; Y-axis) of increasing numbers (100, 200, 500, and 1000; X-axis) of top genes within α -cells (yellow bars), β -cells (green bars), PP-cells (blue bars), or δ -cells (red bars) in reference vs. scRNA-seq (upper panel) and in scRNA-seq vs. snRNA-seq (lower panels) datasets following harmonization to the reference. **G** Pie chart representing proportions of biotypes of genes detected with higher confidence in snRNA-seq compared to scRNA-seq (snRNA-seq enriched genes)

endocrine cells, providing confidence that snRNA-seq is a reliable approach for identifying genes expressed in the various islet cell types.

To evaluate whether snRNA-seq would allow identification of specific types of genes which are covered with higher confidence compared to scRNA-seq, we analyzed the biotypes of the genes detected in the snRNA-seq dataset with a significantly higher (> 1.5-folds) percentage of detection compared to the scRNA-seq dataset. Among the 1896 genes which were covered in the snRNA-seq with a higher confidence compared to the scRNA-seq method, 7.5% accounted for long non-coding RNAs (lncRNAs), whereas 92.1% were protein coding genes (Fig. 3G). Importantly, lncRNA genes were also among the top genes detected in snRNA-seq with higher confidence compared to scRNA-seq protocols. Indeed, the proportions of detection of cytochrome c oxidase assembly factor heme A:farnesyltransferase (COX10) antisense RNA 1 (COX10-AS1) and minichromosome maintenance complex component 3 associated protein (MCM3AP) antisense RNA 1 (MCM3AP-AS1) genes in snRNA-seq were 15.5-fold and 13.7-fold higher than scRNA-seq, respectively (Additional file 5). In addition, nuclear paraspeckle assembly transcript 1 (NEAT1) and maternally expressed 3 (MEG3), also lncRNA genes, were detected in the snRNA-seq dataset with a proportion of 70% and 26.9%, respectively, whereas their detection rates in the scRNA-seq dataset were 35.5% and 7.0%, respectively (Additional file 5). These data suggested that the snRNA-seq method allows for detecting nuclei-enriched lncRNA genes with a higher confidence compared to scRNA-seq. The ability of the nuclear transcriptomic analysis to detect genes enriched in non-coding RNAs provides an important resource for studying the epigenetic regulatory mechanisms in human islets.

snRNA-seq of transplanted human islets

To reveal the transcriptomic signature of frozen engrafted human islets, we undertook snRNA-seq experiments as depicted in Fig. 4A. We used the immunodeficient NSG mouse model which is a widely utilized in vivo model for β -cell regeneration studies [6, 8, 10, 65]. Human islets (1000 IEQs) obtained from 4 different donors were transplanted individually under the kidney capsule of 8-to-12-week-old male

mice and followed up for 4 weeks. At the end of 4 weeks, grafts were removed carefully and divided into two equivalent parts (~ 500 IEQs each) that were subsequently snap frozen. On the day of nuclei isolation, frozen graft fractions (~ 500 IEQs) were placed in ice-cold lysis buffer individually and homogenized immediately to obtain pure nuclei as described above (Fig. 1) and single-nucleus cDNA libraries were generated by using the Chromium Single-Cell 3' Library Kit (Fig. 4A) [41]. Similar to the analyses of cultured islet snRNA-seq, the engrafted islet snRNA-seq data were initially analyzed using quality check pipelines in order to remove low-quality nuclei, including those with a total number of reads (UMI) < 1000, a total number of expressed genes < 500, and a proportion of mitochondrial genes > 20%, in line with previous reports [35, 36, 38] (Additional file 2: Table S2, Additional file 3: Figure S3A-C). It is worth noting that the percentage of expression of mitochondrial genes per nucleus was < 2% of the whole transcriptome (Additional file 2: Figure S3C), in line with the in vitro snRNA-seq results, suggesting a high efficiency of nuclear isolation method and allowing for a greater confidence to interpret and analyze the snRNA-seq data [36]. In addition, by aligning the recovered reads to the murine genome (GRCm38), we filtered out nuclei containing > 25% mouse-specific UMI (Additional file 3: Figure S3D). Finally, Cell Ranger estimated ambient RNA contamination levels at 4.6 ± 1.3 % (Additional file 2: Table S2; Additional file 3: Figure S3E) within nucleus-containing droplets. With this approach, we recovered 3565 (891.2 ± 409.6 per sample) high-quality nuclei, which is comparable with previous studies using a similar [35] or different platforms [35, 42, 66, 67].

Next, clustering with Seurat yielded 7 major clusters (Fig. 4B). Notably, we observed a cluster of nuclei enriched in α -cell-specific genes, namely cluster 1; a nuclear cluster enriched in *SST* and *PPY* genes, namely cluster 2; and a high *INS* and *MAF* bZIP transcription factor A (*MAFA*)-expressing nuclear cluster (natural log of counts ≥ 5), namely cluster 3 (Fig. 4B–D). Interestingly, other β -cell marker genes, including ATP binding cassette subfamily C member 8 (*ABCC8*) and solute carrier family 30 member 8 (*SLC30A8*), were also expressed in cluster 0, which displayed low levels of *INS* expression

