

UC Davis

UC Davis Previously Published Works

Title

Navigating the Depths and Avoiding the Shallows of Pancreatic Islet Cell Transcriptomes

Permalink

<https://escholarship.org/uc/item/3109p6zc>

Journal

Diabetes, 68(7)

ISSN

0012-1797

Authors

Mawla, Alex M
Huising, Mark O

Publication Date

2019-07-01

DOI

10.2337/dbi18-0019

Peer reviewed



Navigating the Depths and Avoiding the Shallows of Pancreatic Islet Cell Transcriptomes

Alex M. Mawla¹ and Mark O. Huising^{1,2}

Diabetes 2019;68:1380–1393 | <https://doi.org/10.2337/dbi18-0019>

Islet gene expression has been widely studied to better understand the transcriptional features that define a healthy β -cell. Transcriptomes of FACS-purified α -, β -, and δ -cells using bulk RNA-sequencing have facilitated our understanding of the complex network of cross talk between islet cells and its effects on β -cell function. However, these approaches were by design not intended to resolve heterogeneity between individual cells. Several recent studies used single-cell RNA sequencing (scRNA-Seq) to report considerable heterogeneity within mouse and human β -cells. In this Perspective, we assess how this newfound ability to assess gene expression at single-cell resolution has enhanced our understanding of β -cell heterogeneity. We conduct a comprehensive assessment of several single human β -cell transcriptome data sets and ask if the heterogeneity reported by these studies showed overlap and concurred with previously known examples of β -cell heterogeneity. We also illustrate the impact of the inevitable limitations of working at or below the limit of detection of gene expression at single cell resolution and their consequences for the quality of single-islet cell transcriptome data. Finally, we offer some guidance on when to opt for scRNA-Seq and when bulk sequencing approaches may be better suited.

Type 1 diabetes (T1D) and type 2 diabetes (T2D) affect roughly 14% of the population and are the seventh leading causes of death in the U.S. (1). T1D is characterized by autoimmune-mediated β -cell destruction within the pancreas. T2D is characterized by increased peripheral insulin resistance, which eventually unmasks and/or precipitates β -cell dysfunction (2). Consequently, the field has mostly focused on β -cells, despite the fact that pancreatic islets of

Langerhans contain at least five different hormone-secreting endocrine cell types, supported by a constellation of auxiliary cells, whose clustering supports coordinated secretion of insulin and glucagon to maintain nutrient homeostasis (3–5). The spatial distribution of these cells within islets varies between human and mouse models, but β -cells are the most abundant endocrine cell type in both species, followed by α -cells, δ -cells, and a lower number of γ -pancreatic polypeptide cells and ϵ -cells (6,7).

While islet isolation is a routine procedure, the close association of all of these endocrine and auxiliary cell types within the islet has long complicated the isolation and purification of homogeneous populations of each islet cell type. Consequently, changes in gene and protein expression within intact isolated islets were often attributed to β -cells, as they are numerically the most abundant islet cell type within the islet. Clearly, this ignores the fact that multiple additional endocrine cells, as well as endothelial cells, macrophages, glia, fibroblasts, and pericytes collectively make up the pancreatic islet (8–11). β -Cell dysregulation and dysfunction are a prominent factor in disrupted insulin secretion and blood glucose control, but major functional and transcriptional changes also occur in α -cells (12,13), as well as vasculature (14), that are difficult to detect or distinguish from changes to β -cells at the level of the intact islet.

RESOLVING DIFFERENCES BETWEEN ISLET ENDOCRINE CELLS

Purification of β -cells had initially been achieved on the basis of autofluorescence (15), an approach that works reasonably well. Subsequent strategies have improved this approach by generating transgenic reporter lines that express fluorescent markers such as GFP or mCherry

¹Department of Neurobiology, Physiology and Behavior, College of Biological Sciences, University of California, Davis, Davis, CA

²Department of Physiology and Membrane Biology, School of Medicine, University of California, Davis, Davis, CA

Corresponding author: Mark O. Huising, mhuising@ucdavis.edu

Received 27 October 2018 and accepted 29 April 2019

This article contains Supplementary Data online at <http://diabetes.diabetesjournals.org/lookup/suppl/doi:10.2337/dbi18-0019/-/DC1>.

© 2019 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. More information is available at <http://www.diabetesjournals.org/content/license>.

specifically in β -cells (16,17). However, neither strategy can copurify pure α - or δ -cells. Several groups have recently resolved this limitation by generating combinations of transgenic reporter mice that made it possible to isolate pure populations of α -, β -, and δ -cells from the same islet by FACS. This has enabled the generation of comprehensive transcriptomes of FACS-purified pools of mouse α -, β -, and δ -cells with >99% purity (17–19). For human islets, the problem of purifying α - and β -cells was resolved independently by the generation of a panel of antibodies that enabled the purification of human α - and β -cells with approximately 90% purity (20–22). The ability to purify human islet cell types has allowed for further exploration in human islet transcriptomics and the subsequent identification of genes that encode proteins exclusively expressed in β -cells (23,24). However, cell-surface markers are currently unable to isolate human δ -cells or other, more rare islet endocrine cells with reasonable purity by flow cytometry.

PREVIOUSLY ESTABLISHED HETEROGENEITY

In addition to the heterogeneity that results from the clustering of many different cell types within a functional islet, it has long been evident that considerable heterogeneity exists within the β -cell population (21,25–29), and likely within non- β populations of islet cells as well. Functional heterogeneity among β -cells occurs with regard to the glucose threshold and insulin secretory response of individual β -cells (25,26,30). Heterogeneity in the expression of a number of markers, such as the peptide hormone neuropeptide Y (NPY), tyrosine hydroxylase (TH), and Dickkopf-3, by individual β -cells has also been reported (31–34).

More recently, a series of articles have rekindled interest in β -cell heterogeneity, with the description of Flattp (Fltp)-expressing β -cells (27), ST8SIA1/CD9-positive β -cells (21), Ucn3/Glut2-negative “virgin” β -cells (35,36), “bottom” β -cells (named for the bottom of two FACS gates used to isolate them [37]), and senescent β -cells (38). This paints a landscape of β -cell heterogeneity that features changes in marker expression over the life span of the β -cell and/or in relation to the functional state of the β -cell in health and disease. Understanding of this heterogeneity would benefit greatly from transcriptional read-outs at single-cell resolution. Indeed, a number of recent articles have reported on single-cell transcriptomes of mouse and human primary islet cells (39–49).

The great promise of sequencing at single-cell resolution is that this should resolve the considerable heterogeneity that exists among the individual β -cells that come together in the islet. Here, we take stock of what these recent single-cell studies have added to our understanding of islet cell biology. We do so by asking two basic questions: 1) Have individual single-cell sequencing studies that are similar in design resulted in comparable outcomes? 2) Did single-cell approaches recapitulate well-known and validated examples of β -cell heterogeneity? In addressing

these two straightforward questions, we discuss areas where single-cell approaches have made clear and tangible contributions to our field. However, we also document examples where single-cell sequencing approaches may fall short of the unrealistically high expectations that exist for this approach. We review and clarify some of the underlying reasons that may have contributed to this disconnect. Finally, we offer some guidance on when a single-cell approach is preferred and what question may be better resolved using a bulk RNA sequencing (RNA-Seq) approach.

VALIDATION OF NOVEL β -CELL HETEROGENEITY IDENTIFIED IN SINGLE-CELL RNA-Seq STUDIES

As a first step in assessing the reported heterogeneity by recent single-cell (sc)RNA-Seq studies of islets, we compared β -cell heterogeneity that was highlighted by the authors of several recent scRNA-Seq studies of human pancreatic islets (43,44,46–48). The overall design for each of these studies was to sequence dissociated islet cells of human subjects at single-cell resolution, even though each of these studies inevitably differed in the technical details and the sequencing methodologies that were used (Supplementary Table 1). Nevertheless, given the agreement in the overall design, we reasoned that true heterogeneity should emerge despite the inevitable variations in methodologies and should be reproducible across individual human donors in each of these studies. After all, if this were not true, all observations that have emerged from scRNA-Seq studies of human islets to date would be limited only to the deceased islet donors who were the subject of these studies and would not extend to the general population.

To our surprise, not a single gene was highlighted after manual annotation by the authors as heterogeneously expressed across all five studies, and only a few genes were highlighted independently by up to three scRNA-Seq studies of human β -cells (Supplementary Fig. 1A and Supplementary Table 2). This observation can be interpreted in two possible ways. It may be that the extent of β -cell heterogeneity is so great that the human β -cell scRNA-Seq studies to date have effectively undersampled this heterogeneity. The alternative explanation is that the detection of variation in gene expression across single β -cells is dominated by noise resulting from operating at or below the limit of detection of gene expression in single-cell expression, causing false negatives to dominate the list of heterogeneously detected β -cell genes. Moreover, the short list of heterogeneously detected genes in β -cells was notably lacking genes encoding proteins known to demonstrate heterogeneous expression patterns among β -cells (e.g., NPY, TH, UCN3, DKK3). This raises the question whether scRNA-Seq approaches were able to accurately detect expression of established markers of heterogeneity among β -cells. While many of these studies have taken their analyses beyond single-gene transcriptomes, i.e., gene set enrichment and multiparametric pathway

analyses, our primary focus was to evaluate whether heterogeneity could accurately be recapitulated.

VALIDATION OF NOVEL HETEROGENEITY IN β -CELLS

To determine the degree with which different scRNA-Seq studies detect overlap in β -cell heterogeneity, we conducted a meta-analysis of the five human scRNA-Seq studies. Given that the overall design of each of these studies was essentially the same and that differences between these studies were limited to the inevitable variation across human donors and variations in the sequencing methods and analyses pipelines, we expected to observe considerable overlap between each data set (Supplementary Table 1). To further reduce any differences, we downloaded and reanalyzed the raw data from each of the studies, generating an integrated analysis that resolved each of the major pancreas populations (Fig. 1A–L and Approach & Tools in Supplementary Data). We also verified that clustering was not driven by the platform used or by donor (Supplementary Fig. 2A and B).

