Published online 20 November 2020

NAR Genomics and Bioinformatics, 2020, Vol. 2, No. 4 **1**
doi: 10.1093/nargab/lqaa097

Integration of single-cell datasets reveals novel transcriptomic signatures of β-cells in human type 2 diabetes

Emanuele Bosi[1], Lorella Marselli[1], Carmela De Luca[1], Mara Suleiman[1], Marta Tesi[1], Mark Ibberson[2], Decio L. Eizirik[3,4], Miriam Cnop[3,5] and Piero Marchetti[1,*]

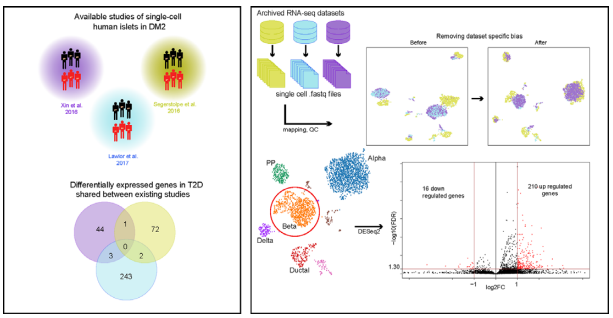
[1]Department of Experimental and Clinical Medicine, Pancreatic Islets Laboratory, University of Pisa, Pisa, I-56124, Italy, [2]Vital-IT Group, SIB Swiss Institute of Bioinformatics, University of Lausanne, Quartier Sorge, CH-1015 Lausanne, Switzerland, [3]ULB Center for Diabetes Research, Université Libre de Bruxelles, Brussels, B-1070, Belgium, [4]Indiana Biosciences Research Institute (IBRI), Indianapolis, IN 46202, USA and [5]Division of Endocrinology, Erasmus Hospital, Université Libre de Bruxelles, Brussels, B-1070, Belgium

Received April 27, 2020; Revised July 26, 2020; Editorial Decision October 26, 2020; Accepted October 30, 2020

ABSTRACT

Pancreatic islet β-cell failure is key to the onset and progression of type 2 diabetes (T2D). The advent of single-cell RNA sequencing (scRNA-seq) has opened the possibility to determine transcriptional signatures specifically relevant for T2D at the β-cell level. Yet, applications of this technique have been underwhelming, as three independent studies failed to show shared differentially expressed genes in T2D β-cells. We performed an integrative analysis of the available datasets from these studies to overcome confounding sources of variability and better highlight common T2D β-cell transcriptomic signatures. After removing low-quality transcriptomes, we retained 3046 single cells expressing 27 931 genes. Cells were integrated to attenuate dataset-specific biases, and clustered into cell type groups. In T2D β-cells (*n* = 801), we found 210 upregulated and 16 downregulated genes, identifying key pathways for T2D pathogenesis, including defective insulin secretion, SREBP signaling and oxidative stress. We also compared these results with previous data of human T2D β-cells from laser capture microdissection and diabetic rat islets, revealing shared β-cell genes. Overall, the present study encourages the pursuit of single β-cell RNA-seq analysis, preventing presently identified sources of variability, to identify transcriptomic changes associated with human T2D and underscores specific traits of dysfunctional β-cells across different models and techniques.

GRAPHICAL ABSTRACT



INTRODUCTION

The last decade showed a sharp increase in our ability to investigate whole transcriptomes at a high resolution. In parallel to the continuous improvements of sequencing platforms, the emergence of single-cell RNA sequencing (scRNA-seq) (1) made it possible to obtain transcript sequences out of individual cells, enabling to capture features of cellular differentiation, pathogenesis and adaptation (2–4), which would have been overlooked using bulk RNA-seq. The applications of such technology are very promising, especially for the study of heterogeneous tissues containing different cell types or the analysis of rare cells, in that it allows to characterize which genes are selectively expressed in different cell types, to reconstruct the trajectories of cell differentiation and response to stimuli (5,6) and to infer underlying regulatory networks (7). Altogether, scRNA-seq has the potential of filling knowledge gaps in our current understanding of how genetics and environmental factors affect the phenotype of single cells, and how these in turn influence the structure–function of tissues and organs (8).

*To whom correspondence should be addressed. Tel: +39 050 995110; Fax: +39 050996355; Email: piero.marchetti@med.unipi.it

The heterogeneous nature of pancreatic tissue makes it an excellent target to be analyzed with scRNA-seq. In fact, the organ is made up of a number of different cell types having either exocrine or endocrine secretory functions. Cells belonging to the latter category are found in the islets of Langerhans, which are cell clusters predominantly composed of α, β, δ and PP cells that secrete glucagon, insulin, somatostatin and pancreatic polypeptide, respectively. The β-cells are the sole source of insulin produced in the human body, and are therefore strictly implicated in the onset and progression of type 2 diabetes (T2D) (9,10). Therefore, the molecular and physiological characterization of β-cells in T2D is central for the identification of specific pathways and functions associated with their failure, which could provide novel insights into T2D pathophysiology for better prevention and treatment of this disease. Importantly, β-cells are probably heterogeneous (11,12), which may affect how putative β-cell subpopulations respond to the predisposing genetic background and metabolic stresses leading to T2D.

So far, scRNA-seq has been applied to human islets from non-diabetic (ND) and T2D donors in three valuable independent studies (13–15) in an effort to identify differentially expressed genes (DEGs) in T2D. A comparison of the sets of DEGs in β-cells from these studies revealed, surprisingly, that not a single gene was shared (16). This discrepancy could be due to the complex etiology of T2D and the (relatively) limited number of donors analyzed; it should also be considered that these studies had different experimental and analytical steps, from the isolation of single cells to the computational analysis of sequenced reads, which inevitably add technical sources of variability that can confound biologically relevant data (17,18).

The single-cell field is witnessing an incredibly fast progression, with the establishment of toolkits such as Scanpy (19) or Seurat (20) that enable the seamless implementation of standardized analytical workflows to scRNA-seq data. This, coupled with the definition of better guidelines and standards (21), not only makes it easier to integrate datasets within a single analytical design to correct for study-specific bias (22), but also removes the influence of technical biases arising from different computational tools and algorithms.

In this study, we aimed to deliver a comprehensive picture of the human pancreatic single β-cell transcriptomes in T2D. To do so, we integrated the three major scRNA-seq studies of human islets in a single dataset that was then analyzed by focusing on β-cells to identify shared DEGs and pathways to reconcile the identified features of T2D β-cells with the current biological knowledge of this condition (23–25). To evaluate the consistency of our findings, we also compared our results with those of (i) another study of β-cells from T2D patients and controls, based on an orthogonal methodology, namely laser capture microdissection (LCM) (26), and (ii) islets from a rat model of pancreatectomy-induced hyperglycemia (27).

MATERIALS AND METHODS

Analysis of sequencing data and dataset integration

The fastq files from the three studies re-analyzed in this work were downloaded using SRA toolkit (https://ncbi.github.io/sra-tools/) for the projects archived in SRA (28) (SRP075377 and SRP075970), or custom bash script for the one deposited in ArrayExpress (29) (E-MTAB-5061). Metadata reporting information for each cell, including donor ID, gender, body mass index and diabetic condition, were downloaded as well from the respective repositories. For one dataset (14), we excluded cells not having values of the quality metadata as 'OK'.

The reads were aligned against the human reference genome GRCh37 (Ensembl 87 annotation) using STAR 2.7.3 (30) with '–quantMode TranscriptomeSAM GeneCounts', obtaining for each study a table reporting per-gene read counts of each cell. From this point onward, all downstream analyses were conducted using ad-hoc Python scripts implementing functions from the toolbox Scanpy (19). The read count files were integrated with the cell metadata and the Ensembl annotation to produce three AnnData files, which were used to perform cell-wise quality control (QC) analyses.

Defining the number of read counts per sample as 'counts', the number of genes with at least one read mapped as 'expressed genes' and the ratio of reads mapped on mitochondrial genes as 'mitochondrial fraction', we considered counts, expressed genes and mitochondrial fraction as technical covariates defining the quality of each cell. Specifically, cells with relatively high number of counts and genes are likely representing multiplets, i.e. two or more cells captured and sequenced assuming a single cell, whereas high mitochondrial fraction and low expressed genes are indicative of lysed cells. We considered the distribution of these variables and their covariation separately for each dataset, defining separate threshold values that allowed to identify and flag cells as 'low quality', which were then excluded from downstream analysis. Contemporary with cell-level QC, genes expressed in less than three cells or expressed only in a single dataset were not considered for downstream analyses.

After QC, the datasets were concatenated and the counts were normalized by scaling the count values to obtain a total count of 10 000 for each cell and then log transformed. This normalized dataset was analyzed to compute the dispersion of each gene with respect to its mean value to annotate genes as highly variable using the 'highly_variable_genes' function of Scanpy. This allowed to define a set of 971 genes displaying a high variability in each single dataset, which was used to perform dataset integration with mutual nearest neighbors (MNN) algorithm (22) as implemented in mnnpy (https://github.com/chriscainx/mnnpy). Visualization of single cells before/after MNN correction was performed with uniform manifold approximation projection (UMAP) as implemented in Scanpy.

Unsupervised clustering and cell type annotation

Detection of single-cell communities was done using the Louvain modularity algorithm (31) implemented in Scanpy (https://github.com/vtraag/louvain-igraph) with resolution = 0.5. The relative contribution of genes in separating clusters was computed with the 'rank_genes_groups' function of Scanpy, manually evaluating their association with major pancreatic cell types using literature information and gene expression markers reported in PanglaoDB (32).

Differential expression in T2D and enrichment analyses

DEGs in T2D β-cells were identified using DESeq2 (33) with the following design: *Counts* ~ *Dataset* + *Diabetes*, where *Counts* is the matrix of raw count data, *Dataset* is a three-level factor (SEG, XIN, LAW) indicating the dataset of origin and *Diabetes* is a two-level factor (T2D, ND) indicating the diabetes status of the donor. Genes were considered as DEG if passing these thresholds: <0.05 for corrected *P*-value (false discovery rate, FDR) and >2 for the absolute value of fold change (FC). DEG identification for individual datasets was performed similarly, with the following design: *Counts* ~ *Diabetes*.

Gene set enrichment analysis (GSEA) was performed using Enrichr with the following datasets: the Gene Ontology 2018 (GO) subsets Biological Process, Molecular Function and Cellular Component (34,35); BioPlanet 2019 (36); KEGG 2019 (*Homo sapiens*) (37–39); Reactome 2016 (40); and a consensus of transcription factor target genes from Encode and ChEA. All gene sets are available at https://amp.pharm.mssm.edu/Enrichr/#stats.

Overlapping expression trends were identified with two-tailed rank–rank hypergeometric overlap (RRHO) (41,42) of genes ranked according to −log(*P*-value) × *direction*, where *direction* is the sign of the expression change in T2D versus ND. A custom R script implementing the RRHO package (v. 1.24.0) was used to compute and analyze the hypergeometric distribution matrix, which allowed us to identify genes corresponding to overlapping expression trends.

Coding and data visualization

To visualize the obtained results, ad-hoc Python scripts were used combining Scipy (43), Pandas, Matplotlib (44), Seaborn and Scanpy (19). The volcano plot of DEGs was produced with R (https://www.r-project.org/). The analytical workflow was organized using a Jupyter notebook.

RESULTS

A merged scRNA-seq dataset of islets from T2D patients

We obtained raw sequencing data of three studies: E-MTAB-5061 (14), SRP075377 (13) and SRP075970 (15) (hereafter referred to as SEG, XIN and LAW, respectively) (Figure 1A), whose deposited sequences represent the transcriptome of single cells passing QC checks as defined by the respective authors. Read count matrices reporting the relative gene expression in each cell were also available, but we decided not to use them since different approaches were used to obtain them from raw data. By re-analyzing the original reads with the same pipeline, we removed variability coming from usage of different tools.