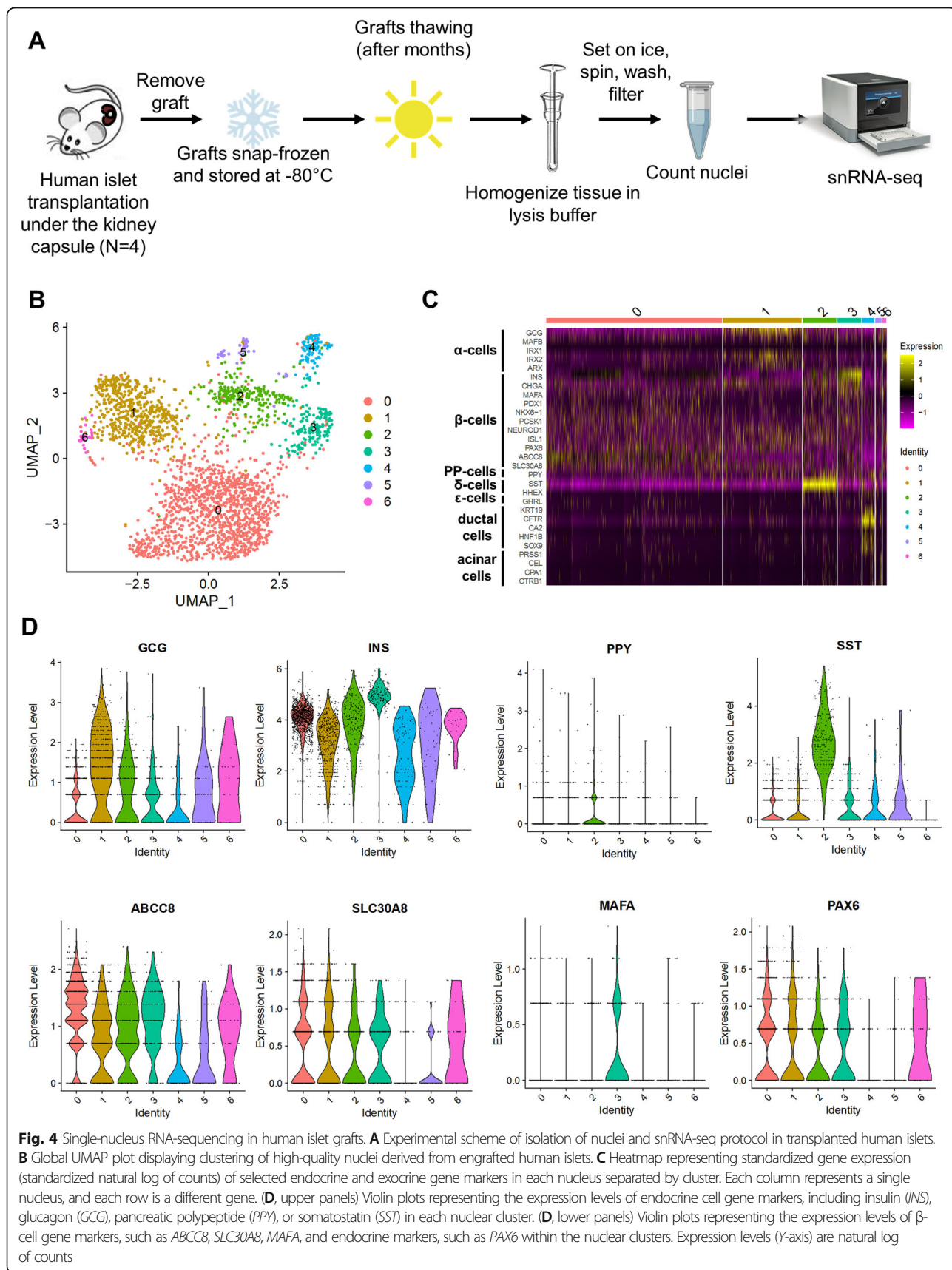


Fig. 4 Single-nucleus RNA-sequencing in human islet grafts. **A** Experimental scheme of isolation of nuclei and snRNA-seq protocol in transplanted human islets. **B** Global UMAP plot displaying clustering of high-quality nuclei derived from engrafted human islets. **C** Heatmap representing standardized gene expression (standardized natural log of counts) of selected endocrine and exocrine gene markers in each nucleus separated by cluster. Each column represents a single nucleus, and each row is a different gene. **(D, upper panels)** Violin plots representing the expression levels of endocrine cell gene markers, including insulin (*INS*), glucagon (*GCG*), pancreatic polypeptide (*PPY*), or somatostatin (*SST*) in each nuclear cluster. **(D, lower panels)** Violin plots representing the expression levels of β -cell gene markers, such as *ABCC8*, *SLC30A8*, *MAFA*, and endocrine markers, such as *PAX6* within the nuclear clusters. Expression levels (Y-axis) are natural log of counts

(natural log of counts < 5) (Fig. 4B–D). In addition, cluster 0 was enriched in endocrine cell-specific genes, such as paired box 6 (*PAX6*), suggesting that such clusters likely included β -cells in a different maturity state. However, low *INS* expression levels were also found in nuclei belonging to cluster 2 (Fig. 4B–D). Regarding the exocrine cell marker gene expression, a cluster of nuclei enriched in ductal cell genes—including CF transmembrane conductance regulator (*CFTR*), SRY-box transcription factor 9 (*SOX9*), and keratin 19 (*KRT19*), was identified in cluster 4, whereas the acinar cell-specific genes were virtually absent (Fig. 4B,C). Finally, we harmonized the in vivo snRNA-seq to the reference dataset to evaluate an unsupervised prediction of the cell type identity. Strikingly, all the endocrine and non-endocrine islet cell types were identified following harmonization, indicating the efficacy of the snRNA-seq approach to recover virtually all islet cells from transplanted samples (Additional file 4: Table S3, Additional file 3: Figure S3F). Taken together, these data indicate that snRNA-seq of human islet grafts reveals the presence of genes marking all islet cell types, and as expected, a virtual absence of exocrine cell genes.

The presence of nuclei with high *SST*-expression and low *INS*-expression in cluster 2 prompted us to undertake immunofluorescence studies to validate polyhormonal expression at the protein level. To this end, we immunostained α -cells, β -cells, and δ -cells in graft sections following transplantation of human islets from the same donors used for the in vivo transcriptomic analysis. We observed overlap between *SST* and *INS* protein immunostaining in transplanted human islets, suggesting that, indeed, cell fate transition occurs over the 4-week in vivo engraftment period, as previously observed [51, 52, 68] (Additional file 3: Figure S4).