We identified two subclusters within the β -cell population (Fig. 1A). By differential testing we identified 52 genes that drove variation between the two subpopulations ($P \leq 0.05$; Supplementary Table 3). Notably, *G6PC2*, *MAFA*, and *NPY* were detected within this list, as were several non- β endocrine and acinar cell markers, such as *GCG*, *SST*, *PPY*, *PRSS*, *SOD2*, and *PDK4*. Among the 52 genes were three genes—retinol-binding protein 4 (*RBP4*) (46–48), delta-like noncanonical Notch ligand 1 (*DLK1*) (43,44,46), and homocysteine-responsive endoplasmic reticulum-resident ubiquitin-like domain member 1 (*HERPUD1*) (44)—that had previously been self-reported as heterogeneously expressed by one of the original human scRNA-Seq studies. While multiparameter signature analyses, such as gene set pathway testing, can be a powerful tool to make meaning out of subtle changes across varying genes, our small list of 52 did not suffice for further downstream analysis.

Surprised by the fact that only a limited number of genes drive variation between these two β -cell subpopulations, and the fact that non- β markers featured prominently in this list, we limited our analysis on only the β -cells from healthy donors (Supplementary Fig. 3A). Because these β -cells are more closely related to each other than, for example, to α - and δ -cells, clustering is confounded significantly by study-related confounders such as sequencing platform, genetic variation among donors, and variations in islet collection and culture parameters, suggesting that these contributions outweighed the contributions of true biological heterogeneity to clustering of β -cells (Supplementary Fig. 3B and C). Indeed, in a Venn diagram of the 2,000 genes that drove clustering of β -cell subpopulations from healthy donors for each individual study, only a distinct minority of 24 genes (1.2%) emerged as common drivers of heterogeneous

expression among β -cells across all five human β -cell scRNA-Seq studies (Supplementary Table 4 and Supplementary Fig. 3D–F). Moreover, approximately half of the genes that drove clustering of β -cells were unique to that particular data set and did not contribute to β -cell subpopulation clustering in each of the other human β -cell scRNA-Seq data sets (Supplementary Fig. 3D). *NPY* was the only gene encoding a known β -cell heterogeneity marker on this list.

To confirm these results, we selected 10 of these 24 genes that had a low to high range of abundance to evaluate how the expression of these genes compared across the same five studies. We observed varying fractions of β -cells with detectable expression (counts per million [CPM] >1) (Supplementary Fig. 3E) and comparable distribution of gene expression in violin plots (Supplementary Fig. 3F) for the majority of these 10 genes. Overrepresentation of *INS* (Supplementary Fig. 4A) may have caused poor library complexity by reducing the detection of other, less abundant genes below the detection limit. This is a general drawback of scRNA-Seq and explains in part why the number of detectable genes in each single cell is several-fold lower than the number of detectable genes in the same sample processed for bulk RNA-Seq (50,51). Two of the studies in our meta-analysis had the foresight to include in their experimental design the parallel processing of bulk samples from the same donors that were used to generate scRNA-Seq libraries, although cold-ischemic and postisolation culture times, as well as processing and dissociation methods, varied between them (43,47). This revealed that the average number of genes detectably expressed in whole-islet bulk islet samples (CPM >1) approximates 15,000, while the number of genes that are detectably expressed in each single β -cell ranges from 2,000 to 6,000 (Supplementary Table 1 and Supplementary Fig. 4B and C), with the subset of genes that is detected in each single β -cell in large part determined by chance (52). To illustrate this heterogeneous detection, we plotted the fraction of single human β -cells with detectable expression for all genes ranked in descending order of abundance (Fig. 2 and Supplementary Fig. 5A–E). This revealed a clear correlation between the average level of gene expression and the rate of detection in single human β -cells across all five studies, with more abundant genes detected in a larger fraction of β -cells. However, only an exceedingly small number of 86 genes (0.46% of all detectable genes; range 1 [0.005%]–153 [0.83%]) on average was detectable across all single β -cells in any given study. This is an obvious concern, as even the most conservative estimates place the number of housekeeping genes—genes required at all times in each cell—at several hundred (53). Stated differently, for the large majority of genes ($>99.5\%$) heterogeneous detection in single β -cells is the norm (Fig. 2 and Supplementary Fig. 5A–E). It is highly unlikely that all of these genes are truly heterogeneously expressed in β -cells. Instead, this observation indicates that heterogeneous detection of expression in single cells

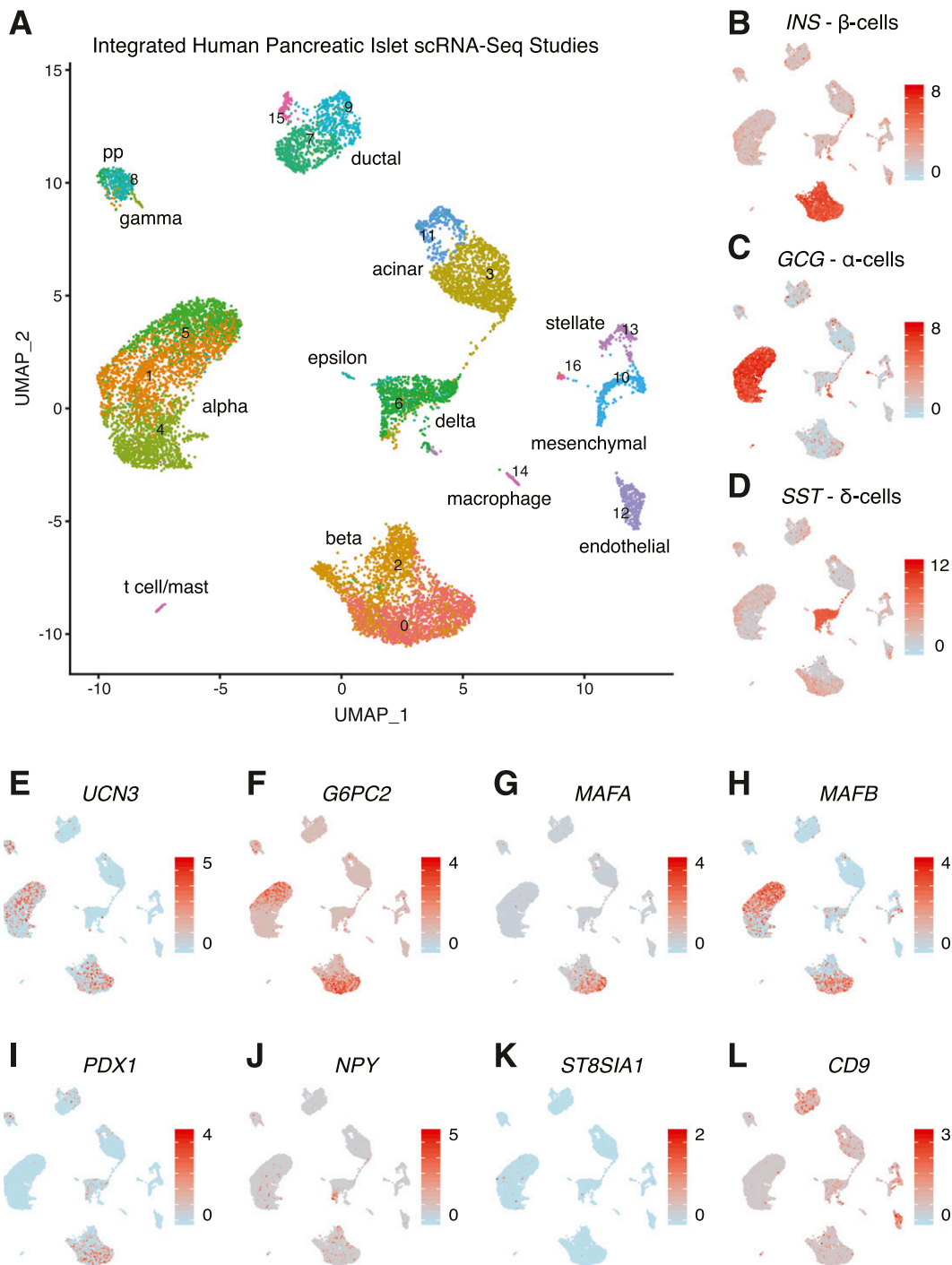


Figure 1—Integrated analysis of five human pancreatic islet scRNA-Seq studies across healthy donors. **A**: Dimensional reduction through uniform manifold approximation and projection across the five human studies identified clear clusters based on cell identity (see Supplementary Data, Approach & Tools for details). Two subpopulation clusters emerged from this analysis within the β -cell group, prompting for differential expression testing that identified 52 genes as significantly driving heterogeneity between the two (Supplementary Fig. 3A) ($P \leq 0.05$). **B–D**: Dimensional reduction of our clustered population to confirm β -, α -, and δ -cell clusters based on hallmark hormone expression. **E–L**: Exploring gene presence across selected, established markers identified in β -cell heterogeneity. PP, pancreatic polypeptide. Data from 43,44,46–48.

may be a poor predictor of actual single β -cell expression. Collectively, these observations suggest that heterogeneity of detection that is observed across single human β -cells may largely reflect the low fidelity of detection that is a consequence of operating at or below the limit of detection

for a majority of transcripts. This may also have driven the limited overlap among the shared set of genes that emerged as common contributors to β -cell clustering across the five human scRNA-Seq studies we assessed in our analysis. Nevertheless, accumulated across all cells in a pool of single

