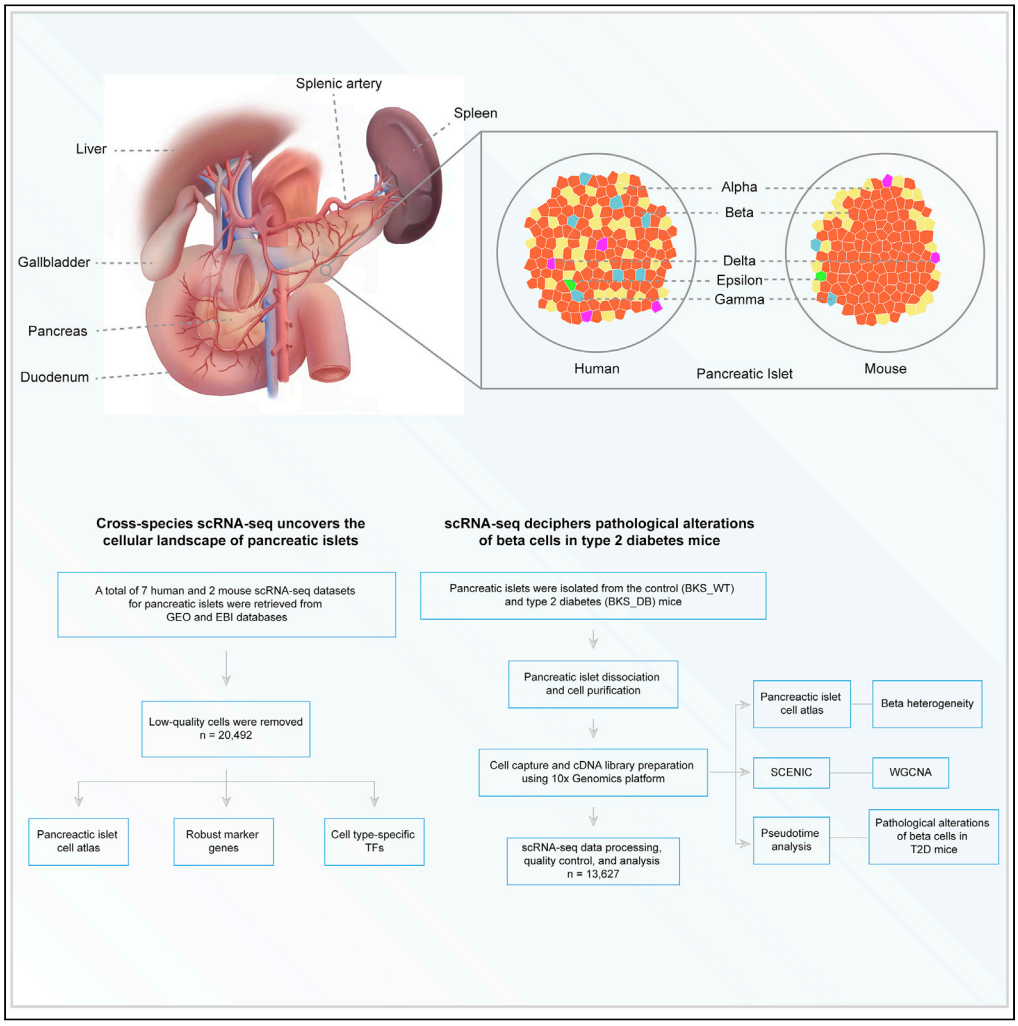
iScience

CellPress
OPEN ACCESS

Article

Single-cell RNA-seq transcriptomic landscape of human and mouse islets and pathological alterations of diabetes



Kai Chen, Junqing Zhang, Youyuan Huang, Xiaodong Tian, Yinmo Yang, Aimei Dong

tianxiaodong@pkufh.com (X.T.)
yangyinmosci@bjmu.edu.cn (Y.Y.)
dongaimei@pkufh.com (A.D.)