As shown in Table 1, the datasets differed in terms of donor selection, single-cell isolation and sequencing library preparation, but also in QC criteria for expressed genes, total read counts and mitochondrial fraction (Figure 1B). For the number of expressed genes, we found a unimodal distribution for XIN and LAW with a peak around 5500, whereas SEG showed a bimodal distribution with peaks around 4000 and 8000. We also found SEG to display a higher variance for total read counts, having a high number of cells with extremely low read counts and outliers with expression up to 6 million reads, and high mitochondrial fraction, with outliers having values close to 1. XIN had no cells with mitochondrial fraction >0.25, reflecting differences in QC criteria used in the original studies.

Table 1. Main features of the datasets used

Dataset	Number of donors (T2D, ND)	Number of cells	Read length	Average million read count (std)
SEG	10 (4, 6)	2209	43	0.558 (0.525)
XIN	18 (6, 12)	1600	75	1.145 (0.551)
LAW	24 (9, 15)	638	75	1.689 (0.653)

The table reports the different features of the datasets used, in terms of number of donors (total, T2D and ND), number of cells with available raw sequencing data, length of sequencing reads and the average number of reads (million) per cell.

We excluded cells with signatures of low viability (i.e. low count depth, high fraction of mitochondrial genes) or multiplets, defined as multiple cells sequenced and labeled as a single cell. We considered each dataset separately to define threshold values for metrics such as number of expressed genes, total read counts and fraction of read counts on mitochondrial genes (see Table 2). We also excluded outlier genes, defined as those expressed in less than three cells or being present in a single dataset.

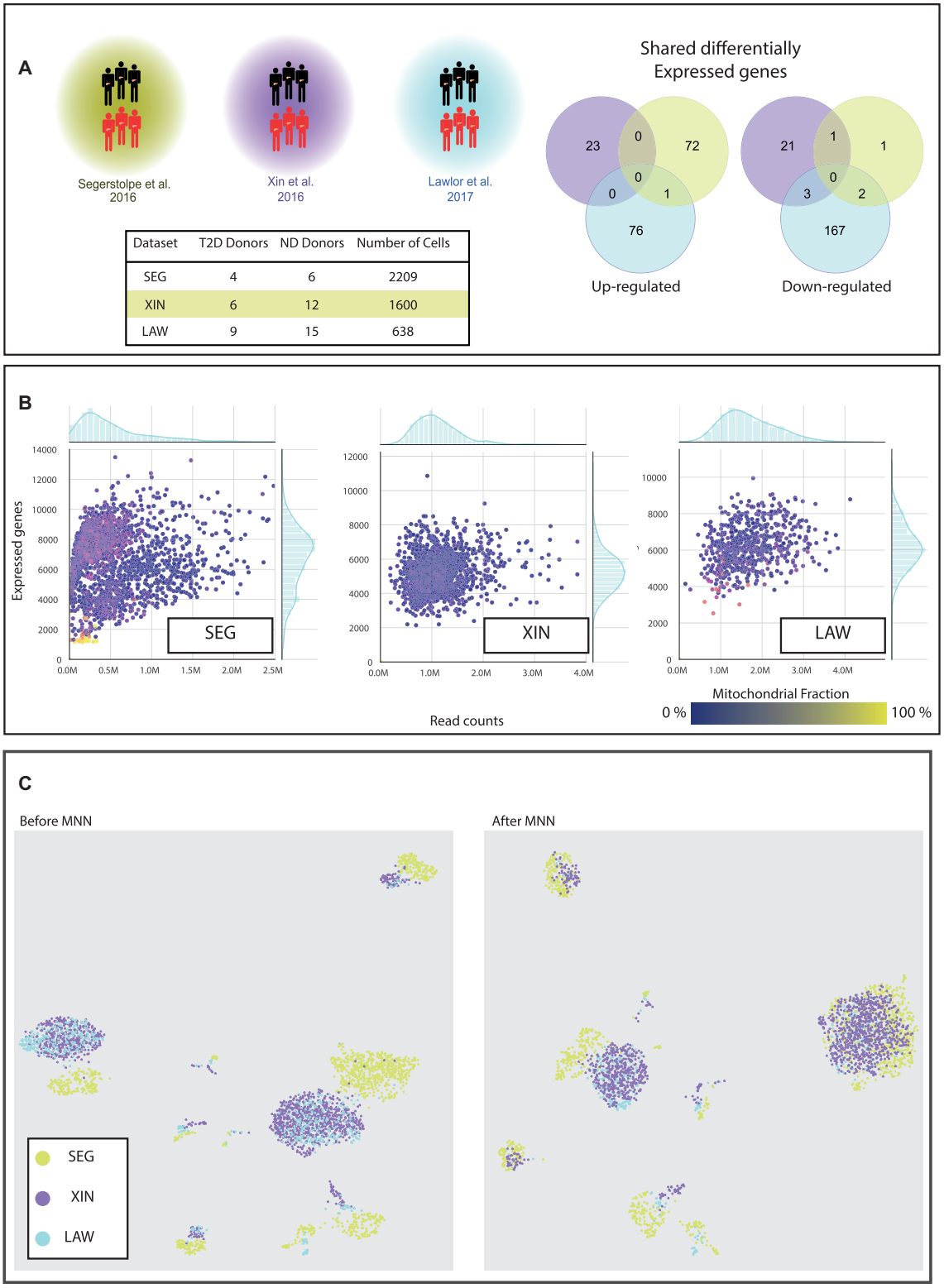
After removing the cells not satisfying the QC criteria, we concatenated the data into a single merged dataset, containing (i) a gene count matrix embedding 3046 cells and 27 931 genes, and (ii) a metadata matrix with the ancillary information available for each cell and donor. The dataset is available as an h5ad file (see the 'Data availability' section).

Normalization of dataset effects and cell-level analyses

Considering that our integrated dataset contains data from three laboratories using different protocols and technologies, gene expression might present systematic differences due to batch effects. In order to correct this technical source of variation, we applied a data integration method (MNN correction) on our dataset to better highlight biological features shared between cells. To visualize the effects of such correction, we identified a set of 971 highly variable genes, which were used for UMAP two-dimensional representation of the cells labeled according to the dataset of origin.

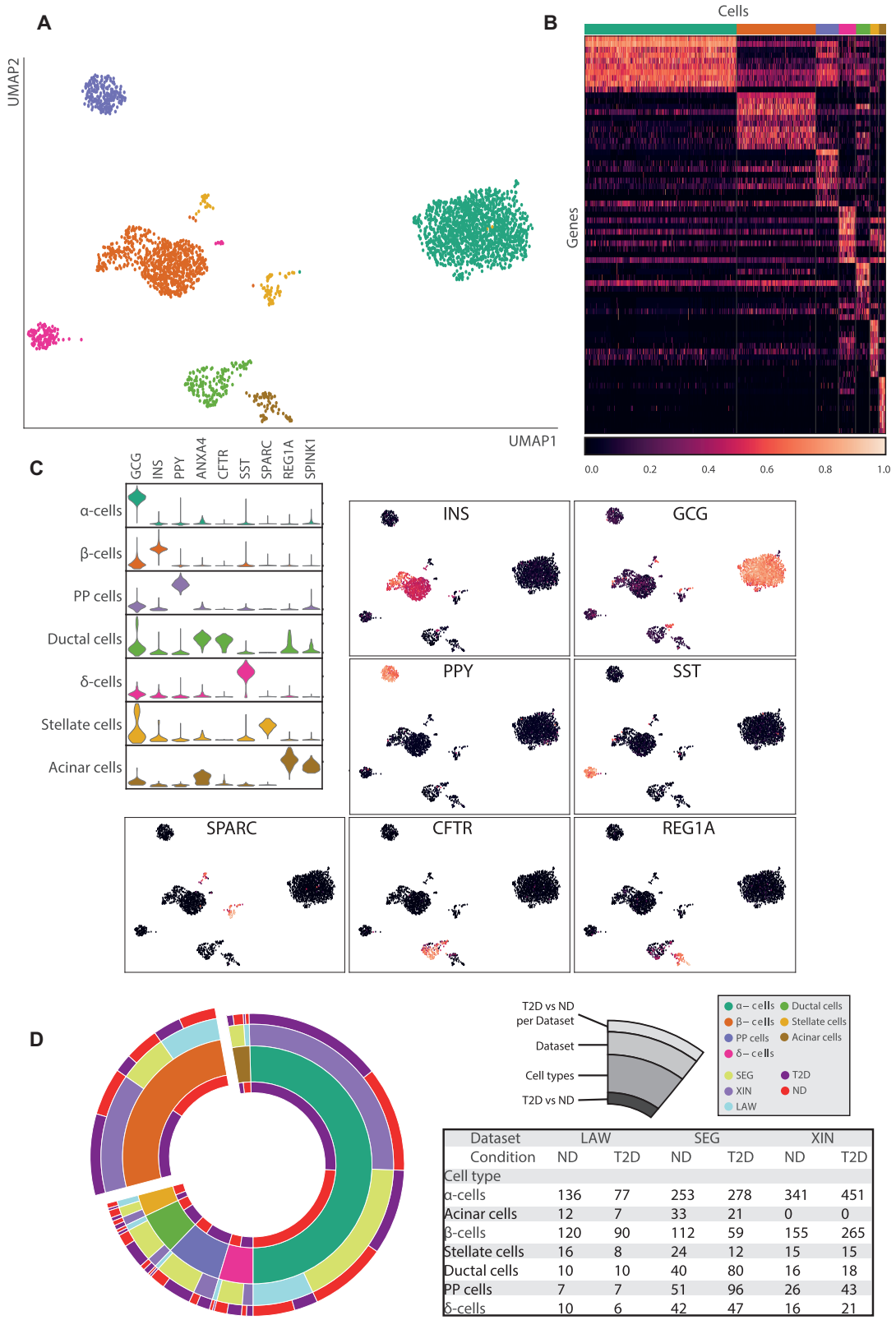
This visualization (Figure 1C) shows clearly separated groups of cells with similar transcriptomes, with LAW and XIN exhibiting remarkable overlap, whereas SEG is more distinct. After MNN correction (Figure 1C), there was a larger overlap of SEG with the other datasets, indicating a reduction of the dataset effect on the transcriptomic differences between cells.

We next performed unsupervised clustering of the cells based on their corrected gene expression profiles, obtaining seven different cell clusters, i.e. α, β, δ, PP, ductal, stellate and acinar cells (Figure 2A). From the transcriptomic signatures of the identified groups, we selected genes maximizing the diversity between groups to annotate their identity (Figure 2B and Supplementary Figure S1). This led to the identification of nine marker genes (GCG, INS, PPY, SST, ANXA4, CFTR, SPARC, REG1A and SPINK1) whose



Dataset	T2D Donors	ND Donors	Number of Cells
SEG	4	6	2209
XIN	6	12	1600
LAW	9	15	638

Figure 1. An integrated dataset of single-cell transcriptomes from studies of human islets (T2D versus ND). (**A**) The three studies considered in this work are Segerstolpe (SEG) (14), Xin (XIN) (13) and Lawlor (LAW) (15) that analyzed the indicated total number of cells from the indicated number of T2D and ND donors. The Venn diagram recapitulates the previous comparison of DEGs (by Wang and Kaestner) (16), showing paucity of shared DEGs. (**B**) The plots report the number of expressed genes, total read counts and the mitochondrial enriched fraction of the cells of each dataset. Extreme values of these parameters indicate low-quality cells, as cells with high read counts and expressed genes might represent multiplets of captured cells, while a high fraction of mitochondrial genes expression is indicative of lysed cells (21). The histograms on the axes of each plot represent the marginal distributions of the corresponding variables (i.e. read counts: top; expressed genes: right). (**C**) UMAP visualization of the integrated dataset before (left) and after (right) the normalization with MNN. Cells are colored according to the dataset of origin to highlight the effect of the MNN normalization.