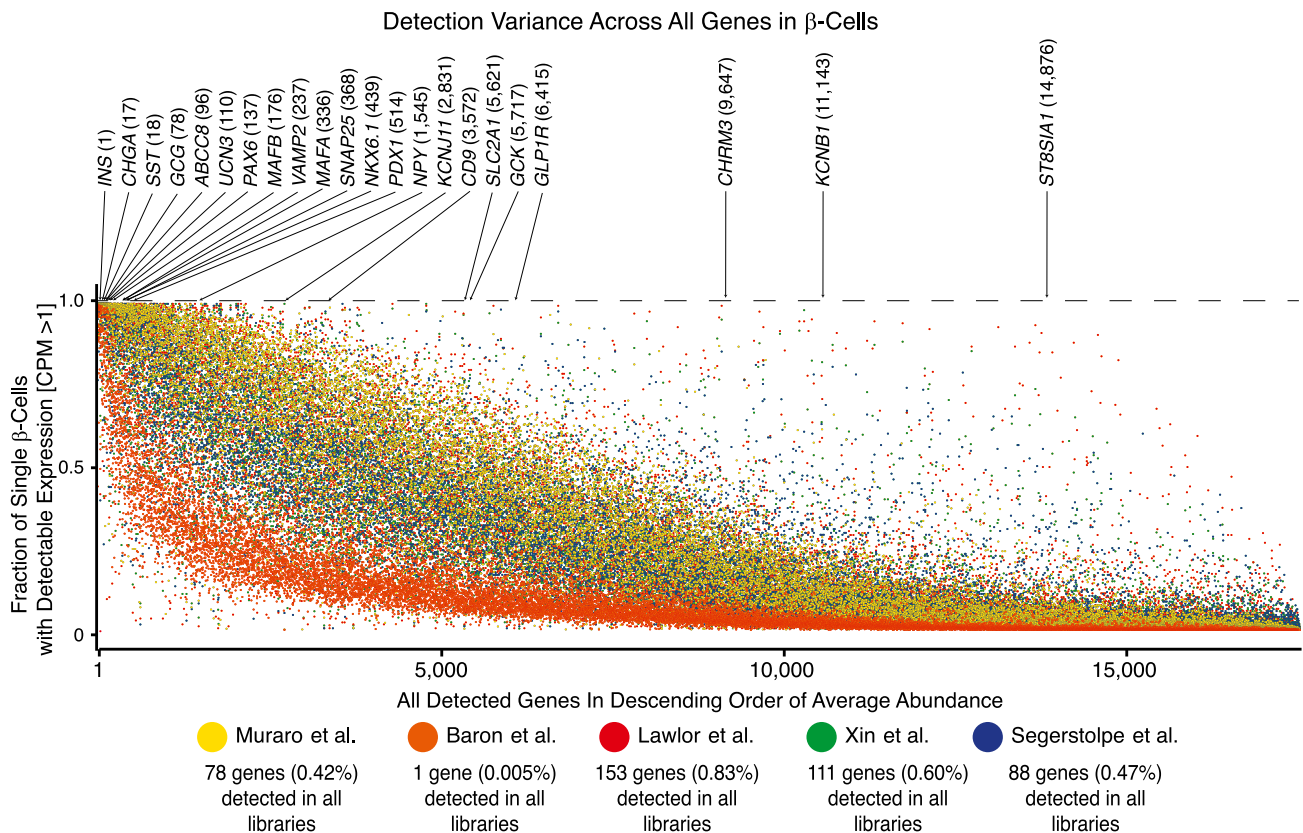


Figure 2—Abacus-style plot for detection of variance across all detectable genes within single-cell libraries. Transcriptome-wide representation of the fraction of single human β -cells where each gene is detected. Genes are ranked in descending order of average expression level; CPM >0 was used as the detection threshold. Key β -cell genes are highlighted across the plot to identify rank order of abundance (see Fig. 4A). The large majority of genes are heterogeneously detected; between 0.005% and 0.83% of all genes are detected in all β -cells in any given study. See Supplementary Fig. 5A–E for expanded view. Data from 43,44,46–48.

β -cell libraries, single β -cells correlate reasonably well with their companion bulk samples and recapitulate the total number of genes detectable in bulk (Supplementary Fig. 4D).

QUALITY OF DETECTION AND COVERAGE OF scRNA-Seq DATA

Prompted by the lack of congruence between scRNA-Seq studies and the fact that partial detection of gene expression across only a subset of β -cells is the norm for the large majority of genes, we further evaluated the quality of the scRNA-Seq data in comparison with conventional bulk RNA-Seq. All library preparation protocols can introduce bias based on mRNA stability, guanine-cytosine (GC) content, and mRNA abundance. Time of islet culturing and processing differ between studies as well, which may also add to these biases. Moreover, some library preparation protocols are 3' based, while others attempt to cover the entire length of the transcript. In addition to this, most scRNA-Seq approaches include two rounds of PCR amplification to exponentially amplify the extremely small amount of input material present in a single cell but that simultaneously amplify noise. We therefore evaluated the coverage of reads across these genes using the University of California, Santa Cruz (UCSC) Genome Browser,

which is an excellent tool to visualize gene expression data (54). We chose to use the data from Segerstolpe et al. (47) for this exercise, as theirs was one of the studies that included parallel single-cell and bulk sequencing from the same human islet samples. Other islet scRNA-Seq data perform similarly (see below). The sequence read coverage of *INS* (insulin), the most abundantly expressed β -cell gene, was homogeneous across the length of the gene model from 3' to 5' and the reads faithfully captured the known intron/exon structure of the coding strands in each of the 161 individual β -cells in this study (47) (Fig. 3A and Fig. 4C–G). In sharp contrast, coverage of the well-known β -cell transcription factor *MAFA* that was uniformly covered in bulk RNA-Seq companion data was marred by serious 3' bias in single- β -cell transcriptomes (Fig. 3B), a consequence of the oligo-dT priming step used to preferentially amplify mRNA from its poly-A tail over contaminating ribosomal RNA species (Fig. 3B). Moreover, no *MAFA* was detected (CPM >1) in 36 out of 161 (22%) β -cells (Fig. 3B), which is hard to reconcile with the general view of *MAFA* as an important β -cell transcription factor necessary for β -cell maturity that is detectable by immunohistochemistry in the nucleus of 88% of human β -cells (55). In parallel, when evaluating

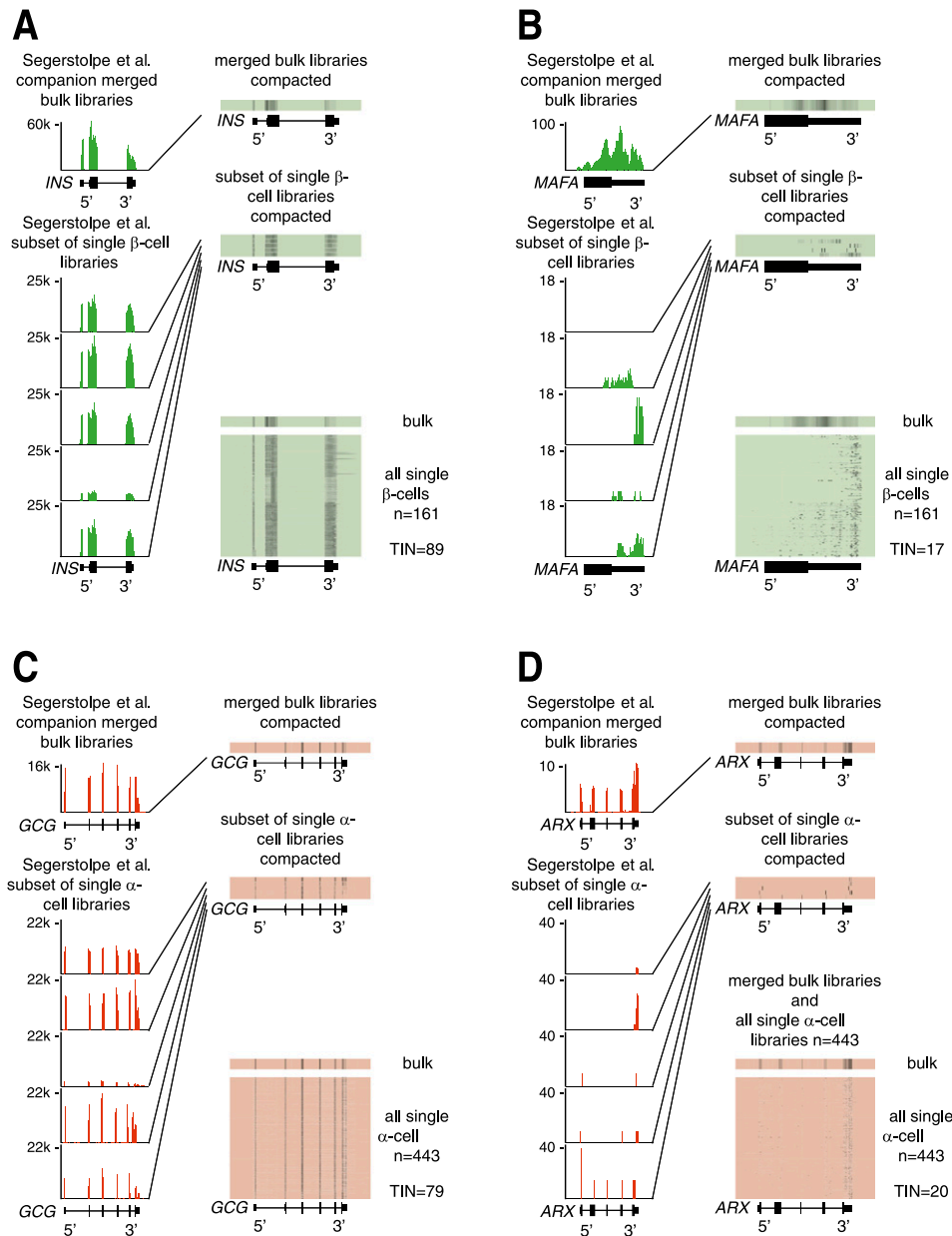


Figure 3—Visualizing the detection across the gene model between single and bulk sequencing. **A:** UCSC Genome Browser comparing the read coverage of insulin (*INS*) between the Segerstolpe et al. (47) healthy single β -cell population ($n = 161$) and their companion bulk islet libraries. **B:** *MafA* is an established β -cell marker that is poorly detected in single human β -cells. **C:** The highly abundant expression of *GCG* uniformly covered in single human α -cells ($n = 443$). **D:** The important α -cell transcription factor *ARX*, in contrast, is captured poorly in healthy single α -cells.

MAFA capture across all five studies, less than 50% of cells in the β -cell cluster had detectable expression as determined through Seurat (56) (Fig. 1G). Using the same approach across the study’s 443 single α -cell libraries revealed uniform coverage of the abundant *GCG* transcript in each single α -cell library (Fig. 3C). However, the transcription factor *ARX*, which is required for α -cell identity (57,58), was not detected at all in 18% of α -cells with evidence of significant 3’ bias in incomplete coverage in the α -cells with detectable *ARX* expression (Fig. 3D).