Comparison of in vitro vs. in vivo single-nucleus RNA-sequencing methods

To determine the reliability of the snRNA-seq dataset generated from transplanted human islets, we compared the transcriptomic profiles of the snRNA-seq output obtained from cultured (in vitro) with that from transplanted (in vivo) human islets. The genes detected in each dataset showed a significant overlap (99.9%; diagram Fig. 5A). Testing the association of the islet cell type gene expression profiles revealed a positive correlation in α -cells and δ -cells (Fig. 5B, E) that was stronger in β -cells and PP-cells between in vitro and the in vivo snRNA-seq datasets (Fig. 5C, D). These data confirm the similarity of transcriptomic profiles between freshly cultured and frozen engrafted human islets and highlight the ability of snRNA-seq procedures to interrogate gene expression of human islet cells.

Discussion

Over the last few decades, the research community is increasingly utilizing human islet studies with the long-term goal of developing novel translational strategies to counteract diabetes [69]. Moreover, the establishment of avant-garde technologies for studying biological processes at the single-cell resolution have provided new tools for exploring human islet cells in physiological conditions and their defects during diabetes disease progression [25–28, 70]. While such studies were commonly conducted on cultured human islets in vitro, the transcriptome of frozen human islet cells after transplantation in mouse models remains largely unexplored and provides an opportunity to study archived tissues. Analyzing the transcriptomic signatures of human islet cells at the single-cell level represents a state-of-the-art tool for gathering insights into dynamic molecular mechanism(s) in response to diverse stimuli, including mitogens, differentiating factors, or stimulators of hormone secretion. In this context, human β -cell proliferation, differentiation/transdifferentiation, or neogenesis have all been topics of investigation in the context of physiological (e.g., pregnancy) or pathophysiological (e.g., insulin resistance/T2D) states. Nevertheless, there continues to be an urgent need for new tools to reliably study archived human or mouse tissues. Here we present single-nucleus RNA-sequencing (snRNA-seq) for interrogating the human islet transcriptome that is especially relevant for small samples that have been frozen following in vivo manipulation.

It has been established that the isolation of nuclei has several advantages over single-cell isolation [35, 67, 71]. First, this method can be simultaneously applied to multiple samples that are collected at different time points, reducing the potential variations introduced by sample handling. The second advantage is that the nucleus isolation method is efficient and requires fewer steps compared to the single-cell protocol. For example, snRNA-seq does not require enzymatic digestion to dissociate tissues into single cells which often results in decreased viability and cell loss [72]. Finally, rapid isolation of nuclei compared to tissue dissociation and isolation of single cells minimizes changes in the transcriptome during the isolation process [31]. Such considerations are important, especially when the volume of tissue available for studies is necessarily limited, such as human islet grafts.

To test the reliability of snRNA-seq in comparison to the well-established scRNA-seq procedures, we undertook a direct comparison between the two methodologies by analyzing the transcriptomics of freshly isolated cells and nuclei obtained from the same human islet donor. Despite the differences in terms of sequenced reads and detected genes per specimen between the two sequencing procedures, likely due to the lower RNA content in the nucleus compared to the cytosol [63], the