Highlights

Cross-species scRNA-seq reveals the complete cell landscape of the islets of Langerhans

We identify the robust marker genes and TFs of each endocrine and exocrine cell type

Pathological alterations of beta cells in type 2 diabetes are explored

A Web-based interactive tool is established for pancreatic islet scRNA-seq datasets

Chen et al., iScience 25, 105366
November 18, 2022
https://doi.org/10.1016/j.isci.2022.105366

iScience

CellPress
OPEN ACCESS

Article

Single-cell RNA-seq transcriptomic landscape of human and mouse islets and pathological alterations of diabetes

Kai Chen,[1,3] Junqing Zhang,[2] Youyuan Huang,[2] Xiaodong Tian,[1,*] Yinmo Yang,[1,*] and Aimei Dong[2,*]

SUMMARY

Single-cell RNA sequencing has paved the way for delineating the pancreatic islet cell atlas and identifying hallmarks of diabetes. However, pathological alterations of type 2 diabetes (T2D) remain unclear. We isolated pancreatic islets from control and T2D mice for single-cell RNA sequencing (scRNA-seq) and retrieved multiple datasets from the open databases. The complete islet cell landscape and robust marker genes and transcription factors of each endocrine cell type were identified. GLRA1 was restricted to beta cells, and beta cells exhibited obvious heterogeneity. The beta subcluster in the T2D mice remarkably decreased the expression of Slc2a2, G6pc2, Mafa, Nkx6-1, Pdx1, and Ucn3 and had higher unfolded protein response (UPR) scores than in the control mice. Moreover, we developed a Web-based interactive tool, creating new opportunities for the data mining of pancreatic islet scRNA-seq datasets. In conclusion, our work provides a valuable resource for a deeper understanding of the pathological mechanism underlying diabetes.

INTRODUCTION

The islets of Langerhans, which account for only 5% of the pancreatic mass, mainly consist of five endocrine cells, including alpha (glucagon), beta (insulin), gamma (pancreatic polypeptide), delta (somatostatin), and epsilon (ghrelin) cells.[1] Pancreatic islet composition and beta cell heterogeneity have been widely studied, but little is known about the less common cell types, such as gamma and epsilon cells.[2] Hormone-producing endocrine cells play major roles in glucose homeostasis, and islet dysfunction is a hallmark of type 1 diabetes (T1D) and type 2 diabetes (T2D) mellitus. Diabetes is characterized by chronic metabolic disorders that cause serious complications in the middle-late stages.[3] T2D is a heterogeneous disease caused by insulin deficiency and resistance. The pathological mechanisms underlying insulin secretion deficiency have been controversial. Whether and how beta cell apoptosis, dedifferentiation, and endoplasmic reticulum (ER) stress are involved in the progression of diabetes needs further investigation. In addition, the imaging methods for the accurate evaluation of beta cell mass and secretion function remain to be explored.

Recently, the advent of single-cell RNA sequencing (scRNA-seq) paved the way for delineating the pancreatic islet cell atlas and identifying cellular heterogeneity at unprecedented resolution.[1,4] For example, Baron et al.[5] isolated pancreatic islets from four human and two mouse donors and conducted scRNA-seq analysis using the inDrop platform. A comprehensive list of pancreatic cell types and specific transcription factors (TFs) was identified. Enge et al.[6] directly measured the effects of aging on the human pancreas using single-cell transcriptome analysis. Human pancreas was derived from eight donors of different ages. They found that endocrine cells from older donors showed remarkably increased levels of transcriptional noise and identified an age-related mutational signature. Many studies[7–11] have revealed the beta heterogeneity and pathological alterations of T2D in human donors using scRNA-seq technologies. Moreover, Qiu and colleagues investigated the key maturation pathways and signatures of pancreatic islet alpha and beta cells using the Smart-seq2 method. However, the sample size and number of captured cells were limited in previous studies at the beginning of the application of scRNA-seq. Type 2 diabetes mice (Lepr knockout model) have been widely used to explore the pathological characteristics of diabetes and diabetic complications, whereas scRNA-seq of this model has not been performed.

[1]Department of General Surgery, Peking University First Hospital, Beijing 100034, China

[2]Department of Endocrinology, Peking University First Hospital, Beijing 100034, China

[3]Lead contact

*Correspondence: tianxiaodong@pkufh.com (X.T.), yangyinmosci@bjmu.edu.cn (Y.Y.), dongaimei@pkufh.com (A.D.)