Dataset	LAW		SEG		XIN	
Condition	ND	T2D	ND	T2D	ND	T2D
Cell type						
α-cells	136	77	253	278	341	451
Acinar cells	12	7	33	21	0	0
β-cells	120	90	112	59	155	265
Stellate cells	16	8	24	12	15	15
Ductal cells	10	10	40	80	16	18
PP cells	7	7	51	96	26	43
δ-cells	10	6	42	47	16	21

Figure 2. Analysis of the integrated dataset at the cell level. (**A**) The UMAP visualization of the integrated dataset (after MNN) with cells color labeled according to the clusters assigned with the unsupervised method Louvain (31). (**B**) The heatmap reports the normalized expression of the 10 most representative genes of each cluster (rows) in each cell (columns). Columns are color labeled according to cell clusters. (**C**) Distribution of the expression of marker genes for each cell type (violin plots). Normalized expression of the most representative marker genes in each cell with a UMAP visualization. (**D**) The sections in the circular plot represent the cell counts stratified in different categories. From outer to inner: donor diabetes condition (T2D, ND) per dataset and cell type, dataset (SEG, XIN, LAW), cell types and donor diabetes condition (T2D, ND) per cell type with aggregated datasets.

Table 2. QC filtering criteria

Dataset	Number of cells before QC	Number of good quality cells	Number of genes before QC	Number of genes after filtering	Range of expressed genes	Range of total read count (million)	Maximum mitochondrial fraction
SEG	2209	1148	33 466	27 931	2500–9000	0.3–2	0.4
XIN	1600	1382	33 466	27 931	2500–7000	0.5–2.2	–
LAW	638	516	31 511	27 931	2500–9000	1–3	0.4

The table reports for each dataset the number of cells and genes before and after the QC filtering, as well as the threshold values used to identify good quality cells. Ranges are pairs of values indicating the lower and the upper bound thresholds.

expression is associated with specific pancreatic cell types (Figure 2C). This information allowed us to characterize the cell type composition of our dataset at different hierarchical resolution (aggregated, dataset and individual level) and related to disease state (T2D versus ND), testing the enrichment of cell types in each condition. We observed substantial differences in the abundance of cell types between different datasets (Figure 2D). SEG had half the proportion of β-cells of other datasets (15% versus 30% and 41%), and substantially higher ductal, δ and PP cells. Acinar cells were missing in XIN.

Differentially expressed genes in T2D β-cells

A previous study comparing scRNA-seq transcriptomic signatures of T2D showed no shared DEGs (16). To assess how much of this variability is effectively due to differences in data processing, we identified DEGs separately for each dataset and compared them. Since no DEGs are shared between datasets (Supplementary Figure S2), variability observed is probably due to different sample source, preparation and other experimental steps rather than data processing.

Using our merged dataset, we regressed out bias due to the dataset of origin. Comparing the T2D versus ND β-cell subpopulations, we identified 226 DEGs (FDR $\leq$ 0.05, FC $\geq$ 2), with 210 upregulated and 16 downregulated genes in T2D (see Figure 3A). Of these, 60 were protein-coding genes (Figure 3B), which were manually curated based on the available literature and databases (i.e. UniProt, STRING, KEGG and GWASdb) to explore their pathophysiological role (Supplementary Table S1). For 35 DEGs, their function could be related to T2D, including β-cell failure mechanisms, such as defective insulin secretion, increased oxidative stress, altered autophagy and apoptosis (24,25,45–47) (Figure 3B). The remaining 25 genes have not been described previously. Of these, 16 genes have an undefined function, whereas the other 9 could be ascribed to cellular processes linked to β-cell dysfunctions (Supplementary Table S1). These include CABIN1, CKS1B, C19orf60 and SDR39U1 (affecting cell survival), SLC31A1 (involved in copper homeostasis), DNAJA4 (associated with ER stress), ZC3H8 (regulating the expression of GATA3), OTUD3 and UBALD1 (affecting ubiquitination) (48–58).

We then performed GSEA on seven datasets using Enrichr. The results of this analysis showed an enrichment of several categories that can be associated with altered β-cell pathways (Figure 3C and Supplementary Table S2), including the control of hydrogen peroxide and respiratory burst (GO:0010310, GO:0060263), the activity of NADPH oxidase for ROS generation (GO:0016176) and the ionotropic glutamate receptor (GO:0008328). Of interest, a pathway that appeared positively enriched in multiple datasets is related to lysosome function, which is associated with autophagy (59).

Generalizing common transcriptomic signatures of T2D β-cells across different methodologies

To verify the extent to which our results may be generalizable, we compared the gene expression signatures of the present integrated transcriptomic dataset with those from two other models. We used (i) the results obtained by microarray gene expression analysis of human T2D and ND β-cell enriched samples yielded by LCM (60), as a model similar to the human single β-cell approach, and (ii) the data recently generated by islet RNA-seq assessment in 90% pancreatectomized, hyperglycemic rats (27), as a less close model.

In a first set of analyses, we used the RRHO approach (41) that allows to compare differentially expressed transcriptomes between independent studies in a threshold-free way and visualize both the significance and direction of the possible overlays. In the work by Marselli *et al.* (26), hereafter referred to as MAR, the analysis of 10 T2D and 10 ND β-cell enriched samples identified 1742 DEGs utilizing a significance threshold of lower confidence bound $\geq$1.2, corresponding to 1086 upregulated and 656 downregulated genes. The comparison of the transcriptomes of MAR T2D versus ND with those of our integrated T2D versus ND single β-cell transcriptome assessment is reported in Figure 4A. The analysis revealed that although there were genes regulated in opposite direction, 191 transcription signatures overlapped between the two studies (173 upregulated and 18 downregulated). Among them, there were 20 genes with *P*-value $\leq$0.05, of which 11 have been previously described in association with T2D traits (61–71) (Supplementary Table S3).

In the study by Ebrahimi *et al.* (27), hereafter referred to as EBR, the authors characterized rat islet transcriptomic changes following 90% pancreatectomy. At 10 weeks after surgery, pancreatectomized, hyperglycemic animals showed many islet transcriptome changes in comparison with controls. There were 7844 DEGs, many of which associated with glucose toxicity, stress, inflammation and β-cell identity. The RRHO comparison of EBR data with the present single β-cell transcriptome results revealed that there were 1014 common genes that were upregulated (Figure 4B). Among them, 118 genes were significantly ($P \leq 0.05$) regulated in both datasets (Supplementary Table S3), and 61

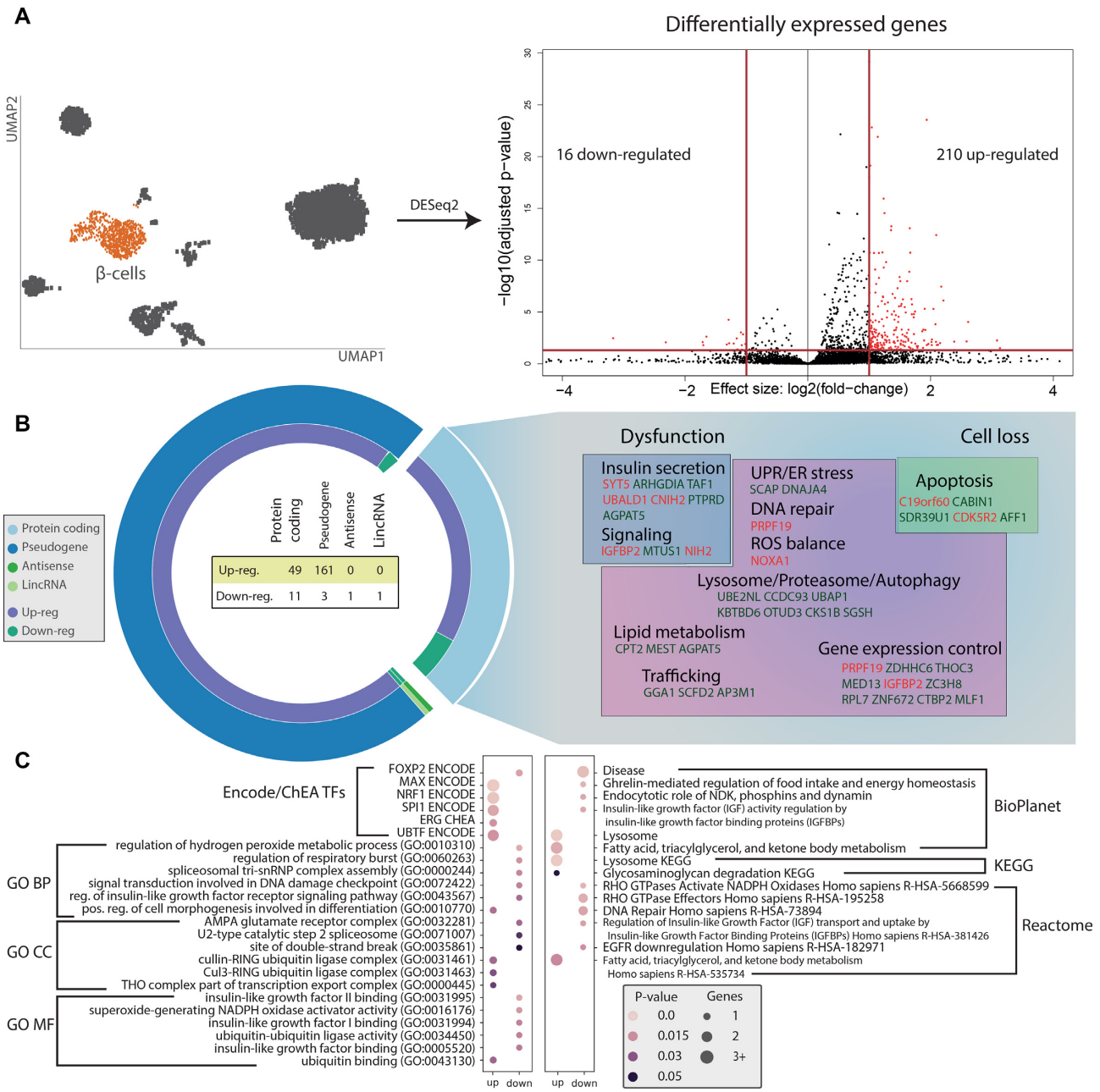


Figure 3. Analyses of the integrated dataset at the gene level. (**A**) Volcano plot reporting the significance and FC values obtained for each β-cell gene with DESeq2. The horizontal and vertical red lines report the threshold used to define DEGs (shown in red). (**B**) Classification of DEGs according to biotype (outer circle) and direction of change (inner circle). The gene symbols associated with T2D-related functions are reported colored differently according to the corresponding expression change direction (green: upregulated; red: downregulated). (**C**) Terms enriched in the DEG set. For each term, the number of associated genes (dot size) and enrichment significance (color scale) are reported. The reported terms correspond to the six most significant terms for each dataset.

of these genes have been previously linked to T2D features (72–126) (Supplementary Table S3).

Next, we compared the three datasets (integrated single cells, MAR and EBR) to identify shared genes (Figure 4C). We considered only genes with $P \leq 0.05$ and found a set of 208 genes, of which 41 and 8 were, respectively, upregulated and downregulated in all three datasets (Supplementary Table S4). Interestingly, several of such genes (20 upregulated and 4 downregulated) have been reported to be linked to diabetes (74,76,91–93,111,112,126–138) (Supplementary Table S4).