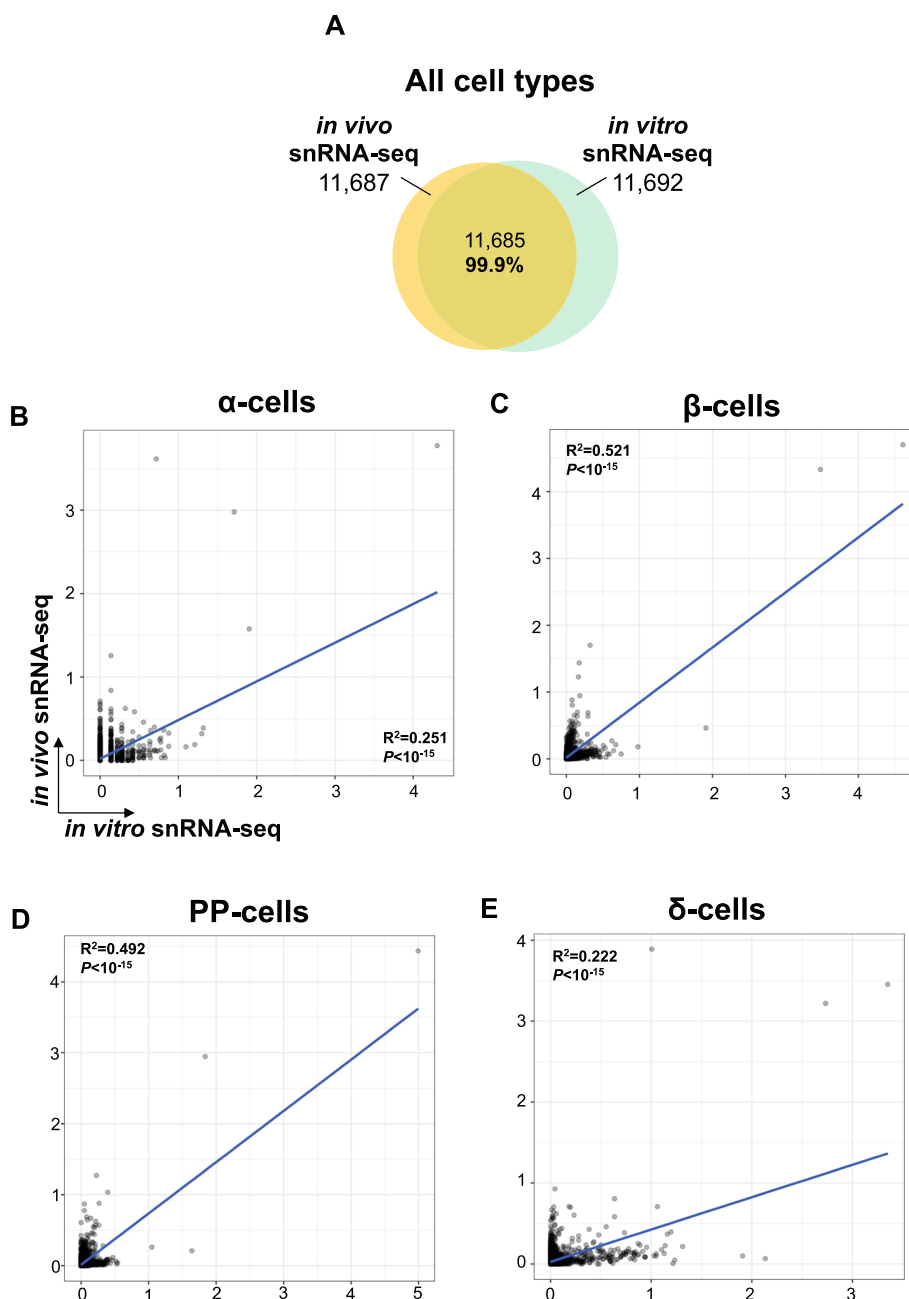


Fig. 5 Comparison between *in vitro* and *in vivo* snRNA-seq datasets. **A** Intersection of the total number of detected genes in the snRNA-seq dataset of cultured human islets (*in vitro* snRNA-seq, green circle) and the snRNA-seq dataset of transplanted human islets (*in vivo* snRNA-seq, yellow circle). **B–D** Scatter plots of harmonized cell type-specific gene expression in **B** α-cells, **C** β-cells, **D** PP-cells, and **E** δ-cells between the *in vitro* and the *in vivo* snRNA-seq datasets. X-axis and Y-axis represent the gene expression levels expressed as natural log of counts in the indicated datasets. The blue line in each plot represents the regression line, whose fit is indicated by the R^2 value (the square of the Pearson correlation coefficient). The P and R values are provided for each correlation

sequencing efficiency was similar between scRNA-seq and snRNA-seq procedures, consistent with reports on other metabolic tissues [35, 67, 71]. We also generated a reference dataset by harmonizing 5 previously published scRNA-seq datasets in human islets in order to compare our transcriptomic results in an unsupervised fashion [56]. This approach indicated that snRNA-seq allowed

for the identification of all the pancreatic islet cell types, including the least abundant, such as the PP-cells. The data also showed an almost total overlap in global gene expression (99.9%) between the two methodologies highlighted by positive correlations of gene signatures, mainly driven by cell-specific genes, between each of the islet cell types. Furthermore, the ability to identify genes

whose percentage of detection in snRNA-seq was > 1.5-fold higher than scRNA-seq indicated those genes were detected by snRNA-seq with a higher confidence in comparison to scRNA-seq. The fact that ~ 7.5% of the snRNA-seq enriched genes were lncRNAs suggested that snRNA-seq represented a potential tool for identifying non-coding RNAs in human islets that would be useful to examine chromatin remodeling, post-transcriptional modifications, and crosstalk with other RNA species.

We then applied snRNA-seq to archived human islet graft samples. As demonstrated for the cultured islet samples, single-nucleus preparations from transplanted human islets allowed the identification of all the islet cell types with a coverage that is comparable to single-cell profiling. Notably, the global and the islet cell-specific single-nucleus transcriptomics of cultured versus transplanted human islet were highly concordant, confirming that snRNA-seq represents a reliable strategy to analyze the gene signature of human islets in vivo at single-cell resolution. Finally, the transcriptome of polyhormonal islet cell clusters, e.g., cluster 2 expressing *INS* and *SST*, could be recapitulated at the protein level by immunofluorescence. Although other in situ hybridization methods, such as RNA-scope, would be more appropriate to validate transcriptomic results, the validation at the protein level provides an important functional perspective to the snRNA-seq data.

Conclusions

We propose snRNA-seq as a reliable tool to explore the transcriptomic profile of human islets especially from frozen archived samples which may not be ideal for single-cell procedures.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13073-021-00941-8>.

Additional file 1: Table S1. Characteristics of human islet donors and human islet preparations used for single-cell or single-nucleus transcriptomic analysis.

Additional file 2: Table S2. Number of different types of droplets, reads and genes generated by scRNA-seq and snRNA-seq in cultured or transplanted human islets.

Additional file 3: Figure S1. Ambient contribution and islet cell gene marker UMAP and violin plots in scRNA-seq and snRNA-seq in cultured human islets. (A,B) Box plots representing the levels of ambient RNA contamination in droplets of (A) scRNA-seq and (B) snRNA-seq datasets. (C) Violin plot representing the duplication rate in scRNA-seq (orange plot) and snRNA-seq (blue plot) methodologies. The rate was calculated by normalizing the average number of UMI per cell/nucleus on the average number of reads per cell/nucleus. (D, F, H, J) UMAP plots displaying expression levels of (D) *GCG*, (F) *INS*, (H) *PPY*, and (J) *SST* within the global distribution in scRNA-seq (left panels) and snRNA-seq (right panels). Expression levels are indicated as natural log of counts and range from 0 (gray) to 5 (purple). (E, G, I, K) Violin plots representing expression levels (natural log of counts, Y-axis) of (E) *GCG*, (G) *INS*, (I) *PPY*, and (K) *SST* within each cluster (X-axis) identified in scRNA-seq (orange plots) or snRNA-seq