REPRODUCING KNOWN β -CELL GENE EXPRESSION

These observations raise the question of whether the heterogeneous detection of mRNA expression in single β -cells reflects true biological heterogeneity in gene expression or instead is a product of the inherent limitations of scRNA-Seq. Therefore, we queried if single β -cell transcriptomes accurately detected genes encoding for proteins that are required by every single β -cell, as well as genes that encode for proteins with well-documented and validated heterogeneous expression across the β -cell population

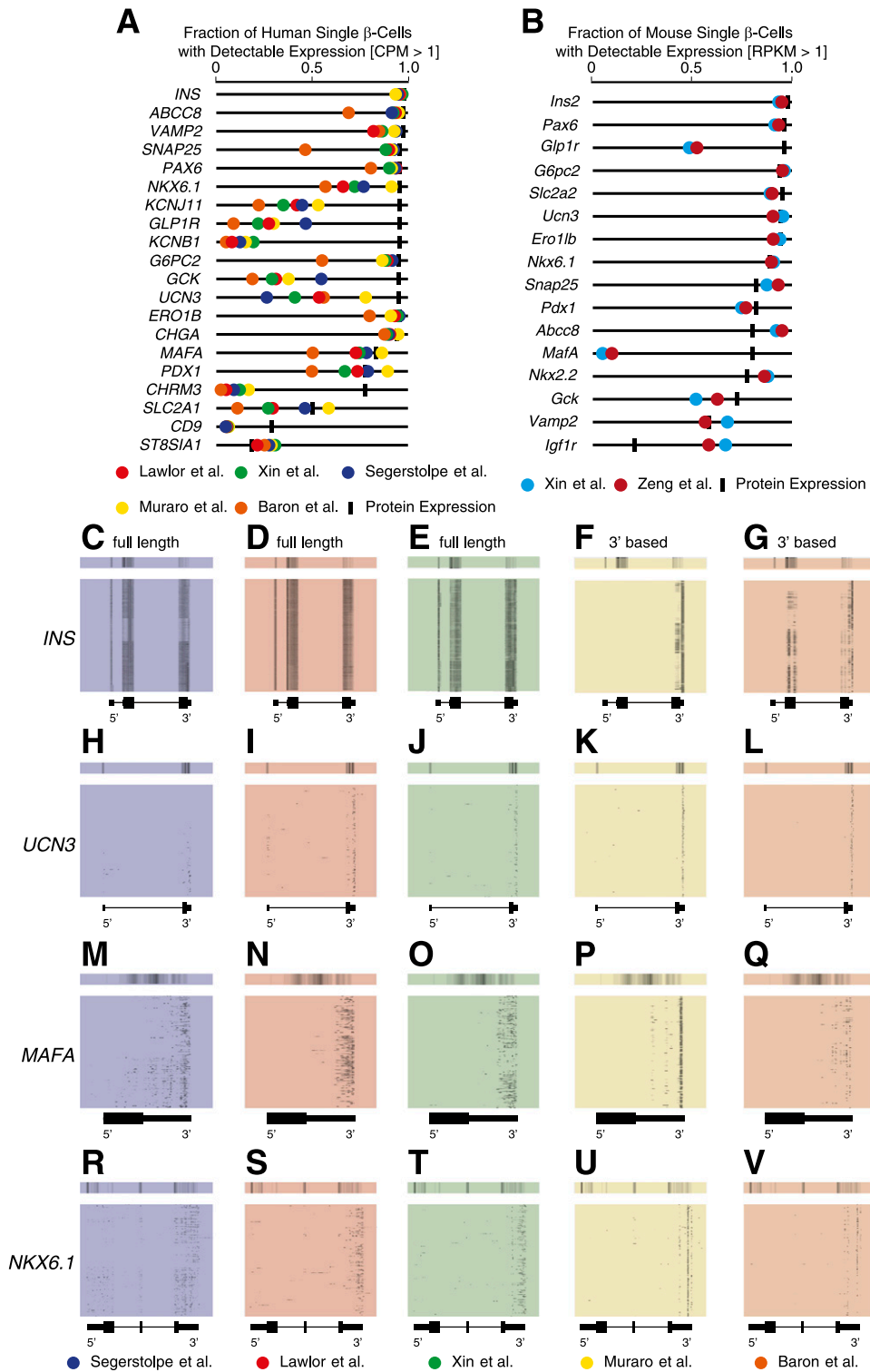


Figure 4—Comparison of detection of established β -cell markers in single β -cells comparing RNA-Seq and immunohistochemistry and coverage across the gene body of established β -cell markers detected by scRNA-Seq. **A**: Fraction of single human β -cells with detectable expression (CPM > 1) across five human β -cell scRNA-Seq studies, compared with the fraction of human β -cells reported to express the corresponding protein (Supplementary Table 5A). The majority of genes are underdetected in human scRNA-Seq compared with the fraction of human β -cells that express the corresponding protein. **B**: The approach used in Fig. 4A, applied to established β -cell markers comparing two mouse scRNA-Seq studies and expression of the proteins encoded by these genes (Supplementary Table 5B). **C–V**: UCSC Genome Browser plots of read coverage across these genes, using single- β -cell libraries and companion bulk sequencing across all studies and platforms, highlighting differences between full-length and 3'-based capture (43,44,46–48).

(Supplementary Table 5). In addition to *INS*, examples include transcription factors such as *PDX1* (57), *NKX6.1* (59), *PAX6* (60), and *MAFA* (61), as well as proteins required for normal stimulus-secretion coupling, insulin processing, and exocytosis such as *SLC2A1* (62), *GLP1R* (63), *ABCC8* (64), *KCNJ11* (65), *GCK* (66), *G6PC2* (67), *KCNB1* (65), *ERO1B* (68), *VAMP2* (69), *SNAP25* (69), and *UCN3* (70). With the exception of *PDX1*, *MAFA*, and *SLC2A1*, all of the proteins encoded by these genes are detected in more than an estimated 95% of human β -cells in healthy islets by immunohistochemical techniques. However, mRNA for all but the most abundantly expressed of these genes is consistently detected in a decidedly smaller fraction of β -cells than stain positive for the protein product they encode (Fig. 4A). *NKX6.1*, *UCN3*, *KCNJ11*, and *KCNB1* transcripts are detected in a particularly low fraction of β -cells. One possible explanation for this is intermittent transcription, where transcription occurs in discrete bursts that underlies stable protein expression (71). However, if this is the case, one would expect uniform coverage gene body capture for the subset of β -cells that would have been captured during the burst phase of expression for that gene. Instead, UCSC Genome Browser plots for these genes indicate widespread 3' bias and underrepresentation of many known β -cell genes, even those that are expressed at medium to high transcript levels such as *UCN3*, *MAFA*, and *NKX6-1* (Fig. 4H–V). This is a likely consequence of working at or below the level of detection of scRNA-Seq approaches. One uncommon example with read coverage across the full gene model was observed in a distinct subset of β -cells for *DLK1*, which reflects a pattern in line with burst transcription (Supplementary Fig. 6A–E). *ST8SIA1* and *CD9*, two genes that encode protein markers recently used to distinguish four distinct human β -cell types (21), are also consistently underdetected in single-human β -cell transcriptomes. A similar set of ubiquitous β -cell genes that are expressed at medium to high levels in mouse β -cell transcriptomes are detected in a higher fraction of β -cells, although large discrepancies remain for *Glp1r* and *Mafa* (Fig. 4B).

ASSESSING SINGLE-CELL SEQUENCING QUALITY

Until this point, we have largely used the fraction of β -cells with detectable expression (CPM >1) of a given gene as a metric of the fidelity of scRNA-Seq (Fig. 2, Supplementary Fig. 1B, and Supplementary Fig. 3E). This revealed that heterogeneous detection and significant 3' bias is the norm for single-human β -cell transcriptomes, irrespective of investigator, approach, or platform (Supplementary Table 1). To better quantify the gap in quality of gene coverage between single-cell and bulk sequencing approaches, we adopted the transcript integrity number (TIN) score (72). This metric ranges between 0 and 100 and is calculated after library preparation and sequencing to reflect the quality and uniformity of read coverage across the gene model. A high TIN score for a gene reflects uniform read coverage across the gene

model, while a low TIN score reflects uneven coverage across the gene model owing to 3' bias, GC bias, or transcript degradation (Fig. 3 and Supplementary Fig. 7). TIN scores strongly correlate with the RNA integrity number, a measure of RNA quality used to assess input RNA quality before library preparation.

To visualize the relationship between gene expression and quality of its representation in single-cell versus bulk RNA-Seq approaches, we compared the correlation of TIN scores and gene expression among five human (43,44,46–48) and two mouse single-cell studies (42,45) with two mouse (18,19) and three human bulk islet RNA-Seq data sets (43,47,73). For bulk RNA-Seq approaches, there is essentially no drop-off in TIN score with lower gene expression (expressed as CPM) until CPM values are <5 (Fig. 5A). In other words, in bulk RNA-Seq approaches, the quality of the coverage of gene expression across the gene model from 5' to 3' is both high and independent of transcript abundance unless gene expression is quite low. In sharp contrast, in scRNA-Seq, there is a very clear effect of the abundance of gene expression on TIN score across the full range of transcript abundance values. Even at highly abundant transcripts with CPM values >100, TIN scores remain well below those of similarly abundant genes detected via bulk RNA-Seq. This reflects the drop-off in the quality of sequence coverage that is the consequence of working at or below the level of detection in scRNA-Seq approaches. Limiting analysis to only genes with a consistently high TIN score would yield more reliable and reproducible results but would also drastically undercut the number of genes that are included in the analysis, as over half of the genes detected in human β -cell scRNA-Seq have TIN scores <20. A comparison of TIN score cutoff versus CPM cutoff to the fraction of remaining genes suggests that TIN score cutoffs are a better metric than CPM cutoffs to separate high- and medium-quality read data (Fig. 5B).

Conversely, for bulk RNA-Seq samples, significant numbers of genes are excluded from the analysis only when the TIN quality threshold is raised over 50 (Fig. 5C).