DISCUSSION

The present work was prompted by the observation that, comparing the results of the three available studies of pancreatic β-cells from individuals affected by T2D and ND controls (13–15), scRNA-seq 'failed' to deliver a shared

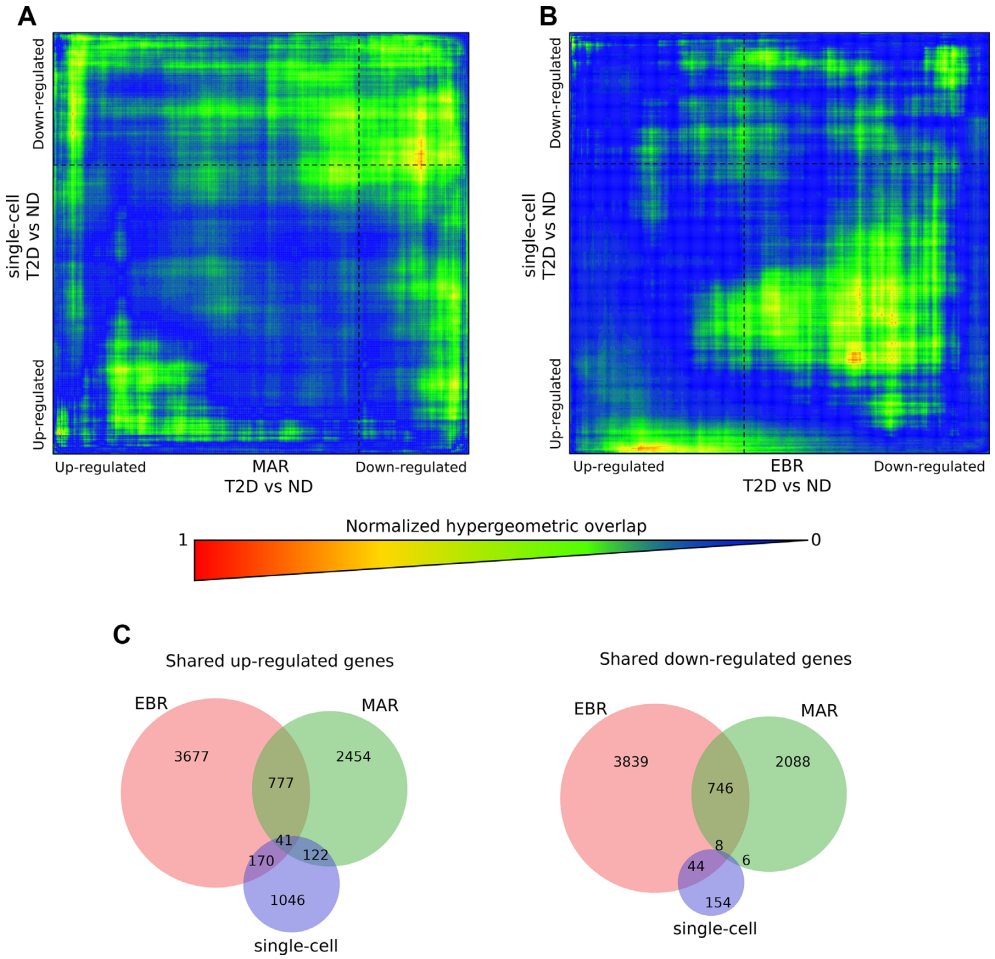


Figure 4. Shared transcriptomic patterns between single β-cells, MAR and EBR. (**A**) RRHO map showing overlap between the single-cell differential expression of T2D versus ND and the differential expression of T2D versus ND in MAR. A detailed description of the MAR dataset is provided in the text. Genes are ranked by FC from most downregulated to most upregulated. The level map colors show normalized −log *P*-values for overlap, with an indication of the smallest *P*-value for clusters with statistically significant overlap between genes upregulated in both datasets (bottom left quadrant), downregulated in both (top right quadrant), upregulated in MAR T2D and downregulated in single-cell T2D (top left quadrant), and downregulated in MAR T2D and upregulated in single-cell T2D (bottom right quadrant). (**B**) RRHO map showing overlap between the single-cell differential expression of T2D versus ND and the differential expression of T2D versus ND in EBR. A detailed description of the EBR dataset is provided in the text. Genes are ranked by FC from most downregulated to most upregulated. The level map colors show normalized −log *P*-values for overlap, with an indication of the smallest *P*-value for clusters with statistically significant overlap between genes upregulated in both datasets (bottom left quadrant), downregulated in both (top right quadrant), upregulated in EBR T2D and downregulated in single-cell T2D (top left quadrant), and downregulated in EBR T2D and upregulated in single-cell T2D (bottom right quadrant). (**C**) Differentially regulated genes shared in the present integrated single β-cell dataset, MAR and EBR data. The Venn diagrams indicate the number of DEGs ($P \leq 0.05$) shared between the single-cell dataset, MAR and EBR. The diagrams separate upregulated (left) and downregulated (right) genes.

view of T2D-associated transcriptomic alterations of β-cells from human islets, possibly due to a number of methodological issues (16). Here, we show that an integrated analysis of the three datasets, based on recently developed algorithms and computational frameworks specific for single-cell transcriptomics (19,21,22), allowed us to identify genes that were differentially expressed in T2D versus ND β-cells, with potential pathophysiological roles.

These new tools mentioned above have made it possible to integrate different datasets and deal with technical covariates, allowing the re-analysis of published data to obtain novel biological insights (22). In particular, the pan-transcriptome we obtained by merging the published scRNA of human β-cells (13–15) has a larger sample size, hence is more robust against biases arising from donor-specific sources of biological variability (including a supposed effect of multiple and various T2D etiologies). For instance, this approach allowed us to reconcile discrepancies between the results of the three different studies. As an example, the genes TUBA1B and LEPROTL1 were downregulated in T2D according to LAW, upregulated according to SEG and not significantly different according to XIN; according to our analysis, these genes are not differentially expressed.

We strived to minimize technical sources of variabilities that could affect our analyses. Indeed, the comparison of QC metrics revealed a divergence between datasets at the cellular level that likely reflect separate selection criteria operated by the respective authors. Briefly, the original studies were not uniform concerning the minimal quality level of

cells to be subjected to downstream analyses, with effects that include a different heterogeneity within cell types observed by the authors. For instance, Segerstolpe *et al.* (14) reported heterogeneity for the β-cell group, whereas Xin *et al.* (13) did not. To obtain cells with comparable quality levels and reduce this bias, we adopted a conservative approach using different thresholds for each dataset to remove low-quality cells. Even with such harmonization, UMAP visualization of transcriptomes revealed a separation of cells according to the dataset of origin, implying that differences in experimental procedures, from islet culturing to RNA-seq library preparation and sequencing, impact the islet cell transcriptomes. To minimize this effect, we used MNN (22,139), but standardized procedures and materials in future studies will increase reproducibility.

Our integrated dataset identified 226 DEGs in β-cells associated with T2D, most of which were overexpressed. Since we used the T2D condition to contrast the transcriptomes, we expected some DEGs to recapitulate known signatures of diabetes. Indeed, among the 60 differentially expressed protein-coding genes there were 35 genes with functions related to β-cell damage, such as impaired insulin secretion, increased oxidative stress, deranged autophagy and apoptosis (24,25,45–47) (Figure 3B). In addition, we identified nine genes (CABIN1, CKS1B, C19orf60, SDR39U1, SLC31A1, DNAJA4, ZC3H8, OTUD3 and UBALD1) not previously associated with diabetes, but that are known to be involved in processes potentially linked to β-cell dysfunction, such as cell turnover, oxidative stress and ER stress (Supplementary Table S1). Hence, the presence of genes with a known relevance in the context of T2D (24,25,45–48,51–53,140–152) provides a confirmation of the validity of the approach we used.

In addition to this, we found 25 genes with no previous association with T2D described in the literature. Among the criteria used to associate novel genes with a potential β-cell role, we considered the presence of interactions with known genes. 'Synthetic lethals' are gene pairs for which the deletion/inactivation of a single gene has no major phenotypic effects, whereas the inactivation of both genes produces a lethal phenotype. This kind of genetic interaction allowed us to link C19orf60 and SDR39U1 with MYC and RAS (52), respectively, which are relevant for β-cell differentiation, proliferation and apoptosis (51,53). We also found novel genes displaying proven functional interactions. For instance, DNAJA4 has been shown in human colon cancer cells to be regulated by SREBP and act as a mediator of lipotoxicity through ER stress (55). UBALD1 and OTUD3 are enzymes involved in ubiquitination with an experimentally validated interaction with MLYCD and PTEN, respectively, and potentially affecting β-cell metabolic pathways, function and turnover (153). The protein encoded by CKS1B binds SKP2, increasing the activity of the E3 ubiquitin ligase Skp1–Cullin-1–Skp2 that degrades p27 (50), a pathway shown to be involved in regulating β-cell mass and function with implications for T2D development (49). Other novel transcripts that we found to be differentially expressed in diabetic β-cells (such as the zinc finger protein ZC3H8 and CABIN1) are involved in mechanisms potentially associated with β-cell dysfunction, such as the regulation of intracellular calcium signaling (56,154,155). Therefore, our integrative analysis partly reconciled the previous fully inconsistent results reported with β-cell scRNA-seq.

We then assessed how the results of our meta-analysis compared with those from a relatively similar human dataset (microarray of laser capture microdissection of T2D and control islets) (26) and a less similar rat model (RNA-seq of whole islets following hyperglycemia induced by partial pancreatectomy) (27). Using the overlapping transcriptomic signature (RRHO) approach, we observed that there were several genes differentially expressed in the same direction in our results and those from MAR (60), where the authors used β-cell enriched preparation from ND and T2D donors. Interestingly, islets from hyperglycemic pancreatectomized rats (27) showed several changes in gene expression similar to those in our integrated human β-cell datasets. Furthermore, based on the analysis of genes differentially expressed ($P \leq 0.05$) in the three datasets, we observed 49 shared genes (48 upregulated). Therefore, despite remarkable differences between the three models (single human β-cells yielded after islet digestion and separated by FACS or a microfluidic platform; β-cell enriched preparations obtained by LCM from the pancreas of organ donors; isolated islets from rats with surgically induced hyperglycemia; use of microarray or RNAseq), a set of shared DEGs remains associated with the dysfunctional β-cell across biological models and experimental techniques. These genes might represent key factors involved in the trajectory of β-cell failure (see Supplementary Table S4). For instance, upregulation of LDHA has a demonstrated involvement in perturbed insulin secretion (92), and overexpression of RPS10 is a marker of a functionally immature phenotype (93). Among the downregulated genes, ATP2A2 and PGRMC1 have a role in insulin secretion (127,131). Other shared genes without previously illustrated associations could provide novel insights into signatures of β-cell dysfunction. As an example, the gene PELP1 encodes a coactivator involved in a number of signaling pathways, including SRC/PI3K/AKT and ERK/MAPK, that can be relevant for defective insulin secretion (129,130).

In conclusion, this work represents the first integration of human islet single-cell transcriptomes to understand β-cell dysfunction in human T2D. The dataset we assembled (available at https://github.com/EBosi/scPanBetaT2D) has allowed to (i) partly reconcile the previously reported inconsistencies in single-cell analysis of human islet cells, (ii) identify novel genes to be investigated in future studies to better characterize the molecular basis of T2D onset and progression, and (iii) underscore specific traits of dysfunctional β-cells across different models and techniques.

DATA AVAILABILITY

The ID numbers of the gene expression series used in this study are E-MTAB-5061 (14), SRP075377 (13), SRP075970 (15), GSE20966 (60) and GSE134966 (27).

The source Python and R code, the intermediary files and the h5ad file encoding the integrated dataset are available on https://github.com/EBosi/scPanBetaT2D.

SUPPLEMENTARY DATA

Supplementary Data are available at NARGAB Online.

ACKNOWLEDGEMENTS

We are grateful to Dr Roberto Semeraro for insightful discussions and comments and to the referees for detailed insights and suggestions to improve the text.