(blue plots) datasets. **Figure S2.** Correlation plots and fractional overlap estimation of islet cell types from scRNA-seq dataset with those from the reference dataset. (A) Scatter plots representing correlation of gene expression levels between scRNA-seq-derived (Y-axis) α -cells (first row from top), β -cells (second row from top), PP-cells (third row from top), and δ -cells (fourth row from top) and reference-derived (X-axis) α -cells (first column from left), β -cells (second column from left), PP-cells (third column from left) or δ -cells (fourth column from left). X-axis and Y-axis represent the expression levels in natural log of counts in the indicated datasets. The blue line in each plot represents the regression line, whose fit is indicated by the R^2 value (the square of the Pearson correlation coefficient). The red circles indicate the islet cell specific marker genes driving the correlation between the same cell type from the two datasets. The red dotted squares highlight the correlation plots used in the main Fig. 3D. (B-E) Fractional overlap expressed in percentages (%), Y-axis) of the (B) top 100, (C) 200, (D) 500, or (E) 1000 genes between the indicated islet cell types from the reference dataset (X-axis) and the α -cells (yellow bars), β -cells (green bars), PP-cells (blue bars), or δ -cells (red bars) from the scRNA-seq dataset. **Figure S3.** Quality check parameters in the transplanted human islet snRNA-seq dataset. (A-D) Violin plots of (A) number of genes, (B) number of reads, (C) proportion of mitochondrial genes, and (D) proportion of mouse genes per nucleus across the 4 transplanted human islet samples in the snRNA-seq dataset. (E) Box plot of the levels of ambient RNA contamination in droplets of the in vivo snRNA-seq dataset within each human islet graft. (F) UMAP plot of nuclear cluster distribution and cell type prediction of in vivo snRNA-seq dataset following harmonization to the reference dataset. **Figure S4.** Islet cell type validation in human islet graft sections by immunofluorescence. Representative images of human islet cells identified as α -cells, β -cells or δ -cells according to the glucagon (GCG, green), insulin (INS, red) and somatostatin (SST, white) labeling. Nuclei are stained in blue. Scale bar is: 50 μ m.

Additional file 4: Table S3. Number of endocrine and exocrine pancreatic cell types predicted in scRNA-seq or snRNA-seq of cultured or transplanted human islets following harmonization.

Additional file 5. List of top 20 snRNA-seq enriched genes ordered by fold change of percentage of detection (upper table) or percentage of detection (lower table) compared to scRNA-seq.

Acknowledgements

We thank Michael DeRan PhD (Broad Institute), Qiong Zhou, and Mary-Elizabeth Patti MD (Joslin Genomics Core) for technical assistance with performing 10X Genomics experiments and Bridget Wagner PhD (Broad Institute) and Amit Chaudhary PhD (Broad Institute) for advice. We thank Christopher Cahill (Joslin Advanced Microscopy Core) for assistance with imaging experiments.

Authors' contributions

G.B., S.K., and E.D. designed the study, performed the experiments, and analyzed the data. G.B. and S.K. wrote the manuscript. H.P. and J.M.D. performed the bioinformatics analyses. R.N.K. designed the study, directed the project and edited the manuscript. All authors edited, read, and approved the final manuscript.

Funding

R.N.K. acknowledges support from NIH Grants RO1 DK067536, RO1 DK117639, RO1 DK105588, and UC4 DK 116278 and the Joslin DRC Cores including the Molecular Phenotyping and Genotyping, the Advanced Microscopy and Bioinformatics Cores (P30 DK 36836).

Availability of data and materials

All raw and processed snRNA-seq data generated from engrafted human islet tissues have been submitted to the NCBI Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE150212>) under accession number GSE150212. The publicly available scRNA-seq datasets on human islets that were used in this the manuscript to integrate our scRNA-seq and snRNA-seq datasets have been deposited in the public Genomic Spatial Event Database (GSE) and in the ArrayExpress Database at the EMBL-EBI. For InDrop, accession number GSE84133 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE84133>) [21]; for CelSeq2, accession number GSE85241 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=>

GSE85241 [23]; for SMART-Seq2, accession number E-MTAB-5061 (<https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-5061/>) [26]; for Fluidigm C1, accession number GSE86469 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE86469>) (27); and for CelSeq, accession number GSE81076 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE81076>) [57] were used.

Declarations

Ethics approval and consent to participate

Ethical approval for the usage of human islets was granted by the Ethics Committee of the Joslin Diabetes Center for Human Studies (#05-05). The study complied with all relevant ethical regulations for work with human cells for research purposes. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements. The human islets that are distributed by the IIDP are from approved cadaveric organ donors from which at least one other organ has been approved for transplantation. Because the donors are brain dead, the IRB's from the institutions that isolate the islets consider the tissue as "Exempt" from Human Studies Approval.

Animal studies and protocols were approved by the Institutional Animal Care and Use Committees of the Joslin Diabetes Center (IACUC #05-01, #2012-09). Animal studies and protocols were approved by the Institutional Animal Care and Use Committees of the Joslin Diabetes Center (IACUC #05-01, #2012-09).

All research performed in this study followed the Declaration of the Helsinki Ethical Principles for Medical Research Involving Human Samples.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Section of Islet Cell and Regenerative Biology, Joslin Diabetes Center and Harvard Medical School, Boston, MA 02215, USA. ²Current Address: Department of Pharmacology, New York Medical College School of Medicine, Valhalla, NY 10595, USA. ³Bioinformatics and Biostatistics Core, Joslin Diabetes Center and Harvard Medical School, Boston, MA, USA. ⁴Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA. ⁵Harvard Stem Cell Institute, Harvard Medical School, Boston, MA, USA.