CROSS-CONTAMINATION IN SINGLE-ISLET CELL TRANSCRIPTOMES

One question that continues to stir debate in the field is whether healthy β -cells transcribe *GCG* at low abundance and conversely if α -cells transcribe *INS*. Indeed, the α -cell cluster in our study clearly contains lower but detectable levels of *INS*, and β -cells had detectable levels of *GCG* (Fig. 1B and C). While cells that coexpress insulin and glucagon protein are regularly observed during embryonic development and in stem cell-derived β -cell-like cultures (32,74,75), they are exceedingly rare in healthy adult islets (76,77). However, this does not rule out translational inhibition of *GCG* in β -cells and *INS* in α -cells. Indeed, bulk RNA-Seq data of FACS-purified mouse α -cells detect *Ins2* expression at 80- to 170-fold lower than *Ins2* in β -cells from the same islets (Fig. 6A) (18). Similarly, *Gcg* is detected

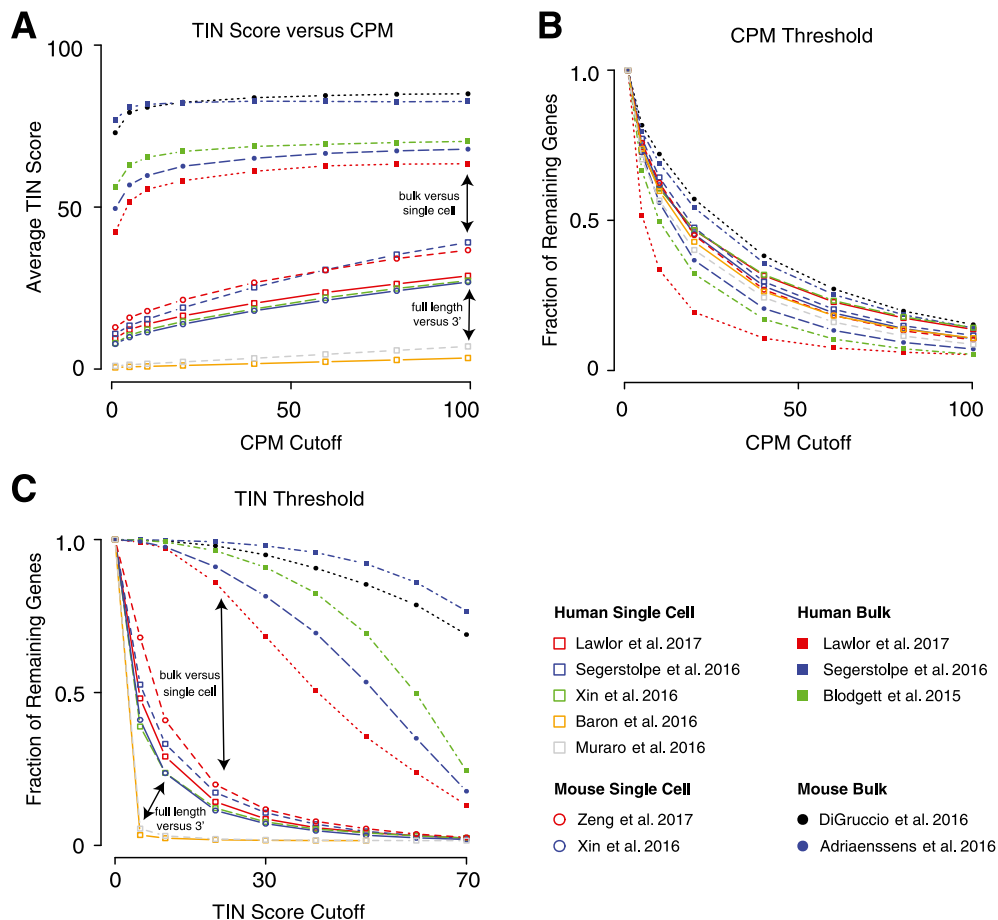


Figure 5—Visualization of the difference in sequence quality between single-cell and bulk RNA-Seq. **A**: Average TIN quality score plotted against transcript abundance in CPM across five human and two mouse single-cell and bulk RNA-Seq data sets. Note the large difference in quality between single-cell and bulk RNA-Seq data and that TIN scores in bulk are uniformly high irrespective of transcript abundance. **B**: Progressive application of a CPM threshold leads to the exclusion of similar relative numbers of genes between single-cell and bulk data sets (after correcting for the fact that single-cell libraries detect between 2,000 and 6,000 genes, while bulk libraries detect around 16,000 genes). **C**: Progressive application of a TIN threshold leads to the rapid exclusion of genes from single-cell data sets but only effects bulk RNA-Seq data sets at much more stringent TIN score cutoffs. Data from human (43,44,46–48) and mouse (42,45) scRNA-Seq and human (43,47,73) and mouse (18,19) bulk RNA-Seq.

in FACS-purified β -cells at 220-fold lower levels than its expression in α -cells (Fig. 6B) (18). This relatively low level of detectable reads could be caused by cross-contamination during FACS purification. While doublets, including those consisting of an α -cell and a β -cell, are normally gated out before collection, a well-calibrated FACS running at a conservative speed has an error rate less than 1%. In the context of FACS purification of dissociated islet suspensions, this means that fewer than 1% of the events that are sorted as β -cells are in fact a non- β -cell, possibly an α -cell. Since *GCG* accounts for up to 20% of all reads in the α -cell pool (17,18), a couple of contaminating α -cells could suffice to explain the detection of *Gcg* in transcriptomes of bulk FACS-purified β -cells.

Single-cell approaches ostensibly do not suffer from this confounder as they assess transcription in individual cells. Indeed, 0.2–1.5% of all reads in single human β -cells map to *GCG* and 0.001–1.109% of reads in single human α -cells map to *INS*. These observations at face value have been

suggested as definitive proof that β -cells express *GCG* and α -cells express *INS*. However, Macosko et al. (78), in their original article describing the Drop-Seq approach, conducted a key control experiment that is often overlooked but is of direct relevance in this discussion. They approached the question of contamination at the single-cell level by mixing human HEK cells and mouse 3T3 cells prior to droplet formation and single-cell sequencing. They observed that an average of 0.26–2.44% of the reads in each and every single cell mapped uniquely to the genome of the other species (Fig. 6C). As they demonstrate, this can only be explained by the integration of free-floating or naked mRNA derived from cells that were disrupted by generating cell suspensions into libraries constructed from single cells that did not actually express the message (78). This problem is not unique to the Drop-Seq approach but will affect any procedure where tissues are dissociated into a single-cell suspension in preparation of single-cell sequencing or FACS sorting in bulk RNA-Seq approaches