FUNDING

This work was supported by non-profit organizations and public bodies for funding of scientific research conducted within the European Union: the Innovative Medicines Initiative 2 Joint Undertaking, RHAPSODY [115881 to EB, LM, CdL, MS, MT, MI, DLE, MC, PM], INNODIA [115797 to LM, DLE, MC, PM] and INNODIA HARVEST [945268 to LM, DLE,MC, PM] - this Joint Undertaking receives support from the Union's Horizon 2020 research and innovation programme, 'EFPIA', 'JDRF' and 'The Leona M. and Harry B. Helmsley Charitable Trust" (INNODIA, INNODIA HARVEST), the 'EFPIA' and the Swiss State Secretariat for Education, Research and Innovation under contract number 16.0097 (RHAPSODY); the European Union's Horizon 2020 research and innovation programme, project T2DSystems [667191 to EB, LM, CdL,MS,MT, DLE,MC, PM]; the Walloon Region through the FRFS-WELBIO Fund for Strategic Fundamental Research [grant numbers CR-2015A-06s, CR-2019C-04 to DLE]; the Welbio-Fonds National de la Recherche Scientifique, Belgium and Dutch Diabetes Fonds, Holland [2018.10.002 to DLE]; the Brussels CapitalRegion-Innoviris project Diatype [2017-PFS-24 to DLE, MC]; Fonds National de laRecherche Scientifique, Belgium [to MC].

Conflict of interest statement. None declared.

REFERENCES

1. Tang,F., Barbacioru,C., Wang,Y., Nordman,E., Lee,C., Xu,N., Wang,X., Bodeau,J., Tuch,B.B., Siddiqui,A. *et al.* (2009) mRNA-seq whole-transcriptome analysis of a single cell. *Nat. Methods*, **6**, 377–382.
2. Jaitin,D.A., Kenigsberg,E., Keren-Shaul,H., Elefant,N., Paul,F., Zaretsky,I., Mildner,A., Cohen,N., Jung,S., Tanay,A. *et al.* (2014) Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science*, **343**, 776–779.
3. Peng,J., Sun,B.-F., Chen,C.-Y., Zhou,J.-Y., Chen,Y.-S., Chen,H., Liu,L., Huang,D., Jiang,J., Cui,G.-S. *et al.* (2019) Single-cell RNA-seq highlights intra-tumoral heterogeneity and malignant progression in pancreatic ductal adenocarcinoma. *Cell Res.*, **29**, 725–738.
4. Shalek,A.K., Satija,R., Shuga,J., Trombetta,J.J., Gennert,D., Lu,D., Chen,P., Gertner,R.S., Gaublomme,J.T., Yosef,N. *et al.* (2014) Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. *Nature*, **510**, 363–369.
5. Haghverdi,L., Büttner,M., Wolf,F.A., Buettner,F. and Theis,F.J. (2016) Diffusion pseudotime robustly reconstructs lineage branching. *Nat. Methods*, **13**, 845–848.
6. Rabineau,P.A. and Flammer,J. (1989) [Sources of error in automated static perimetry]. *Klin. Monbl. Augenheilkd.*, **194**, 307–319.
7. Mahata,B., Zhang,X., Kolodziejczyk,A.A., Proserpio,V., Haim-Vilmovsky,L., Taylor,A.E., Hebenstreit,D., Dingler,F.A., Moignard,V., Göttgens,B. *et al.* (2014) Single-cell RNA sequencing reveals T helper cells synthesizing steroids *de novo* to contribute to immune homeostasis. *Cell Rep.*, **7**, 1130–1142.
8. Hwang,B., Lee,J.H. and Bang,D. (2018) Single-cell RNA sequencing technologies and bioinformatics pipelines. *Exp. Mol. Med.*, **50**, 96.
9. Halban,P.A., Polonsky,K.S., Bowden,D.W., Hawkins,M.A., Ling,C., Mather,K.J., Powers,A.C., Rhodes,C.J., Sussel,L. and Weir,G.C. (2014) β-Cell failure in type 2 diabetes: postulated mechanisms and prospects for prevention and treatment. *J. Clin. Endocrinol. Metab.*, **99**, 1983–1992.
10. Marchetti,P., Suleiman,M., De Luca,C., Baronti,W., Bosi,E., Tesi,M. and Marselli,L. (2020) A direct look at the dysfunction and pathology of the β cells in human type 2 diabetes. *Semin. Cell Dev. Biol.*, **103**, 83–93.
11. Avrahami,D., Klochendler,A., Dor,Y. and Glaser,B. (2017) Beta cell heterogeneity: an evolving concept. *Diabetologia*, **60**, 1363–1369.
12. Tritschler,S., Theis,F.J., Lickert,H. and Böttcher,A. (2017) Systematic single-cell analysis provides new insights into heterogeneity and plasticity of the pancreas. *Mol. Metab.*, **6**, 974–990.
13. Xin,Y., Kim,J., Okamoto,H., Ni,M., Wei,Y., Adler,C., Murphy,A.J., Yancopoulos,G.D., Lin,C. and Gromada,J. (2016) RNA sequencing of single human islet cells reveals type 2 diabetes genes. *Cell Metab.*, **24**, 608–615.
14. Segerstolpe,Å., Palasantza,A., Eliasson,P., Andersson,E.-M., Andréasson,A.-C., Sun,X., Picelli,S., Sabirsh,A., Clausen,M., Bjursell,M.K. *et al.* (2016) Single-cell transcriptome profiling of human pancreatic islets in health and type 2 diabetes. *Cell Metab.*, **24**, 593–607.
15. Lawlor,N., George,J., Bolisetty,M., Kursawe,R., Sun,L., Sivakamasundari,V., Kycia,I., Robson,P. and Stitzel,M.L. (2017) Single-cell transcriptomes identify human islet cell signatures and reveal cell-type-specific expression changes in type 2 diabetes. *Genome Res.*, **27**, 208–222.
16. Wang,Y.J. and Kaestner,K.H. (2019) Single-cell RNA-seq of the pancreatic islets—a promise not yet fulfilled? *Cell Metab.*, **29**, 539–544.
17. Ziegenhain,C., Vieth,B., Parekh,S., Reinius,B., Guillaumet-Adkins,A., Smets,M., Leonhardt,H., Heyn,H., Hellmann,I. and Enard,W. (2017) Comparative analysis of single-cell RNA sequencing methods. *Mol. Cell*, **65**, 631.e4–643.e4.
18. Mawla,A.M. and Huising,M.O. (2019) Navigating the depths and avoiding the shallows of pancreatic islet cell transcriptomes. *Diabetes*, **68**, 1380–1393.
19. Wolf,F.A., Angerer,P. and Theis,F.J. (2018) SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.*, **19**, 15.
20. Stuart,T., Butler,A., Hoffman,P., Hafemeister,C., Papalexi,E., Mauck,W.M. 3rd, Hao,Y., Stoeckius,M., Smibert,P. and Satija,R. (2019) Comprehensive integration of single-cell data. *Cell*, **177**, 1888.e21–1902.e21.
21. Luecken,M.D. and Theis,F.J. (2019) Current best practices in single-cell RNA-seq analysis: a tutorial. *Mol. Syst. Biol.*, **15**, e8746.
22. Haghverdi,L., Lun,A.T.L., Morgan,M.D. and Marioni,J.C. (2018) Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat. Biotechnol.*, **36**, 421–427.
23. Cnop,M., Welsh,N., Jonas,J.-C., Jörns,A., Lenzen,S. and Eizirik,D.L. (2005) Mechanisms of pancreatic beta-cell death in type 1 and type 2 diabetes: many differences, few similarities. *Diabetes*, **54**, S97–S107.
24. Taneera,J., Lang,S., Sharma,A., Fadista,J., Zhou,Y., Ahlqvist,E., Jonsson,A., Lyssenko,V., Vikman,P., Hansson,O. *et al.* (2012) A systems genetics approach identifies genes and pathways for type 2 diabetes in human islets. *Cell Metab.*, **16**, 122–134.
25. Solimena,M., Schulte,A.M., Marselli,L., Ehehalt,F., Richter,D., Kleeberg,M., Mziaut,H., Knoch,K.-P., Parnis,J., Bugliani,M. *et al.* (2018) Systems biology of the IMIDIA biobank from organ donors and pancreatectomised patients defines a novel transcriptomic signature of islets from individuals with type 2 diabetes. *Diabetologia*, **61**, 641–657.
26. Marselli,L., Sgroi,D.C., Bonner-Weir,S. and Weir,G.C. (2009) Laser capture microdissection of human pancreatic beta-cells and RNA preparation for gene expression profiling. *Methods Mol. Biol.*, **560**, 87–98.
27. Ebrahimi,A.G., Hollister-Lock,J., Sullivan,B.A., Tsuchida,R., Bonner-Weir,S. and Weir,G.C. (2020) Beta cell identity changes with mild hyperglycemia: implications for function, growth, and vulnerability. *Mol. Metab.*, **35**, 100959.