Received: 17 July 2020 Accepted: 13 July 2021

Published online: 10 August 2021

References

- Bluestone JA, Herold K, Eisenbarth G. Genetics, pathogenesis and clinical interventions in type 1 diabetes. *Nature*. 2010;464(7293):1293–300. <https://doi.org/10.1038/nature08933>.
- Butler AE, Janson J, Bonner-Weir S, Ritzel R, Rizza RA, Butler PC. β -cell deficit and increased β -cell apoptosis in humans with type 2 diabetes. *Diabetes*. 2003;52:102–10.
- Shirakawa J, Kulkarni RN. Novel factors modulating human β -cell proliferation. *Diabetes Obes Metab*. 2016;18(Suppl 1):71–7. <https://doi.org/10.1111/dom.12731>.
- Benthuisen JR, Carrano AC, Sander M. Advances in β cell replacement and regeneration strategies for treating diabetes. *J Clin Invest*. 2016;126(10):3651–60. <https://doi.org/10.1172/JCI87439>.
- Basile G, Kulkarni RN, Morgan NG. How, When, and Where Do Human β -Cells Regenerate? *Curr Diab Rep*. 2019;19(48). <https://doi.org/10.1007/s11892-019-1176-8>.
- El Ouaamari A, Dirice E, Gedeon N, Hu J, Zhou JY, Shirakawa J, et al. SerpinB1 Promotes Pancreatic beta Cell Proliferation. *Cell Metab*. 2016;23(1):194–205. <https://doi.org/10.1016/j.cmet.2015.12.001>.
- Dhawan S, Dirice E, Kulkarni RN, Bhushan A. Inhibition of TGF- β signaling promotes human pancreatic β -cell replication. *Diabetes*. 2016;65(5):1208–18. <https://doi.org/10.2337/db15-1331>.
- Dirice E, Walpita D, Vetere A, Meier BC, Kahraman S, Hu J, et al. Inhibition of DYRK1A stimulates human beta-cell proliferation. *Diabetes*. 2016;65(6):1660–71. <https://doi.org/10.2337/db15-1127>.
- Wang P, Alvarez-Perez JC, Felsenfeld DP, Liu H, Sivendran S, Bender A, et al. A high-throughput chemical screen reveals that harmine-mediated inhibition of DYRK1A increases human pancreatic beta cell replication. *Nat Med*. 2015;21(4):383–8. <https://doi.org/10.1038/nm.3820>.
- Wang P, Karakose E, Liu H, Swartz E, Acekifi C, Zlatanic V, et al. Combined Inhibition of DYRK1A, SMAD, and Trithorax Pathways Synergizes to Induce Robust Replication in Adult Human Beta Cells. *Cell Metab*. 2019;29(3):638–52. <https://doi.org/10.1016/j.cmet.2018.12.005>.
- Acekifi C, Wang P, Karakose E, Manning Fox JE, González BJ, Liu H, et al. GLP-1 receptor agonists synergize with DYRK1A inhibitors to potentiate functional human β cell regeneration. *Sci Transl Med*. 2020;12(530).
- Zhu S, Russ HA, Wang X, Zhang M, Ma T, Xu T, et al. Human pancreatic beta-like cells converted from fibroblasts. *Nat Commun*. 2016;7(1):10080. <https://doi.org/10.1038/ncomms10080>.
- Rezania A, Bruin JE, Arora P, Rubin A, Batushansky I, Asadi A, et al. Reversal of diabetes with insulin-producing cells derived in vitro from human pluripotent stem cells. *Nat Biotechnol*. 2014;32(11):1121–33. <https://doi.org/10.1038/nbt.3033>.
- 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(2):428–39. <https://doi.org/10.1016/j.cell.2014.09.040>.
- Millman JR, Xie C, Van Dervort A, Gürtler M, Pagliuca FW, Melton DA. Generation of stem cell-derived β -cells from patients with type 1 diabetes. *Nat Commun*. 2016;7(1):11463. <https://doi.org/10.1038/ncomms11463>.
- Nair GG, Liu JS, Russ HA, Tran S, Saxton MS, Chen R, et al. Recapitulating endocrine cell clustering in culture promotes maturation of human stem-cell-derived β cells. *Nat Cell Biol*. 2019;21(2):263–74. <https://doi.org/10.1038/s41556-018-0271-4>.
- Velazco-Cruz L, Song J, Maxwell KG, Goedegebuure MM, Augsornworawat P, Högrebé NJ, et al. Acquisition of Dynamic Function in Human Stem Cell-Derived β Cells. *Stem Cell Rep*. 2019;12(2):351–65. <https://doi.org/10.1016/j.stemcr.2018.12.012>.
- Mawla AM, Huisig MO. Navigating the depths and avoiding the shallows of pancreatic islet cell transcriptomes. *Diabetes*. 2019;68(7):1380–93. <https://doi.org/10.2337/db18-0019>.
- 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(5):1194–1205.e4.
- 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(5):1160–1175.e11.
- 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(4):346–60. <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE84133>. <https://doi.org/10.1016/j.cels.2016.08.011>.
- 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(2):178–87. <https://doi.org/10.15252/embr.201540946>.
- 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(4):385–94. <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE85241>. <https://doi.org/10.1016/j.cels.2016.09.002>.
- Marquina-Sanchez B, Fortelny N, Farlik M, et al. Single-cell RNA-seq with spike-in cells enables accurate quantification of cell-specific drug effects in pancreatic islets. *Genome Biol*. 2020;21(106). <https://doi.org/10.1186/s13059-020-02006-2>.
- 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(2):321–330.e14.
- 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(4):593–607. <https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-5061>. <https://doi.org/10.1016/j.cmet.2016.08.020>.
- 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(2):208–22. <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE86469>. <https://doi.org/10.1101/gr.212720.116>.
- 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;3(4):608–15. <https://doi.org/10.1016/j.cmet.2016.08.018>.