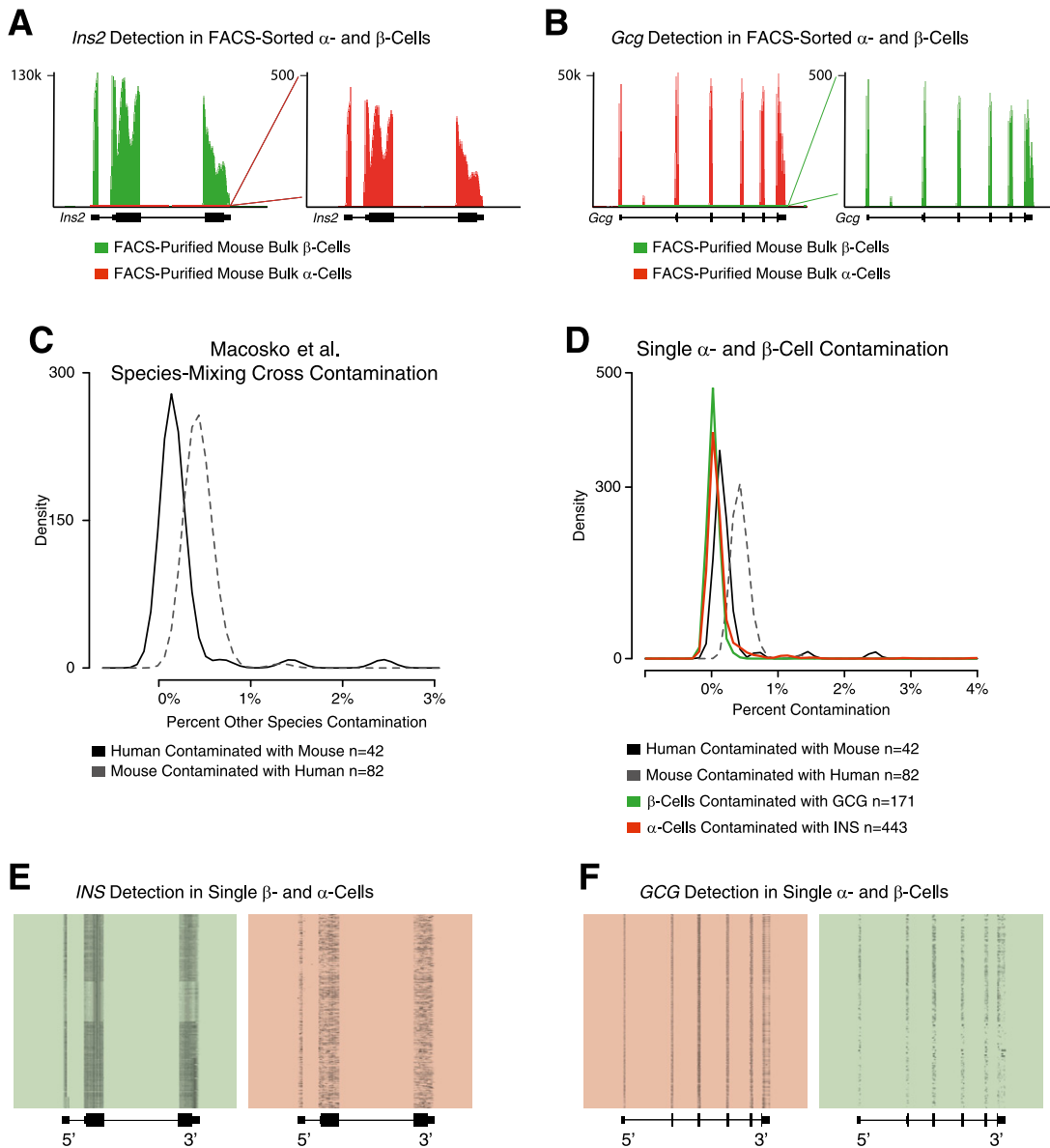


Figure 6—Detection of *GCG* in β -cells and *INS* in α -cells in bulk and scRNA-Seq. *A*: Detection of *Ins2* in FACS-purified mouse α -cells at 260-fold lower levels compared with β -cells from the same islets. Data from DiGrucchio et al. (18). *B*: Detection of *Gcg* in FACS-purified mouse β -cells at 100-fold lower levels compared with α -cells from the same islets. Data from DiGrucchio et al. (18). *C*: Species mixing experiments between cell lines of mouse and human origin have illustrated that 0.26–2.44% of all reads detected in single-cell libraries uniquely mapped to the other species. Data from Macosko et al. (78). *D*: Fraction of reads within 161 single- β -cell libraries that maps to *GCG*, and fraction of reads within 443 single- α -cell libraries that maps to *INS*. Data from Segerstolpe et al. (47). The *INS* and *GCG* cross-detection is below the threshold of species contamination established by Macosko et al. (78). Given the extremely high abundance of *INS* and *GCG*, their detection in single α - and β -cells, respectively, may not be evidence of actual gene expression. *E*: Coverage of *INS* across all single β - and α -cell libraries. Data from Segerstolpe et al. (47). *F*: Coverage of *GCG* across all single α - and β -cell libraries. Data from Segerstolpe et al. (47).

(78). This relatively low level of cross-contamination will likely not meaningfully affect detection of the large majority of genes. However, *INS* and *GCG* are expressed so abundantly in β - and α -cells, respectively, that their cross-detection could be explained entirely by contamination of free-floating mRNA (Fig. 6*D–F*). These observations do not rule out true *GCG* expression by β -cells. However, the detection of *GCG* in single- β -cell transcriptomes at levels below those estimated through the species

cross-contamination paradigm established by Macosko et al. (78) cannot be taken as proof that β -cells actually express *GCG* mRNA.

FUTURE OUTLOOK

The fact that we can now detect and attempt to quantify gene expression in single cells is in itself a remarkable achievement. A survey of β -cell gene expression at single-cell resolution across hundreds or even thousands of

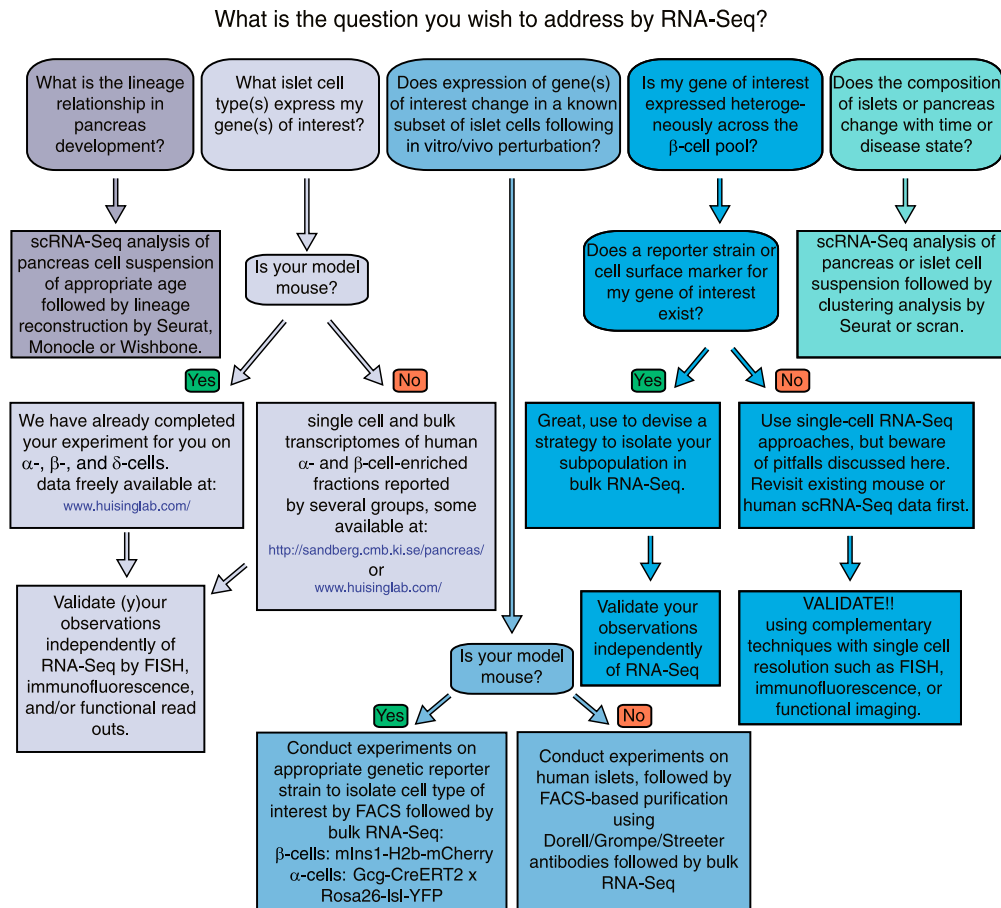


Figure 7—Workflow to align the appropriate sequencing approaches with the stated experimental goal. Starting from five common experimental scenarios, this flowchart offers guidance to the reader to facilitate the choice for the different experimental strategies available, considering their benefits and drawbacks. FISH, fluorescence in situ hybridization.

individual cells is a very enticing prospect that would resolve some of the long-known heterogeneity among β -cells with regard to their functional state or proliferative status. However, in attempting to detect gene expression in single β -cells, it has become obvious that we are operating at or below the limit of reliable detection for a large majority of genes. This comes at a steep price with regard to the quality of the single-cell sequence data that is obtained, irrespective of the investigating laboratory or the chosen single-cell approach.

In this Perspective, we have illustrated these inherent limitations of scRNA-Seq applied to adult human islet cells by pointing out the underestimation of the number of detected genes per single cell and by applying TIN scores as a quantitative measure of the incomplete coverage and 3' bias that affects all genes, from rare to highly abundant. By comparison, the quality of the gene coverage in bulk RNA-Seq samples is so much better that it is quite possible that the coverage and data quality of scRNA-Seq may not approach that of bulk RNA-Seq for some time. Therefore, for each experiment investigators need to determine if transcript detection at single-cell resolution is worth these

inevitable drawbacks (Fig. 7). Given the large quality gap between single-cell versus bulk transcriptome data, we would advocate for a bulk transcriptome approach, if compatible with your experimental question, in spite of the perceived novelty of single-cell sequencing. Evidently, if transcriptional heterogeneity among β - or α -cells is the central focus of a study, scRNA-Seq experiments may be the only choice, unless a known marker for these subpopulations can be leveraged to isolate these cells by FACS for bulk sequencing. Nevertheless, our illustration that—with the exception of a handful of the most highly abundant transcripts—every single gene is detected in only a fraction of β -cells questions the ability of scRNA-Seq to discern true heterogeneous expression amid widespread heterogeneous detection. Case in point is the fact that none of the many markers of known heterogeneity were independently identified by any of the “unbiased” scRNA-Seq approaches, with some acknowledging their inability to do so (43). Therefore, any observation derived from single-cell or bulk RNA-Seq experiments should—wherever possible—be subject to rigorous validation using independent approaches that can achieve single-cell resolution, such

as RNA fluorescence in situ hybridization to detect the message, immunofluorescence to detect the protein encoded by that mRNA, and/or live cell functional imaging to correlate gene expression with functional readout that indicates the presence of the corresponding protein.

Our intent in drawing attention to the limitations of scRNA-Seq approaches applied to islet cells is certainly not to dissuade our colleagues from relying on observations obtained by scRNA-Seq approaches in their studies of islet function. It should not be a surprise that working at the extreme limits of our technical capabilities comes at a price. Ongoing improvements in library preparation, including the generation of protocols that no longer rely on multiple rounds of PCR amplification, should constitute a significant improvement (79,80). As tissue collection and processing time will influence gene expression and mRNA stability, standardization of the collection of human islets to the extent possible will increase the conformity of gene representation across both bulk and scRNA-Seq studies. New methods, such as split-pool ligation-based transcriptome sequencing (SPLiT-Seq) (81), may be able to overcome some of the limitations of current scRNA-Seq protocols. While SPLiT-Seq still requires tissue dissociation, it instead compartmentalizes the RNA into single-cell libraries within the native cell rather than relying on droplets or wells. This may help mitigate issues such as poor library complexity and 3' bias and may reduce contamination with naked mRNA (78) (Fig. 6). Spatial transcriptomics may also provide more reliable avenues for scRNA-Seq, as they avoid confounders associated with islet dissociation and would allow the field an unbiased perspective to determine whether heterogeneity of gene expression is spatially driven (82), as was recently suggested (35,83). Newer 3'-based methods that allow for higher throughput, and greater sample sizes at reasonable cost, have allowed the identification of rarer populations, such as ϵ -cells (84). Targeted sequencing approaches, such as droplet-assisted RNA targeting by single-cell sequencing (DART-Seq), significantly improve coverage by targeting the limited depth of scRNA-Seq to a subset of preselected transcripts of interest (85). Computational methods are being developed to take into account and correct for confounding factors, such as donor genetic variation, dropout, and technical noise, although avoiding confounders will always be preferable to correcting for them through bioinformatic means (86,87).

Despite the current limitation of the approach, scRNA-Seq experiments have successfully resolved gene expression in human δ -cells (47,88), for which purification methods to obtain bulk samples do not exist. Moreover, scRNA-Seq has recapitulated known differentiation trajectories in the development of many organs and tissues (89–93). This includes the pancreas, where scRNA-Seq has been able to trace Ngn3+ progenitor populations at different embryonic ages to preferentially differentiate into α - or β -cells (94), thus recapitulating and validating a phenomenon that had previously been independently described

by careful developmental biology experiments (95,96). *Gjd2*, *Scg2*, *Ociad2*, and *Fev*, novel genes whose contribution to embryonic pancreas development had not been known, have also emerged from scRNA-Seq efforts (80,94). Moreover, pseudo-time strategies, where single cells are placed on a lineage based on their transcriptional stage instead of their chronological age, have successfully resolved aspects of postnatal β -cell maturation (42).