28. Leinonen,R., Sugawara,H., Shumway,M. and International Nucleotide Sequence Database Collaboration (2011) The Sequence Read Archive. *Nucleic Acids Res.*, **39**, D19–D21.
29. Parkinson,H., Kapushesky,M., Shojatalab,M., Abeygunawardena,N., Coulson,R., Farne,A., Holloway,E., Kolesnykov,N., Lilja,P., Lukk,M. *et al.* (2007) ArrayExpress—a public database of microarray experiments and gene expression profiles. *Nucleic Acids Res.*, **35**, D747–D750.
30. Dobin,A., Davis,C.A., Schlesinger,F., Drenkow,J., Zaleski,C., Jha,S., Batut,P., Chaisson,M. and Gingeras,T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.
31. Lu,H., Halappanavar,M. and Kalyanaraman,A. (2015) Parallel heuristics for scalable community detection. *Parallel Comput.*, **47**, 19–37.
32. Franzén,O., Gan,L.-M. and Björkegren,J.L.M. (2019) PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data. *Database*, **2019**, baz046.
33. Love,M.I., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
34. Ashburner,M., Ball,C.A., Blake,J.A., Botstein,D., Butler,H., Cherry,J.M., Davis,A.P., Dolinski,K., Dwight,S.S., Eppig,J.T. *et al.* (2000) Gene Ontology: tool for the unification of biology. *Nat. Genet.*, **25**, 25–29.
35. The Gene Ontology Consortium (2019) The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res.*, **47**, D330–D338.
36. Huang,R., Grishagin,I., Wang,Y., Zhao,T., Greene,J., Obenauer,J.C., Ngan,D., Nguyen,D.-T., Guha,R., Jadhav,A. *et al.* (2019) The NCATS BioPlanet—an integrated platform for exploring the universe of cellular signaling pathways for toxicology, systems biology, and chemical genomics. *Front. Pharmacol.*, **10**, 445.
37. Kanehisa,M. and Goto,S. (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.*, **28**, 27–30.
38. Kanehisa,M., Sato,Y., Furumichi,M., Morishima,K. and Tanabe,M. (2019) New approach for understanding genome variations in KEGG. *Nucleic Acids Res.*, **47**, D590–D595.
39. Kanehisa,M. (2019) Toward understanding the origin and evolution of cellular organisms. *Protein Sci.*, **28**, 1947–1951.
40. Jassal,B., Matthews,L., Viteri,G., Gong,C., Lorente,P., Fabregat,A., Sidiropoulos,K., Cook,J., Gillespie,M., Haw,R. *et al.* (2020) The Reactome pathway knowledgebase. *Nucleic Acids Res.*, **48**, D498–D503.
41. Plaisier,S.B., Taschereau,R., Wong,J.A. and Graeber,T.G. (2010) Rank–rank hypergeometric overlap: identification of statistically significant overlap between gene-expression signatures. *Nucleic Acids Res.*, **38**, e169.
42. Cahill,K.M., Huo,Z., Tseng,G.C., Logan,R.W. and Seney,M.L. (2018) Improved identification of concordant and discordant gene expression signatures using an updated rank–rank hypergeometric overlap approach. *Sci. Rep.*, **8**, 9588.
43. Virtanen,P., Gommers,R., Oliphant,T.E., Haberland,M., Reddy,T., Cournapeau,D., Burovski,E., Peterson,P., Weckesser,W., Bright,J. *et al.* (2020) SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods*, **17**, 261–272.
44. Hunter,J.D. (2007) Matplotlib: a 2D graphics environment. *Comput. Sci. Eng.*, **9**, 90–95.
45. Cnop,M., Abdulkarim,B., Bottu,G., Cunha,D.A., Igoillo-Esteve,M., Masini,M., Turatsinze,J.-V., Griebel,T., Villate,O., Santin,I. *et al.* (2014) RNA sequencing identifies dysregulation of the human pancreatic islet transcriptome by the saturated fatty acid palmitate. *Diabetes*, **63**, 1978–1993.
46. Poitout,V. and Robertson,R.P. (2008) Glucolipotoxicity: fuel excess and beta-cell dysfunction. *Endocr. Rev.*, **29**, 351–366.
47. Weir,G.C., Marselli,L., Marchetti,P., Katsuta,H., Jung,M.H. and Bonner-Weir,S. (2009) Towards better understanding of the contributions of overwork and glucotoxicity to the β-cell inadequacy of type 2 diabetes. *Diabetes Obes. Metab.*, **11**, 82–90.
48. Kim,M.-J., Jo,D.-G., Hong,G.-S., Kim,B.J., Lai,M., Cho,D.-H., Kim,K.-W., Bandyopadhyay,A., Hong,Y.-M., Kim,D.H. *et al.* (2002) Calpain-dependent cleavage of cain/cabin1 activates calcineurin to mediate calcium-triggered cell death. *Proc. Natl Acad. Sci. U.S.A.*, **99**, 9870–9875.
49. Zhong,L., Georgia,S., Tschen,S.-I., Nakayama,K., Nakayama,K. and Bhushan,A. (2007) Essential role of Skp2-mediated p27 degradation in growth and adaptive expansion of pancreatic beta cells. *J. Clin. Invest.*, **117**, 2869–2876.
50. Malek,E., Abdel-Malek,M.A.Y., Jagannathan,S., Vad,N., Karns,R., Jegga,A.G., Broyl,A., van Duin,M., Sonneveld,P., Cottini,F. *et al.* (2017) Pharmacogenomics and chemical library screens reveal a novel SCFSKP2 inhibitor that overcomes bortezomib resistance in multiple myeloma. *Leukemia*, **31**, 645–653.
51. Jinesh,G.G. and Kamat,A.M. (2017) RalBP1 and p19-VHL play an oncogenic role, and p30-VHL plays a tumor suppressor role during the blebbishield emergency program. *Cell Death Discov.*, **3**, 17023.
52. Luo,J., Emanuele,M.J., Li,D., Creighton,C.J., Schlabach,M.R., Westbrook,T.F., Wong,K.-K. and Elledge,S.J. (2009) A genome-wide RNAi screen identifies multiple synthetic lethal interactions with the Ras oncogene. *Cell*, **137**, 835–848.
53. Kessler,J.D., Kahle,K.T., Sun,T., Meerbrey,K.L., Schlabach,M.R., Schmitt,E.M., Skinner,S.O., Xu,Q., Li,M.Z., Hartman,Z.C. *et al.* (2012) A SUMOylation-dependent transcriptional subprogram is required for Myc-driven tumorigenesis. *Science*, **335**, 348–353.
54. Tanaka,A., Kaneto,H., Miyatsuka,T., Yamamoto,K., Yoshiuchi,K., Yamasaki,Y., Shimomura,I., Matsuoka,T.-A. and Matsuhisa,M. (2009) Role of copper ion in the pathogenesis of type 2 diabetes. *Endocr. J.*, **56**, 699–706.
55. Jakobsen,C.H., Størvold,G.L., Bremseth,H., Follestad,T., Sand,K., Mack,M., Olsen,K.S., Lundemo,A.G., Iversen,J.G., Krokan,H.E. *et al.* (2008) DHA induces ER stress and growth arrest in human colon cancer cells: associations with cholesterol and calcium homeostasis. *J. Lipid Res.*, **49**, 2089–2100.
56. Huda,N., Hosen,M.I., Yasmin,T., Sarkar,P.K., Hasan,A.K.M.M. and Nabi,A.H.M.N. (2018) Genetic variation of the transcription factor GATA3, not STAT4, is associated with the risk of type 2 diabetes in the Bangladeshi population. *PLoS One*, **13**, e0198507.
57. Yuan,L., Lv,Y., Li,H., Gao,H., Song,S., Zhang,Y., Xing,G., Kong,X., Wang,L., Li,Y. *et al.* (2015) Deubiquitylase OTUD3 regulates PTEN stability and suppresses tumorigenesis. *Nat. Cell Biol.*, **17**, 1169–1181.
58. Prentki,M., Joly,E., El-Assaad,W. and Roduit,R. (2002) Malonyl-CoA signaling, lipid partitioning, and glucolipotoxicity: role in beta-cell adaptation and failure in the etiology of diabetes. *Diabetes*, **51**, S405–S413.
59. Schröder,B.A., Wrocklage,C., Hasilik,A. and Saftig,P. (2010) The proteome of lysosomes. *Proteomics*, **10**, 4053–4076.
60. Marselli,L., Thorne,J., Dahiya,S., Sgroi,D.C., Sharma,A., Bonner-Weir,S., Marchetti,P. and Weir,G.C. (2010) Gene expression profiles of beta-cell enriched tissue obtained by laser capture microdissection from subjects with type 2 diabetes. *PLoS One*, **5**, e11499.
61. Xu,J., Han,J., Epstein,P.N. and Liu,Y.Q. (2006) Regulation of PDK mRNA by high fatty acid and glucose in pancreatic islets. *Biochem. Biophys. Res. Commun.*, **344**, 827–833.
62. Huang,L.-Y., Wang,Y.-P., Wei,B.-F., Yang,J., Wang,J.-Q., Wu,B.-H., Zhang,Z.-Z., Hou,Y.-Y., Sun,W.-M., Hu,R.-M. *et al.* (2013) Deficiency of IRTKS as an adaptor of insulin receptor leads to insulin resistance. *Cell Res.*, **23**, 1310–1321.
63. Kelkar,N., Standen,C.L. and Davis,R.J. (2005) Role of the JIP4 scaffold protein in the regulation of mitogen-activated protein kinase signaling pathways. *Mol. Cell. Biol.*, **25**, 2733–2743.
64. Ben,J., Jiang,B., Wang,D., Liu,Q., Zhang,Y., Qi,Y., Tong,X., Chen,L., Liu,X., Zhang,Y. *et al.* (2019) Major vault protein suppresses obesity and atherosclerosis through inhibiting IKK–NF-κB signaling mediated inflammation. *Nat. Commun.*, **10**, 1801.
65. Song,C., Yan,H., Wang,H., Zhang,Y., Cao,H., Wan,Y., Kong,L., Chen,S., Xu,H., Pan,B. *et al.* (2018) AQR is a novel type 2 diabetes-associated gene that regulates signaling pathways critical for glucose metabolism. *J. Genet. Genomics*, **45**, 111–120.
66. Possik,E., Jalali,Z., Nouët,Y., Yan,M., Gingras,M.-C., Schmeisser,K., Panaite,L., Dupuy,F., Kharitidi,D., Chotard,L. *et al.* (2014) Folliculin regulates ampk-dependent autophagy and metabolic stress survival. *PLoS Genet.*, **10**, e1004273.
67. Engelbrechtsen,L., Mahendran,Y., Jonsson,A., Gjesing,A.P., Weeke,P.E., Jørgensen,M.E., Færch,K., Witte,D.R., Holst,J.J., Jørgensen,T. *et al.* (2018) Common variants in the hERG (KCNH2)