29. 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(10):3028–38. <https://doi.org/10.2337/db16-0405>.
30. Grindberg RV, Yee-Greenbaum JL, McConnell MJ, Novotny M, O'Shaughnessy AL, Lambert GM, et al. RNA-sequencing from single nuclei. *Proc Natl Acad Sci*. 2013;110(49):19802–7. <https://doi.org/10.1073/pnas.1319700110>.
31. Adam M, Potter AS, Potter SS. Psychrophilic proteases dramatically reduce single-cell RNA-seq artifacts: A molecular atlas of kidney development. *Dev Camb*. 2017;144:3625–32.
32. Bonnycastle LL, Gildea DE, Yan T, Narisu N, Swift AJ, Wolfsberg TG, et al. Single-cell transcriptomics from human pancreatic islets: Sample preparation matters. *Biol Methods Protoc*. 2019;4(1). <https://doi.org/10.1093/biometods/bpz019>.
33. Al-Dalahmah O, Sosunov AA, Shaik A, Ofori K, Liu Y, Vonsattel JP, et al. Single-nucleus RNA-seq identifies Huntington disease astrocyte states. *Acta Neuropathol Commun*. 2020;18:19.
34. Kebschull JM, Richman EB, Ringach N, Friedmann D, Albarran E, Kolluru SS, et al. Cerebellar nuclei evolved by repeatedly duplicating a conserved cell-type set. *Science*. 2020;370(6523).
35. Wu H, Kirita Y, Donnelly EL, Humphreys BD. Advantages of Single-Nucleus over Single-Cell RNA Sequencing of Adult Kidney: Rare Cell Types and Novel Cell States Revealed in Fibrosis. *J Am Soc Nephrol*. 2019;30(1):23–32. <https://doi.org/10.1681/ASN.2018090912>.
36. Lake BB, Chen S, Hoshi M, Plongthongkum N, Salamon D, Knoten A, et al. A single-nucleus RNA-sequencing pipeline to decipher the molecular anatomy and pathophysiology of human kidneys. *Nat Commun*. 2019;10(1):2832. <https://doi.org/10.1038/s41467-019-10861-2>.
37. Muto Y, Wilson PC, Ledru N, Wu H, Dimke H, Waikar SS, et al. Single cell transcriptional and chromatin accessibility profiling redefine cellular heterogeneity in the adult human kidney. *Nat Commun*. 2021;12(1):2190. <https://doi.org/10.1038/s41467-021-22368-w>.
38. Hu P, Liu J, Zhao J, Wilkins BJ, Lupino K, Wu H, et al. Single-nucleus transcriptomic survey of cell diversity and functional maturation in postnatal mammalian hearts. *Genes Dev*. 2018;32(19–20):1344–57. <https://doi.org/10.1101/gad.316802.118>.
39. Kim M, Franke V, Brandt B, Lowenstein ED, Schöwel V, Spuler S, et al. Single-nucleus transcriptomics reveals functional compartmentalization in synctial skeletal muscle cells. *Nat Commun*. 2020;11(1):6375. <https://doi.org/10.1038/s41467-020-20064-9>.
40. Petrany MJ, Swoboda CO, Sun C, Chetal K, Chen X, Weirauch MT, et al. Single-nucleus RNA-seq identifies transcriptional heterogeneity in multinucleated skeletal myofibers. *Nat Commun*. 2020;11(1):6374. <https://doi.org/10.1038/s41467-020-20063-w>.
41. Korrapati S, Taukulis I, Olszewski R, Pyle M, Gu S, Singh R, et al. Single Cell and Single Nucleus RNA-Seq Reveal Cellular Heterogeneity and Homeostatic Regulatory Networks in Adult Mouse Stria Vascularis. *Front Mol Neurosci*. 2019;12:316. <https://doi.org/10.3389/fnmol.2019.00316>.
42. Liang Q, Dharmat R, Owen L, Shakoor A, Li Y, Kim S, et al. Single-nuclei RNA-seq on human retinal tissue provides improved transcriptome profiling. *Nat Commun*. 2019;10(1):5743. <https://doi.org/10.1038/s41467-019-12917-9>.
43. Andrews TS, Atif J, Liu JC, Perciani CT, Ma X-Z, Thoeni C, et al. Single Cell, Single Nucleus and Spatial RNA Sequencing of the Human Liver Identifies Hepatic Stellate Cell and Cholangiocyte Heterogeneity. *bioRxiv*. 2021; Available from: <https://www.biorxiv.org/content/early/2021/03/28/2021.03.27.436882>.
44. Diamanti K, Inda Díaz JS, Raine A, Pan G, Wadelius C, Cavalli M. Single nucleus transcriptomics data integration recapitulates the major cell types in human liver. *Hepatal Res*. 2021;51(2):233–8. <https://doi.org/10.1111/hepr.13585>.
45. Wang A, Chiou J, Poirion OB, Buchanan J, Valdez MJ, Verheyden JM, et al. Single-cell multiomic profiling of human lungs reveals cell-type-specific and age-dynamic control of SARS-CoV2 host genes. *Morrissey EE, De Langhe S, editors. eLife*. 2020;9:e62522.
46. Sárvari AK, Van Hauwaert EL, Markussen LK, Gammelmark E, Marcher A-B, Ebbesen MF, et al. Plasticity of Epididymal Adipose Tissue in Response to Diet-Induced Obesity at Single-Nucleus Resolution. *Cell Metab*. 2021;33(2):437–453.e5.
47. Khalyfa A, Warren W, Andrade J, Bottoms CA, Rice ES, Cortese R, et al. Transcriptomic Changes of Murine Visceral Fat Exposed to Intermittent Hypoxia at Single Cell Resolution. *Int J Mol Sci*. 2021;22(1) Available from: <https://www.mdpi.com/1422-0067/22/1/261>.
48. Rajbhandari P, Arneson D, Hart SK, Ahn IS, Diamante G, Santos LC, et al. Single cell analysis reveals immune cell-adipocyte crosstalk regulating the transcription of thermogenic adipocytes. *eLife*. 2019;8:e49501.
49. Sun W, Dong H, Balaz M, Slyper M, Drokhylyansky E, Colleluori G, et al. snRNA-seq reveals a subpopulation of adipocytes that regulates thermogenesis. *Nature*. 2020;587(7832):98–102. <https://doi.org/10.1038/s41586-020-2856-x>.