In summary, our goal with this Perspective has been to raise awareness among a general audience of diabetes researchers of some of the limitations of scRNA-Seq and discuss potential solutions to overcome the current limitations. It is amazing that we are now capable of detecting islet cell gene expression at single-cell resolution. It therefore should not come as a surprise that there is inevitably a price to pay for the benefit of single-cell resolution. The limitations we discussed should be well known to investigators who have been at the forefront of single-cell sequencing. However, they are likely less appreciated by a general audience of diabetes researchers not as well versed in bioinformatics, who nevertheless use scRNA-Seq data generated by others or are adopting scRNA-Seq for their own future experiments. Next-generation sequencing at single-cell resolution has the potential to reveal unprecedented insight into biological processes that until recently had remained out of reach. We hope that the considerations discussed in this Perspective will help our colleagues align their sequencing approaches with realistic experimental goals.

Acknowledgments. The authors thank Dr. Talitha van der Meulen, University of California, Davis, for constructive comments on the manuscript.

Funding. This work was supported by grants from the National Institutes of Health National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK110276) and JDRF (CDA-2-2013-54) to M.O.H. A.M.M. was supported by the Stephen F. and Bettina A. Sims Immunology Fellowship.

Duality of Interest. No potential conflicts of interest relevant to this article were reported.

References

1. Menke A, Casagrande S, Geiss L, Cowie CC. Prevalence of and trends in diabetes among adults in the United States, 1988–2012. *JAMA* 2015;314:1021–1029
2. Halban PA, Polonsky KS, Bowden DW, et al. β -Cell failure in type 2 diabetes: postulated mechanisms and prospects for prevention and treatment. *Diabetes Care* 2014;37:1751–1758
3. Schuit FC, Huypens P, Heimberg H, Pipeleers DG. Glucose sensing in pancreatic β -cells: a model for the study of other glucose-regulated cells in gut, pancreas, and hypothalamus. *Diabetes* 2001;50:1–11
4. Röder PV, Wu B, Liu Y, Han W. Pancreatic regulation of glucose homeostasis. *Exp Mol Med* 2016;48:e219
5. Johnston NR, Mitchell RK, Haythorne E, et al. Beta cell hubs dictate pancreatic islet responses to glucose. *Cell Metab* 2016;24:389–401
6. Brissova M, Fowler MJ, Nicholson WE, et al. Assessment of human pancreatic islet architecture and composition by laser scanning confocal microscopy. *J Histochem Cytochem* 2005;53:1087–1097
7. Brereton MF, Vergari E, Zhang Q, Clark A. Alpha-, delta- and PP-cells: are they the architectural cornerstones of islet structure and co-ordination? *J Histochem Cytochem* 2015;63:575–591

8. Cabrera O, Berman DM, Kenyon NS, Ricordi C, Berggren PO, Caicedo A. The unique cytoarchitecture of human pancreatic islets has implications for islet cell function. *Proc Natl Acad Sci U S A* 2006;103:2334–2339
9. Arrojo e Drigo R, Ali Y, Diez J, Srinivasan DK, Berggren PO, Boehm BO. New insights into the architecture of the islet of Langerhans: a focused cross-species assessment. *Diabetologia* 2015;58:2218–2228
10. Ranjan AK, Joglekar MV, Hardikar AA. Endothelial cells in pancreatic islet development and function. *Islets* 2009;1:2–9
11. Unanue ER. Macrophages in endocrine glands, with emphasis on pancreatic islets. *Microbiol Spectr* 2016;4:1–11
12. Brissova M, Haliyur R, Saunders D, et al. α Cell function and gene expression are compromised in type 1 diabetes. *Cell Reports* 2018;22:2667–2676
13. Lam CJ, Cox AR, Jacobson DR, Rankin MM, Kushner JA. Highly proliferative α -cell-related islet endocrine cells in human pancreata. *Diabetes* 2018;67:674–686
14. Dai C, Brissova M, Reinert RB, et al. Pancreatic islet vasculature adapts to insulin resistance through dilation and not angiogenesis. *Diabetes* 2013;62:4144–4153
15. Pipeleers D. The biosociology of pancreatic B cells. *Diabetologia* 1987;30:277–291
16. Hara M, Wang X, Kawamura T, et al. Transgenic mice with green fluorescent protein-labeled pancreatic β -cells. *Am J Physiol Endocrinol Metab* 2003;284:E177–E183
17. Benner C, van der Meulen T, Cacéres E, Tigyi K, Donaldson CJ, Huisin MO. The transcriptional landscape of mouse beta cells compared to human beta cells reveals notable species differences in long non-coding RNA and protein-coding gene expression. *BMC Genomics* 2014;15:620
18. DiGrucio MR, Mawla AM, Donaldson CJ, et al. Comprehensive alpha, beta and delta cell transcriptomes reveal that ghrelin selectively activates delta cells and promotes somatostatin release from pancreatic islets. *Mol Metab* 2016;5:449–458
19. Adriaenssens AE, Svendsen B, Lam BY, et al. Transcriptomic profiling of pancreatic alpha, beta and delta cell populations identifies delta cells as a principal target for ghrelin in mouse islets. *Diabetologia* 2016;59:2156–2165
20. Dorrell C, Abraham SL, Lanxon-Cookson KM, Canaday PS, Streeter PR, Grompe M. Isolation of major pancreatic cell types and long-term culture-initiating cells using novel human surface markers. *Stem Cell Res (Amst)* 2008;1:183–194
21. Dorrell C, Schug J, Canaday PS, et al. Human islets contain four distinct subtypes of β cells. *Nat Commun* 2016;7:11756
22. Banerjee M, Otonkoski T. A simple two-step protocol for the purification of human pancreatic beta cells. *Diabetologia* 2009;52:621–625
23. Wang YJ, Golson ML, Schug J, et al. Single-cell mass cytometry analysis of the human endocrine pancreas. *Cell Metab* 2016;24:616–626
24. Clardy SM, Mohan JF, Vinegoni C, et al. Rapid, high efficiency isolation of pancreatic β -cells. *Sci Rep* 2015;5:13681
25. Pipeleers D, Kiekens R, Ling Z, Wilkens A, Schuit F. Physiologic relevance of heterogeneity in the pancreatic beta-cell population. *Diabetologia* 1994;37(Suppl. 2):S57–S64
26. Van Schravendijk CF, Kiekens R, Pipeleers DG. Pancreatic beta cell heterogeneity in glucose-induced insulin secretion. *J Biol Chem* 1992;267:21344–21348
27. Bader E, Migliorini A, Gegg M, et al. Identification of proliferative and mature β -cells in the islets of Langerhans. *Nature* 2016;535:430–434
28. Benninger RK, Piston DW. Cellular communication and heterogeneity in pancreatic islet insulin secretion dynamics. *Trends Endocrinol Metab* 2014;25:399–406
29. Olsson R, Carlsson PO. A low-oxygenated subpopulation of pancreatic islets constitutes a functional reserve of endocrine cells. *Diabetes* 2011;60:2068–2075
30. Kiekens R, In 't Veld P, Mahler T, Schuit F, Van De Winkel M, Pipeleers D. Differences in glucose recognition by individual rat pancreatic B cells are associated with intercellular differences in glucose-induced biosynthetic activity. *J Clin Invest* 1992;89:117–125
31. Bouwens L, Klöppel G. Islet cell neogenesis in the pancreas. *Virchows Arch* 1996;427:553–560
32. Teitelman G, Alpert S, Polak JM, Martinez A, Hanahan D. Precursor cells of mouse endocrine pancreas coexpress insulin, glucagon and the neuronal proteins tyrosine hydroxylase and neuropeptide Y, but not pancreatic polypeptide. *Development* 1993;118:1031–1039
33. Hermann M, Pirkebner D, Draxl A, et al. Dickkopf-3 is expressed in a subset of adult human pancreatic beta cells. *Histochem Cell Biol* 2007;127:513–521
34. Rodnoi P, Rajkumar M, Moin ASM, Georgia SK, Butler AE, Dhawan S. Neuropeptide Y expression marks partially differentiated β cells in mice and humans. *JCI Insight* 2017;2:1–13
35. van der Meulen T, Mawla AM, DiGrucio MR, et al. Virgin beta cells persist throughout life at a neogenic niche within pancreatic islets. *Cell Metab* 2017;25:911–926.e6
36. Beamish CA, Strutt BJ, Arany EJ, Hill DJ. Insulin-positive, Glut2-low cells present within mouse pancreas exhibit lineage plasticity and are enriched within extra-islet endocrine cell clusters. *Islets* 2016;8:65–82
37. Rui J, Deng S, Arazi A, Perdigoto AL, Liu Z, Herold KC. β cells that resist immunological attack develop during progression of autoimmune diabetes in NOD mice. *Cell Metab* 2017;25:727–738
38. Thompson PJ, Shah A, Ntranos V, Van Gool F, Atkinson M, Bhushan A. Targeted elimination of senescent beta cells prevents type 1 diabetes. *Cell Metab* 2019;29:1045–1060.e10
39. Enge M, Arda HE, Mignardi M, et al. Single-cell analysis of human pancreas reveals transcriptional signatures of aging and somatic mutation patterns. *Cell* 2017;171:321–330.e14
40. Qiu W-L, Zhang Y-W, Feng Y, Li L-C, Yang L, Xu C-R, et al. Deciphering pancreatic islet beta cell and alpha cell maturation pathways and characteristic features at the single-cell level. *Cell Metab* 2017;25:1194–1205.e4
41. Ackermann AM, Wang Z, Schug J, Naji A, Kaestner KH. Integration of ATAC-seq and RNA-seq identifies human alpha cell and beta cell signature genes. *Mol Metab* 2016;5:233–244
42. Zeng C, Mulas F, Sui Y, et al. Pseudotemporal ordering of single cells reveals metabolic control of postnatal β cell proliferation. *Cell Metab* 2017;25:1160–1175.e11
43. Lawlor N, George J, Bolisetty M, et al. Single-cell transcriptomes identify human islet cell signatures and reveal cell-type-specific expression changes in type 2 diabetes. *Genome Res* 2017;27:208–222
44. Baron M, Veres A, Wolock SL, et al. A single-cell transcriptomic map of the human and mouse pancreas reveals inter- and intra-cell population structure. *Cell Syst* 2016;3:346–360.e4
45. Xin Y, Kim J, Ni M, et al. Use of the Fluidigm C1 platform for RNA sequencing of single mouse pancreatic islet cells. *Proc Natl Acad Sci U S A* 2016;113:3293–3298
46. Xin Y, Kim J, Okamoto H, et al. RNA sequencing of single human islet cells reveals type 2 diabetes genes. *Cell Metab* 2016;24:608–615
47. Segerstolpe Å, Palasantza A, Eliasson P, et al. Single-cell transcriptome profiling of human pancreatic islets in health and type 2 diabetes. *Cell Metab* 2016;24:593–607
48. Muraro MJ, Dharmadhikari G, Grün D, et al. A single-cell transcriptome atlas of the human pancreas. *Cell Syst* 2016;3:385–394.e3
49. Wang YJ, Schug J, Won KJ, et al. Single-cell transcriptomics of the human endocrine pancreas. *Diabetes* 2016;65:3028–3038
50. Daley T, Smith AD. Predicting the molecular complexity of sequencing libraries. *Nat Methods* 2013;10:325–327
51. Daley T, Smith AD. Modeling genome coverage in single-cell sequencing. *Bioinformatics* 2014;30:3159–3165
52. Stegle O, Teichmann SA, Marioni JC. Computational and analytical challenges in single-cell transcriptomics. *Nat Rev Genet* 2015;16:133–145
53. Eisenberg E, Levanon EY. Human housekeeping genes, revisited. *Trends Genet* 2013;29:569–574