voltage-gated potassium channel are associated with altered fasting and glucose-stimulated plasma incretin and glucagon responses. *BMC Genet.*, **19**, 15.
68. Hsu,H.-C., Chiou,J.-F., Wang,Y.-H., Chen,C.-H., Mau,S.-Y., Ho,C.-T., Chang,P.-J., Liu,T.-Z. and Chen,C.-H. (2013) Folate deficiency triggers an oxidative-nitrosative stress-mediated apoptotic cell death and impedes insulin biosynthesis in RINm5F pancreatic islet β-cells: relevant to the pathogenesis of diabetes. *PLoS One*, **8**, e77931.
69. Nakagawa,Y., Shimano,H., Yoshikawa,T., Ide,T., Tamura,M., Furusawa,M., Yamamoto,T., Inoue,N., Matsuzaka,T., Takahashi,A. *et al.* (2006) TFE3 transcriptionally activates hepatic IRS-2, participates in insulin signaling and ameliorates diabetes. *Nat. Med.*, **12**, 107–113.
70. Fryirs Michelle,A., Barter Philip,J., Mathiyalagan,Appavoo, Tuch Bernard,E., Fatiha,Tabet, Heather Alison,K. and Kerry-Anne,Rye (2010) Effects of high-density lipoproteins on pancreatic β-cell insulin secretion. *Arterioscler. Thromb. Vasc. Biol.*, **30**, 1642–1648.
71. Knoch,K.-P., Nath-Sain,S., Petzold,A., Schneider,H., Beck,M., Wegbrod,C., Sönmez,A., Münster,C., Friedrich,A., Roivainen,M. *et al.* (2014) PTBP1 is required for glucose-stimulated cap-independent translation of insulin granule proteins and Coxsackieviruses in beta cells. *Mol. Metab.*, **3**, 518–530.
72. Annerén,C. and Welsh,M. (2001) Increased cytokine-induced cytotoxicity of pancreatic islet cells from transgenic mice expressing the Src-like tyrosine kinase GTK. *Mol. Med.*, **7**, 301–310.
73. Yalçin,A., Şarkici,G. and Kolaç,U.K. (2020) PKR inhibitors suppress endoplasmic reticulum stress and subdue glucolipotoxicity-mediated impairment of insulin secretion in pancreatic beta cells. *Turk. J. Biol.*, **44**, 93–102.
74. Kato,M., Wang,L., Putta,S., Wang,M., Yuan,H., Sun,G., Lanting,L., Todorov,I., Rossi,J.J. and Natarajan,R. (2010) Post-transcriptional up-regulation of Tsc-22 by Ybx1, a target of miR-216a, mediates TGF-β-induced collagen expression in kidney cells. *J. Biol. Chem.*, **285**, 34004–34015.
75. Nagy,N., de la Zerda,A., Kaber,G., Johnson,P.Y., Hu,K.H., Kratochvil,M.J., Yadava,K., Zhao,W., Cui,Y., Navarro,G. *et al.* (2018) Hyaluronan content governs tissue stiffness in pancreatic islet inflammation. *J. Biol. Chem.*, **293**, 567–578.
76. Shrestha,N., Liu,T., Ji,Y., Reinert,R.B., Torres,M., Li,X., Zhang,M., Tang,C.-H.A., Hu,C.-C.A., Liu,C. *et al.* (2020) Sel1L–Hrd1 ER-associated degradation maintains β cell identity via TGF-β signaling. *J. Clin. Invest.*, **130**, 3499–3510.
77. Wang,L., Li,Y., Guo,B., Zhang,J., Zhu,B., Li,H., Ding,Y., Meng,B., Zhao,H., Xiang,L. *et al.* (2020) Myeloid-derived growth factor promotes intestinal glucagon-like peptide-1 production in male mice with type 2 diabetes. *Endocrinology*, **161**, bqaa003.
78. Nishida,Y., Aida,K., Kihara,M. and Kobayashi,T. (2014) Antibody-validated proteins in inflamed islets of fulminant type 1 diabetes profiled by laser-capture microdissection followed by mass spectrometry. *PLoS One*, **9**, e107664.
79. Srivastava,M., Eidelman,O., Leighton,X., Glasman,M., Goping,G. and Pollard,H.B. (2002) Anx7 is required for nutritional control of gene expression in mouse pancreatic islets of Langerhans. *Mol. Med.*, **8**, 781–797.
80. Malenczyk,K., Szodorai,E., Schnell,R., Lubec,G., Szabó,G., Hökfelt,T. and Harkany,T. (2018) Secretagogin protects Pdx1 from proteasomal degradation to control a transcriptional program required for β cell specification. *Mol. Metab.*, **14**, 108–120.
81. Kim,S.Y., Lee,J.-H., Merrins,M.J., Gavrilova,O., Bisteau,X., Kaldis,P., Satin,L.S. and Rane,S.G. (2017) Loss of cyclin-dependent kinase 2 in the pancreas links primary β-cell dysfunction to progressive depletion of β-cell mass and diabetes. *J. Biol. Chem.*, **292**, 3841–3853.
82. Nagy,L., Márton,J., Vida,A., Kis,G., Bokor,É., Kun,S., Gönczi,M., Docsa,T., Tóth,A., Antal,M. *et al.* (2018) Glycogen phosphorylase inhibition improves beta cell function. *Br. J. Pharmacol.*, **175**, 301–319.
83. Sharma,M., Naslavsky,N. and Caplan,S. (2008) A role for EHD4 in the regulation of early endosomal transport. *Traffic*, **9**, 995–1018.
84. Llagostera,E., Catalucci,D., Marti,L., Liesa,M., Camps,M., Ciaraldi,T.P., Kondo,R., Reddy,S., Dillmann,W.H., Palacin,M. *et al.* (2007) Role of myotonic dystrophy protein kinase (DMPK) in glucose homeostasis and muscle insulin action. *PLoS One*, **2**, e1134.
85. Xiao,Y., Wang,C., Chen,J.-Y., Lu,F., Wang,J., Hou,N., Hu,X., Zeng,F., Ma,D., Sun,X. *et al.* (2018) Deficiency of PRKD2 triggers hyperinsulinemia and metabolic disorders. *Nat. Commun.*, **9**, 2015.
86. Higashitsuji,H., Higashitsuji,H., Masuda,T., Liu,Y., Itoh,K. and Fujita,J. (2007) Enhanced deacetylation of p53 by the anti-apoptotic protein HSCO in association with histone deacetylase 1. *J. Biol. Chem.*, **282**, 13716–13725.
87. Fiaschi-Taesch,N.M., Salim,F., Kleinberger,J., Troxell,R., Cozar-Castellano,I., Selk,K., Cherok,E., Takane,K.K., Scott,D.K. and Stewart,A.F. (2010) Induction of human beta-cell proliferation and engraftment using a single G1/S regulatory molecule, cdk6. *Diabetes*, **59**, 1926–1936.
88. Manning,A.K., Hivert,M.-F., Scott,R.A., Grimsby,J.L., Bouatia-Naji,N., Chen,H., Rybin,D., Liu,C.-T., Bielak,L.F., Prokopenko,I. *et al.* (2012) A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nat. Genet.*, **44**, 659–669.
89. Crose,L.E.S., Galindo,K.A., Kephart,J.G., Chen,C., Fitamant,J., Bardeesy,N., Bentley,R.C., Galindo,R.L., Chi,J.-T.A. and Linardic,C.M. (2014) Alveolar rhabdomyosarcoma-associated PAX3–FOXO1 promotes tumorigenesis via Hippo pathway suppression. *J. Clin. Invest.*, **124**, 285–296.
90. Guan,M., Keaton,J.M., Dimitrov,L., Hicks,P.J., Xu,J., Palmer,N.D., Wilson,J.G., Freedman,B.I., Bowden,D.W. and Ng,M.C.Y. (2018) An exome-wide association study for type 2 diabetes-attributed end-stage kidney disease in African Americans. *Kidney Int. Rep.*, **3**, 867–878.
91. Klinger,S., Poussin,C., Debril,M.-B., Dolci,W., Halban,P.A. and Thorens,B. (2008) Increasing GLP-1-induced beta-cell proliferation by silencing the negative regulators of signaling cAMP response element modulator-alpha and DUSP14. *Diabetes*, **57**, 584–593.
92. Ainscow,E.K., Zhao,C. and Rutter,G.A. (2000) Acute overexpression of lactate dehydrogenase-A perturbs beta-cell mitochondrial metabolism and insulin secretion. *Diabetes*, **49**, 1149–1155.
93. Puri,S., Roy,N., Russ,H.A., Leonhardt,L., French,E.K., Roy,R., Bengtsson,H., Scott,D.K., Stewart,A.F. and Hebrok,M. (2018) Replication confers β cell immaturity. *Nat. Commun.*, **9**, 485.
94. Rimbert,A., Dalila,N., Wolters,J.C., Huijkman,N., Smit,M., Kloosterhuis,N., Riemsma,M., van der Veen,Y., Singla,A., van Dijk,F. *et al.* (2020) A common variant in CCDC93 protects against myocardial infarction and cardiovascular mortality by regulating endosomal trafficking of low-density lipoprotein receptor. *Eur. Heart J.*, **41**, 1040–1053.
95. Han,J., Back,S.H., Hur,J., Lin,Y.-H., Gildersleeve,R., Shan,J., Yuan,C.L., Krokowski,D., Wang,S., Hatzoglou,M. *et al.* (2013) ER-stress-induced transcriptional regulation increases protein synthesis leading to cell death. *Nat. Cell Biol.*, **15**, 481–490.
96. Sasaki,T., Lian,S., Khan,A., Llop,J.R., Samuelson,A.V., Chen,W., Klionsky,D.J. and Kishi,S. (2017) Autolysosome biogenesis and developmental senescence are regulated by both Spns1 and v-ATPase. *Autophagy*, **13**, 386–403.
97. Spaeth,J.M., Hunter,C.S., Bonatakis,L., Guo,M., French,C.A., Slack,I., Hara,M., Fisher,S.E., Ferrer,J., Morrisey,E.E. *et al.* (2015) The FOXP1, FOXP2 and FOXP4 transcription factors are required for islet alpha cell proliferation and function in mice. *Diabetologia*, **58**, 1836–1844.
98. Negi,S., Jetha,A., Aikin,R., Hasilo,C., Sladek,R. and Paraskevas,S. (2012) Analysis of beta-cell gene expression reveals inflammatory signaling and evidence of dedifferentiation following human islet isolation and culture. *PLoS One*, **7**, e30415.
99. Cho,S.-J., Rossi,A., Jung,Y.-S., Yan,W., Liu,G., Zhang,J., Zhang,M. and Chen,X. (2013) Ninjurin1, a target of p53, regulates p53 expression and p53-dependent cell survival, senescence, and radiation-induced mortality. *Proc. Natl Acad. Sci. U.S.A.*, **110**, 9362–9367.
100. Yamamoto,K., Sato,T., Matsui,T., Sato,M., Okada,T., Yoshida,H., Harada,A. and Mori,K. (2007) Transcriptional induction of mammalian ER quality control proteins is mediated by single or combined action of ATF6alpha and XBP1. *Dev. Cell*, **13**, 365–376.
101. Baron,M., Maillet,J., Huyvaert,M., Dechaume,A., Boutry,R., Loiselle,H., Durand,E., Toussaint,B., Vaillant,E., Philippe,J. *et al.* (2019) Loss-of-function mutations in MRAP2 are pathogenic in