50. Brissova M, Niland JC, Cravens J, Olack B, Sowinski J, Evans-Molina C. The Integrated Islet Distribution Program Answers the Call for Improved Human Islet Phenotyping and Reporting of Human Islet Characteristics in Research Articles. *Diabetes*. 2019;68(7):1363–5. <https://doi.org/10.2337/dbi19-0019>.
51. Dirice E, De Jesus DF, Kahraman S, Basile G, Ng RWS, El Ouamari A, et al. Human duct cells contribute to β cell compensation in insulin resistance. *JCI Insight*. 2019;4:99576.
52. Mezza T, Muscogiuri G, Sorice GP, Clemente G, Hu J, Pontecorvi A, et al. Insulin resistance alters islet morphology in nondiabetic humans. *Diabetes*. 2014;63(3):994–1007. <https://doi.org/10.2337/db13-1013>.
53. Fleming SJ, Marioni JC, Babadi M. CellBender remove-background: a deep generative model for unsupervised removal of background noise from scRNA-seq datasets. *bioRxiv*. 2019; Available from: <https://www.biorxiv.org/content/early/2019/10/03/791699>.
54. Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol*. 2019;20(1):296. <https://doi.org/10.1186/s13059-019-1874-1>.
55. McGinnis CS, Murrow LM, Gartner ZJ. DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst*. 2019;8(4):329–337.e4.
56. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM III, et al. Comprehensive Integration of Single-Cell Data. *Cell*. 2019;177(7):1888–1902.e21.
57. Grün D, Muraro MJ, Boisset J-C, Wiebrands K, Lyubimova A, Dharmadhikari G, et al. De Novo Prediction of Stem Cell Identity using Single-Cell Transcriptome Data. *Cell Stem Cell*. 2016;19(2):266–77. <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE81076>. <https://doi.org/10.1016/j.stem.2016.05.010>.
58. Narayanan A, Blanco-Carmona E, Demirdizen E, Sun X, Herold-Mende C, Schlesner M, et al. Nuclei Isolation from Fresh Frozen Brain Tumors for Single-Nucleus RNA-seq and ATAC-seq. *J Vis Exp*. 2020;162:e61542.
59. Slyper M, Porter CBM, Ashenberg O, Waldman J, Drokhylyansky E, Wakiro I, et al. A single-cell and single-nucleus RNA-Seq toolbox for fresh and frozen human tumors. *Nat Med*. 2020;26(5):792–802. <https://doi.org/10.1038/s41591-020-0844-1>.
60. Krishnaswami SR, Grindberg RV, Novotny M, Venepally P, Lacar B, Bhutani K, et al. Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons. *Nat Protoc*. 2016;11(3):499–524. <https://doi.org/10.1038/nprot.2016.015>.
61. Pisanía A, Weir GC, O'Neil JJ, Omer A, Tchpashvili V, Lei J, et al. Quantitative analysis of cell composition and purity of human pancreatic islet preparations. *Lab Invest*. 2010;90(11):1661–75. <https://doi.org/10.1038/labinvest.2010.124>.
62. Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M, et al. The RIN: An RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol Biol*. 2006;7(1):3. <https://doi.org/10.1186/1471-2199-7-3>.
63. Piwnicka M, Darzynkiewicz Z, Melamed MR. RNA and DNA content of isolated cell nuclei measured by multiparameter flow cytometry. *Cytometry*. 1983;3(4):269–75. <https://doi.org/10.1002/cyto.990030407>.
64. Han F, Lillard SJ. In-situ sampling and separation of RNA from individual mammalian cells. *Anal Chem*. 2000;72(17):4073–9. <https://doi.org/10.1021/ac000428g>.
65. Greiner DL, Brehm MA, Hosur V, Harlan DM, Powers AC, Shultz LD. Humanized mice for the study of type 1 and type 2 diabetes. *Ann N Y Acad Sci*. 2011;1245(1):55–8. <https://doi.org/10.1111/j.1749-6632.2011.06318.x>.
66. Habib N, Li Y, Heidenreich M, Swiech L, Avraham-Davidi I, Trombetta JJ, et al. Div-Seq: Single-nucleus RNA-Seq reveals dynamics of rare adult newborn neurons. *Science*. 2016;353(6302):925–8. <https://doi.org/10.1126/science.aad7038>.
67. Bakken TE, Hodge RD, Miller JA, Yao Z, Nguyen TN, Aevermann B, et al. Single-nucleus and single-cell transcriptomes compared in matched cortical cell types. *PLoS ONE*. 2018;13(12):1–24.
68. Mezza T, Cinti F, Cefalo CMA, Pontecorvi A, Kulkarni RN, Giacari A. B-cell fate in human insulin resistance and type 2 diabetes: A perspective on islet plasticity. *Diabetes*. 2019;68(6):1121–9. <https://doi.org/10.2337/db18-0856>.

69. Hart NJ, Powers AC. Use of human islets to understand islet biology and diabetes: progress, challenges and suggestions. *Diabetologia*. 2019;62(2): 212–22. <https://doi.org/10.1007/s00125-018-4772-2>.
70. Russell MA, Redick SD, Blodgett DM, Richardson SJ, Leete P, Krogvold L, et al. HLA class II antigen processing and presentation pathway components demonstrated by transcriptome and protein analyses of islet β -cells from donors with type 1 diabetes. *Diabetes*. 2019;68(5):988–1001. <https://doi.org/10.2337/db18-0686>.
71. Lake BB, Codeluppi S, Yung YC, Gao D, Chun J, Kharchenko PV, et al. A comparative strategy for single-nucleus and single-cell transcriptomes confirms accuracy in predicted cell-type expression from nuclear RNA. *Sci Rep*. 2017;7(1):1–8.
72. Reichard A, Asosingh K. Best Practices for Preparing a Single Cell Suspension from Solid Tissues for Flow Cytometry. *Cytometry A*. 2019;95(2):219–26. <https://doi.org/10.1002/cyto.a.23690>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