54. Kent WJ, Sugnet CW, Furey TS, et al. The human genome browser at UCSC. *Genome Res* 2002;12:996–1006
55. Guo S, Dai C, Guo M, et al. Inactivation of specific β cell transcription factors in type 2 diabetes. *J Clin Invest* 2013;123:3305–3316
56. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* 2018;36:411–420
57. Dai C, Brissova M, Hang Y, et al. Islet-enriched gene expression and glucose-induced insulin secretion in human and mouse islets. *Diabetologia* 2012;55:707–718
58. Collombat P, Mansouri A, Hecksher-Sorensen J, et al. Opposing actions of *Arx* and *Pax4* in endocrine pancreas development. *Genes Dev* 2003;17:2591–2603
59. Taylor BL, Liu FF, Sander M. *Nkx6.1* is essential for maintaining the functional state of pancreatic beta cells. *Cell Reports* 2013;4:1262–1275
60. Ahmad Z, Rafeeq M, Collombat P, Mansouri A. *Pax6* inactivation in the adult pancreas reveals ghrelin as endocrine cell maturation marker. *PLoS One* 2015;10:e0144597
61. Bonnavion R, Jaafar R, Kerr-Conte J, et al. Both *PAX4* and *MAFA* are expressed in a substantial proportion of normal human pancreatic alpha cells and deregulated in patients with type 2 diabetes. *PLoS One* 2013;8:e72194
62. McCulloch LJ, van de Bunt M, Braun M, Frayn KN, Clark A, Gloy AL. *GLUT2 (SLC2A2)* is not the principal glucose transporter in human pancreatic beta cells: implications for understanding genetic association signals at this locus. *Mol Genet Metab* 2011;104:648–653
63. Tornehave D, Kristensen P, Rømer J, Knudsen LB, Heller RS. Expression of the *GLP-1* receptor in mouse, rat, and human pancreas. *J Histochem Cytochem* 2008;56:841–851
64. Guiot Y, Stevens M, Marfour I, et al. Morphological localisation of sulfonylurea receptor 1 in endocrine cells of human, mouse and rat pancreas. *Diabetologia* 2007;50:1889–1899
65. Yan L, Figueroa DJ, Austin CP, et al. Expression of voltage-gated potassium channels in human and rhesus pancreatic islets. *Diabetes* 2004;53:597–607
66. Arden C, Harbottle A, Baltrusch S, Tiedge M, Agius L. Glucokinase is an integral component of the insulin granules in glucose-responsive insulin secretory cells and does not translocate during glucose stimulation. *Diabetes* 2004;53:2346–2352
67. Hutton JC, Eisenbarth GS. A pancreatic β -cell-specific homolog of glucose-6-phosphatase emerges as a major target of cell-mediated autoimmunity in diabetes. *Proc Natl Acad Sci U S A* 2003;100:8626–8628
68. Zito E, Chin KT, Blais J, Harding HP, Ron D. *ER01- β* , a pancreas-specific disulfide oxidase, promotes insulin biogenesis and glucose homeostasis. *J Cell Biol* 2010;188:821–832
69. Lena G, Reale P, Maggiora Vergano T, Anastasio M. [Fetal distress and spastic paralysis]. *Quad Clin Ostet Ginecol* 1967;22:755–766 [in Italian]
70. Blum B, Hrvatin S, Schuetz C, Bonal C, Rezanja A, Melton DA. Functional beta-cell maturation is marked by an increased glucose threshold and by expression of *urocortin 3*. *Nat Biotechnol* 2012;30:261–264
71. Depken M, Parrondo JM, Grill SW. Intermittent transcription dynamics for the rapid production of long transcripts of high fidelity. *Cell Reports* 2013;5:521–530
72. Wang L, Nie J, Sicotte H, et al. Measure transcript integrity using RNA-seq data. *BMC Bioinformatics* 2016;17:58
73. Blodgett DM, Nowosielska A, Afik S, et al. Novel observations from next-generation RNA sequencing of highly purified human adult and fetal islet cell subsets. *Diabetes* 2015;64:3172–3181
74. Chiang MK, Melton DA. Single-cell transcript analysis of pancreas development. *Dev Cell* 2003;4:383–393
75. De Krijger RR, Aanstoot HJ, Kranenburg G, Reinhard M, Visser WJ, Bruining GJ. The midgestational human fetal pancreas contains cells coexpressing islet hormones. *Dev Biol* 1992;153:368–375
76. Katsuta H, Akashi T, Katsuta R, et al. Single pancreatic beta cells co-express multiple islet hormone genes in mice. *Diabetologia* 2010;53:128–138
77. Herrera PL. Adult insulin- and glucagon-producing cells differentiate from two independent cell lineages. *Development* 2000;127:2317–2322
78. Macosko EZ, Basu A, Satija R, et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* 2015;161:1202–1214
79. Tabula Muris Consortium; Overall coordination; Logistical coordination; Organ collection and processing; Library preparation and sequencing; Computational data analysis; Cell type annotation; Writing group; Supplemental text writing group; Principal investigators. Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nature* 2018;562:367–372
80. Byrnes LE, Wong DM, Subramaniam M, et al. Lineage dynamics of murine pancreatic development at single-cell resolution. *Nat Commun* 2018;9:3922
81. Rosenberg AB, Roco CM, Muscat RA, et al. Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science* 2018;360:176–182
82. Rodrigues SG, Stickels RR, Goeva A, et al. Slide-seq: a scalable technology for measuring genome-wide expression at high spatial resolution. *Science* 2019;363:1463–1467
83. Cui Y, Hu D, Markillie LM, et al. Fluctuation localization imaging-based fluorescence in situ hybridization (flFISH) for accurate detection and counting of RNA copies in single cells. *Nucleic Acids Res* 2018;46:e7
84. Dominguez Gutierrez G, Kim J, Lee AH, et al. Gene signature of the human pancreatic ϵ cell. *Endocrinology* 2018;159:4023–4032
85. Saikia M, Burnham P, Keshavjee SH, et al. Simultaneous multiplexed amplicon sequencing and transcriptome profiling in single cells. *Nat Methods* 2019;16:59–62
86. Liu P, Song R, Elison GL, Peng W, Acar M. Noise reduction as an emergent property of single-cell aging. *Nat Commun* 2017;8:680
87. Büttner M, Miao Z, Wolf FA, Teichmann SA, Theis FJ. A test metric for assessing single-cell RNA-seq batch correction. *Nat Methods* 2019;16:43–49
88. Lawlor N, Khetan S, Ucar D, Stitzel ML. Genomics of islet (dys)function and type 2 diabetes. *Trends Genet* 2017;33:244–255
89. Setty M, Tadmor MD, Reich-Zeliger S, et al. Wishbone identifies bifurcating developmental trajectories from single-cell data. *Nat Biotechnol* 2016;34:637–645
90. Schwaib PC, Ordóñez-Morán P, Huelsken J, Deplancke B. Cross-tissue identification of somatic stem and progenitor cells using a single-cell RNA-sequencing derived gene signature. *Stem Cells* 2017;35:2390–2402
91. Farrell JA, Wang Y, Riesenfeld SJ, et al. Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. *Science* 2018;360:eaar3131
92. Semrau S, Goldmann JE, Soumillon M, Mikkelsen TS, Jaenisch R, van Oudenaarden A. Dynamics of lineage commitment revealed by single-cell transcriptomics of differentiating embryonic stem cells. *Nat Commun* 2017;8:1096
93. Baron CS, Kester L, Klaus A, et al. Single-cell transcriptomics reveal the dynamic of haematopoietic stem cell production in the aorta. *Nat Commun* 2018;9:2517
94. Scavuzzo MA, Hill MC, Chmielowiec J, et al. Endocrine lineage biases arise in temporally distinct endocrine progenitors during pancreatic morphogenesis. *Nat Commun* 2018;9:3356
95. Gasa R, Mrejen C, Leachman N, et al. Proendocrine genes coordinate the pancreatic islet differentiation program in vitro. *Proc Natl Acad Sci U S A* 2004;101:13245–13250
96. Sommer L, Ma Q, Anderson DJ. Neurogenins, a novel family of atonal-related bHLH transcription factors, are putative mammalian neuronal determination genes that reveal progenitor cell heterogeneity in the developing CNS and PNS. *Mol Cell Neurosci* 1996;8:221–241