hyperphagic obesity with hyperglycemia and hypertension. *Nat. Med.*, **25**, 1733–1738.
102. Wang,W., Huang,X., Xin,H.-B., Fu,M., Xue,A. and Wu,Z.-H. (2015) TRAF family member-associated NF-κB activator (TANK) inhibits genotoxic nuclear factor κB activation by facilitating deubiquitinase USP10-dependent deubiquitination of TRAF6 ligase. *J. Biol. Chem.*, **290**, 13372–13385.
103. Boothe,T., Lim,G.E., Cen,H., Skovsø,S., Piske,M., Li,S.N., Nabi,I.R., Gilon,P. and Johnson,J.D. (2016) Inter-domain tagging implicates caveolin-1 in insulin receptor trafficking and Erk signaling bias in pancreatic beta-cells. *Mol. Metab.*, **5**, 366–378.
104. Kee,A.J., Chagan,J., Chan,J.Y., Bryce,N.S., Lucas,C.A., Zeng,J., Hook,J., Treutlein,H., Laybutt,D.R., Stehn,J.R. *et al.* (2018) On-target action of anti-tropomyosin drugs regulates glucose metabolism. *Sci. Rep.*, **8**, 4604.
105. Grieß,K., Polanski,C., Markgraf,D., Lammert,E., Roden,M., Stark,H., Brüning,J. and Belgardt,B.F. (2018) The role of ceramide synthases in pancreatic beta cell demise. In: *Diabetes Kongress 2018. 53. Jahrestagung der DDG*. Georg Thieme Verlag KG, Stuttgart, Vol. **13**, p. FV 18.
106. Kasai,K., Fujita,T., Gomi,H. and Izumi,T. (2008) Docking is not a prerequisite but a temporal constraint for fusion of secretory granules. *Traffic*, **9**, 1191–1203.
107. Yoshikumi,Y., Mashima,H., Ueda,N., Ohno,H., Suzuki,J., Tanaka,S., Hayashi,M., Sekine,N., Ohnishi,H., Yasuda,H. *et al.* (2005) Roles of CTPL/Sfxn3 and Sfxn family members in pancreatic islet. *J. Cell. Biochem.*, **95**, 1157–1168.
108. Mitok,K.A., Freiberger,E.C., Schueler,K.L., Rabaglia,M.E., Stapleton,D.S., Kwiecien,N.W., Malec,P.A., Hebert,A.S., Broman,A.T., Kennedy,R.T. *et al.* (2018) Islet proteomics reveals genetic variation in dopamine production resulting in altered insulin secretion. *J. Biol. Chem.*, **293**, 5860–5877.
109. Schueler,M., Braun,D.A., Chandrasekar,G., Gee,H.Y., Klasson,T.D., Halbritter,J., Bieder,A., Porath,J.D., Airik,R., Zhou,W. *et al.* (2015) DCDC2 mutations cause a renal–hepatic ciliopathy by disrupting Wnt signaling. *Am. J. Hum. Genet.*, **96**, 81–92.
110. Klein,D., Misawa,R., Bravo-Egana,V., Vargas,N., Rosero,S., Piroso,J., Ichii,H., Umland,O., Zhijie,J., Tsinoremas,N. *et al.* (2013) MicroRNA expression in alpha and beta cells of human pancreatic islets. *PLoS One*, **8**, e55064.
111. Diaferia,G.R., Jimenez-Caliani,A.J., Ranjitkar,P., Yang,W., Hardiman,G., Rhodes,C.J., Crisa,L. and Cirulli,V. (2013) β1 integrin is a crucial regulator of pancreatic β-cell expansion. *Development*, **140**, 3360–3372.
112. Santo-Domingo,J., Galindo,A.N., Cominetti,O., De Marchi,U., Cutillas,P., Dayon,L. and Wiederkehr,A. (2019) Glucose-dependent phosphorylation signaling pathways and crosstalk to mitochondrial respiration in insulin secreting cells. *Cell Commun. Signal.*, **17**, 14.
113. Slayton,M., Gupta,A., Balakrishnan,B. and Puri,V. (2019) CIDE proteins in human health and disease. *Cells*, **8**, 238.
114. Rorsman,P. and Ashcroft,F.M. (2018) Pancreatic β-cell electrical activity and insulin secretion: of mice and men. *Physiol. Rev.*, **98**, 117–214.
115. Fujino,T., Asaba,H., Kang,M.-J., Ikeda,Y., Sone,H., Takada,S., Kim,D.-H., Ioka,R.X., Ono,M., Tomoyori,H. *et al.* (2003) Low-density lipoprotein receptor-related protein 5 (LRP5) is essential for normal cholesterol metabolism and glucose-induced insulin secretion. *Proc. Natl Acad. Sci. U.S.A.*, **100**, 229–234.
116. Miele,C., Raciti,G.A., Cassese,A., Romano,C., Giacco,F., Oriente,F., Paturzo,F., Andreozzi,F., Zabatta,A., Troncone,G. *et al.* (2007) PED/PEA-15 regulates glucose-induced insulin secretion by restraining potassium channel expression in pancreatic beta-cells. *Diabetes*, **56**, 622–633.
117. Fadista,J., Vikman,P., Laakso,E.O., Mollet,I.G., Esguerra,J.L., Taneera,J., Storm,P., Osmark,P., Ladenvall,C., Prasad,R.B. *et al.* (2014) Global genomic and transcriptomic analysis of human pancreatic islets reveals novel genes influencing glucose metabolism. *Proc. Natl Acad. Sci. U.S.A.*, **111**, 13924–13929.
118. Lim,G.E., Piske,M., Lulo,J.E., Ramshaw,H.S., Lopez,A.F. and Johnson,J.D. (2016) Ywhaz/14-3-3ζ deletion improves glucose tolerance through a GLP-1-dependent mechanism. *Endocrinology*, **157**, 2649–2659.
119. Wong,V.S.C., Oh,A.H., Chassot,A.A., Chaboissier,M.C. and Brubaker,P.L. (2011) R-spondin1 deficiency in mice improves glycaemic control in association with increased beta cell mass. *Diabetologia*, **54**, 1726–1734.
120. Marchetti,P., Bugliani,M., De Tata,V., Suleiman,M. and Marselli,L. (2017) Pancreatic beta cell identity in humans and the role of type 2 diabetes. *Front. Cell Dev. Biol.*, **5**, 55.
121. Liu,P., Shi,L., Cang,X., Huang,J., Wu,X., Yan,J., Chen,L., Cui,S. and Ye,X. (2017) CtBP2 ameliorates palmitate-induced insulin resistance in HepG2 cells through ROS mediated JNK pathway. *Gen. Comp. Endocrinol.*, **247**, 66–73.
122. Huang,J.-Q., Zhou,J.-C., Wu,Y.-Y., Ren,F.-Z. and Lei,X.G. (2018) Role of glutathione peroxidase 1 in glucose and lipid metabolism-related diseases. *Free Radic. Biol. Med.*, **127**, 108–115.
123. Metz,S.A. and Kowluru,A. (1999) Inosine monophosphate dehydrogenase: a molecular switch integrating pleiotropic GTP-dependent beta-cell functions. *Proc. Assoc. Am. Physicians*, **111**, 335–346.
124. Lin,Y.-M., Chen,Y.-R., Lin,J.-R., Wang,W.-J., Inoko,A., Inagaki,M., Wu,Y.-C. and Chen,R.-H. (2008) eIF3k regulates apoptosis in epithelial cells by releasing caspase 3 from keratin-containing inclusions. *J. Cell Sci.*, **121**, 2382–2393.
125. Tapia-Limonchi,R., Díaz,I., Cahuana,G.M., Bautista,M., Martín,F., Soria,B., Tejedo,J.R. and Bedoya,F.J. (2014) Impact of exposure to low concentrations of nitric oxide on protein profile in murine and human pancreatic islet cells. *Islets*, **6**, e995997.
126. Zhu,X., Hu,R., Brissova,M., Stein,R.W., Powers,A.C., Gu,G. and Kaverina,I. (2015) Microtubules negatively regulate insulin secretion in pancreatic β cells. *Dev. Cell*, **34**, 656–668.
127. Zhang,M., Robitaille,M., Showalter,A.D., Huang,X., Liu,Y., Bhattacharjee,A., Willard,F.S., Han,J., Froese,S., Wei,L. *et al.* (2014) Progesterone receptor membrane component 1 is a functional part of the glucagon-like peptide-1 (GLP-1) receptor complex in pancreatic β cells. *Mol. Cell. Proteomics*, **13**, 3049–3062.
128. Shi,N., Guo,X. and Chen,S.-Y. (2014) Olfactomedin 2, a novel regulator for transforming growth factor-β-induced smooth muscle differentiation of human embryonic stem cell-derived mesenchymal cells. *Mol. Biol. Cell*, **25**, 4106–4114.
129. Ferdaoussi,M. and MacDonald,P.E. (2017) Toward connecting metabolism to the exocytotic site. *Trends Cell Biol.*, **27**, 163–171.
130. Finkbeiner,E., Haindl,M. and Muller,S. (2011) The SUMO system controls nucleolar partitioning of a novel mammalian ribosome biogenesis complex. *EMBO J.*, **30**, 1067–1078.
131. Tong,X., Kono,T., Anderson-Baucum,E.K., Yamamoto,W., Gilon,P., Lebeche,D., Day,R.N., Shull,G.E. and Evans-Molina,C. (2016) SERCA2 deficiency impairs pancreatic β-cell function in response to diet-induced obesity. *Diabetes*, **65**, 3039–3052.
132. Yamada,E., Okada,S., Bastie,C.C., Vatish,M., Nakajima,Y., Shibusawa,R., Ozawa,A., Pessin,J.E. and Yamada,M. (2016) Fyn phosphorylates AMPK to inhibit AMPK activity and AMP-dependent activation of autophagy. *Oncotarget*, **7**, 74612–74629.
133. Kamath,V., Kyathanahalli,C.N., Jayaram,B., Syed,I., Olson,L.K., Ludwig,K., Klumpp,S., Krieglstein,J. and Kowluru,A. (2010) Regulation of glucose- and mitochondrial fuel-induced insulin secretion by a cytosolic protein histidine phosphatase in pancreatic beta-cells. *Am. J. Physiol. Endocrinol. Metab.*, **299**, E276–E286.
134. Arous,C. and Halban,P.A. (2015) The skeleton in the closet: actin cytoskeletal remodeling in β-cell function. *Am. J. Physiol. Endocrinol. Metab.*, **309**, E611–E620.
135. Kurian,S.M., Ferreri,K., Wang,C.-H., Todorov,I., Al-Abdullah,I.H., Rawson,J., Mullen,Y., Salomon,D.R. and Kandeel,F. (2017) Gene expression signature predicts human islet integrity and transplant functionality in diabetic mice. *PLoS One*, **12**, e0185331.
136. Scavuzzo,M.A., Hill,M.C., Chmielowiec,J., Yang,D., Teaw,J., Sheng,K., Kong,Y., Bettini,M., Zong,C., Martin,J.F. *et al.* (2018) Endocrine lineage biases arise in temporally distinct endocrine progenitors during pancreatic morphogenesis. *Nat. Commun.*, **9**, 3356.
137. Kao,D.-I., Lacko,L.A., Ding,B.-S., Huang,C., Phung,K., Gu,G., Rafii,S., Stuhlmann,H. and Chen,S. (2015) Endothelial cells control pancreatic cell fate at defined stages through EGFL7 signaling. *Stem Cell Rep.*, **4**, 181–189.

138. Stancill,J.S., Cartailler,J.-P., Clayton,H.W., O'Connor,J.T., Dickerson,M.T., Dadi,P.K., Osipovich,A.B., Jacobson,D.A. and Magnuson,M.A. (2017) Chronic β-cell depolarization impairs β-cell identity by disrupting a network of Ca^{2+}-regulated genes. *Diabetes*, **66**, 2175–2187.
139. Butler,A., Hoffman,P., Smibert,P., Papalexi,E. and Satija,R. (2018) Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.*, **36**, 411–420.
140. Juan-Mateu,J., Alvelos,M.I., Turatsinze,J.-V., Villate,O., Lizarraga-Mollinedo,E., Grieco,F.A., Marroquí,L., Bugliani,M., Marchetti,P. and Eizirik,D.L. (2018) SRp55 regulates a splicing network that controls human pancreatic β-cell function and survival. *Diabetes*, **67**, 423–436.
141. Lan,H., Rabaglia,M.E., Stoehr,J.P., Nadler,S.T., Schueler,K.L., Zou,F., Yandell,B.S. and Attie,A.D. (2003) Gene expression profiles of nondiabetic and diabetic obese mice suggest a role of hepatic lipogenic capacity in diabetes susceptibility. *Diabetes*, **52**, 688–700.
142. Shankar,A. and Teppala,S. (2011) Relationship between urinary bisphenol A levels and diabetes mellitus. *J. Clin. Endocrinol. Metab.*, **96**, 3822–3826.
143. Daneshpajooh,M., Bacos,K., Bysani,M., Bagge,A., Ottosson Laakso,E., Vikman,P., Eliasson,L., Mulder,H. and Ling,C. (2017) HDAC7 is overexpressed in human diabetic islets and impairs insulin secretion in rat islets and clonal beta cells. *Diabetologia*, **60**, 116–125.
144. Wu,H., Rahman,H.N.A., Dong,Y., Liu,X., Lee,Y., Wen,A., To,K.H., Xiao,L., Birsner,A.E., Bazinet,L. *et al.* (2018) Epsin deficiency promotes lymphangiogenesis through regulation of VEGFR3 degradation in diabetes. *J. Clin. Invest.*, **128**, 4025–4043.
145. Tewari,R., Bachert,C. and Linstedt,A.D. (2015) Induced oligomerization targets Golgi proteins for degradation in lysosomes. *Mol. Biol. Cell*, **26**, 4427–4437.
146. Flannick,J., Mercader,J.M., Fuchsberger,C., Udler,M.S., Mahajan,A., Wessel,J., Teslovich,T.M., Caulkins,L., Koesterer,R., Barajas-Olmos,F. *et al.* (2019) Exome sequencing of 20,791 cases of type 2 diabetes and 24,440 controls. *Nature*, **570**, 71–76.
147. El Hajj,N., Pliushch,G., Schneider,E., Dittrich,M., Müller,T., Korenkov,M., Aretz,M., Zechner,U., Lehnen,H. and Haaf,T. (2013) Metabolic programming of MEST DNA methylation by intrauterine exposure to gestational diabetes mellitus. *Diabetes*, **62**, 1320–1328.
148. Takahashi,M., Kamei,Y. and Ezaki,O. (2005) Mest/Peg1 imprinted gene enlarges adipocytes and is a marker of adipocyte size. *Am. J. Physiol. Endocrinol. Metab.*, **288**, E117–E124.
149. Elbein,S.C., Gamazon,E.R., Das,S.K., Rasouli,N., Kern,P.A. and Cox,N.J. (2012) Genetic risk factors for type 2 diabetes: a trans-regulatory genetic architecture? *Am. J. Hum. Genet.*, **91**, 466–477.
150. Chen,K., Jin,P., He,H.-H., Xie,Y.-H., Xie,X.-Y. and Mo,Z.-H. (2011) Overexpression of Insig-1 protects β cell against glucolipotoxicity via SREBP-1c. *J. Biomed. Sci.*, **18**, 57.
151. Shin,M., Kang,H.S., Park,J.H., Bae,J.H., Song,D.K. and Im,S.S. (2017) Recent insights into insulin-like growth factor binding protein 2 transcriptional regulation. *Endocrinol. Metab. (Seoul)*, **32**, 11–17.
152. Tabassum,R., Jaiswal,A., Chauhan,G., Dwivedi,O.P., Ghosh,S., Marwaha,R.K., Tandon,N. and Bharadwaj,D. (2012) Genetic variant of AMD1 is associated with obesity in urban Indian children. *PLoS One*, **7**, e33162.
153. Hermjakob,H., Montecchi-Palazzi,L., Lewington,C., Mudali,S., Kerrien,S., Orchard,S., Vingron,M., Roechert,B., Roepstorff,P., Valencia,A. *et al.* (2004) IntAct: an open source molecular interaction database. *Nucleic Acids Res.*, **32**, D452–D455.
154. Hwang,E.S., Choi,A. and Ho,I.-C. (2002) Transcriptional regulation of GATA-3 by an intronic regulatory region and fetal liver zinc finger protein 1. *J. Immunol.*, **169**, 248–253.
155. Soleimanpour,S.A., Crutchlow,M.F., Ferrari,A.M., Raum,J.C., Groff,D.N., Rankin,M.M., Liu,C., De León,D.D., Naji,A., Kushner,J.A. *et al.* (2010) Calcineurin signaling regulates human islet β-cell survival. *J. Biol. Chem.*, **285**, 40050–40059.