https://doi.org/10.1016/j.isci.2022.105366

As mentioned earlier, there is a tremendous amount of scRNA-seq data for pancreatic islets to be deeply analyzed. With the evolution of bioinformatics methods, scRNA-seq data will provide more valuable information for the identification of rare cell types,[12,13] diagnostic and prognostic biomarkers,[14] cell-to-cell communication,[15–17] and cellular developmental trajectories.[18,19] In previous scRNA-seq studies for pancreatic islets,[5–11,20] newly established bioinformatics methods were not applied. In this study, all available human and mouse scRNA-seq datasets were retrieved from open databases. We revealed the complete cell landscape of the islets of Langerhans and identified the common marker genes and TFs of each endocrine and exocrine cell type across all datasets. The expression and distribution of glycine receptors and transporters in the pancreas were also investigated. In addition, we isolated pancreatic islets from the control (BKS_WT) and T2D (BKS_DB) mice and conducted scRNA-seq analysis. Pathological alterations, including glycolysis, hypoxia, ER stress, and dedifferentiation, of T2D at single-cell resolution were explored. A Web-based interactive tool was established, which created new opportunities for data mining of pancreatic islet scRNA-seq datasets.

RESULTS

Cross-species scRNA-seq delineated pancreatic islet cell atlas

To uncover the cellular landscape of the islets of Langerhans, especially for rare gamma and epsilon cells, multiple scRNA-seq datasets of pancreatic islets were retrieved from open databases. A total of 20,492 cells from seven human and two mouse datasets were obtained after filtering out the low-quality cells. Pancreatic islets were isolated from human and mouse donors of different ages, developmental stages, and diabetes types (Figure S1A). Uniform manifold approximation and projection (uMAP) was used to delineate the pancreatic islet cell atlas for each dataset. Cells were segregated into distinct clusters in two dimensions, according to transcript profiles (Figure 1A). We did not detect any obvious specimen-derived bias after integrating scRNA-seq data using the Harmony method (Figure S1B). The cell types were identified by cross-referencing the signature genes of each cluster with known markers of cell populations reported in previous studies,[5–7] such as GCG (Gcg) for alpha, INS (Ins1) for beta, PPY (Ppy) for gamma, SST (Sst) for delta, KRT19 (Krt19) for ductal cells, PRSS1 (Try4) for acinar cells, LUM (Lum) and DCN (Dcn) for fibroblasts, RGS5 (Rgs5) for stellate cells, CDH5 (Cdh5) for endothelial cells, AIF1 (Aif1) for macrophages, CD79A (Cd79a) and MS4A1 (Ms4a1) for B cells, and CD3D (Cd3d) for T cells. (Figures S1C, S1D, S2A–S2D, and S3A–S3C). We identified alpha, beta, gamma, delta, ductal, and acinar cells in most datasets (Figure 1A). In addition, few fibroblasts, macrophages, stellate, and endothelial cells were identified in some datasets, indicating that exocrine and mesenchymal cells could not be completely removed during pancreatic islet isolation. No cluster specifically expressing GHRL (Ghrl), which is a maker of epsilon cell, was found (Figure 1A). Violin plots showed that alpha-specific GCG (Gcg) was also expressed in other cell subpopulations, and similar results were observed for INS (Ins1/2), PPY (Ppy), and SST (Sst). However, ductal, acinar, fibroblast, macrophage, and endothelial cell markers were expressed specifically in the corresponding subpopulation (Figures S1C, S1D, S2A–S2D, and S3A–S3C). Thus, more attention should be paid to performing droplet-based scRNA-seq for pancreatic islets because ambient RNA contamination is far more likely to occur in endocrine cells. Then, the top five cell subpopulation-specific genes were generated using the FindMakers function (Figures S3D–S3G and S4A–S4E).

Consistent with previous studies,[21,22] we found that pancreatic islet composition was not uniform across individuals and changed in response to physiological and pathological conditions, such as aging and diabetes (Figure 1B). There was a difference in pancreatic islet composition between humans and mice. Human islets consisted mostly of alpha and beta cells, and few gamma and delta cells, whereas beta cells were the predominant cell type in the mouse islets (Figures 1B and 1C). Notably, compared with juvenile (ages 1 month, 5 years, and 6 years) and young adult patients (ages 21 and 22 years), adult/middle-aged patients (ages 38, 44, and 54 years) had a lower proportion of beta and a higher proportion of alpha in Enge's dataset.[6] Very few endocrine cells, including alpha, beta, and delta cells, were captured in an adult patient with T1D in Wang's dataset,[10] indicating islet deficiency. There was no significant difference in the proportion of beta between normal (ND) and T2D groups in many datasets[5,7,9,11] (Figure 1B).

Robust marker genes of endocrine and exocrine cells in the pancreas were identified

Identification of cell-specific genes is a prerequisite for sorting cell subpopulations and exploring their biological functions. However, marker genes of gamma and delta cells are largely unknown. Here,

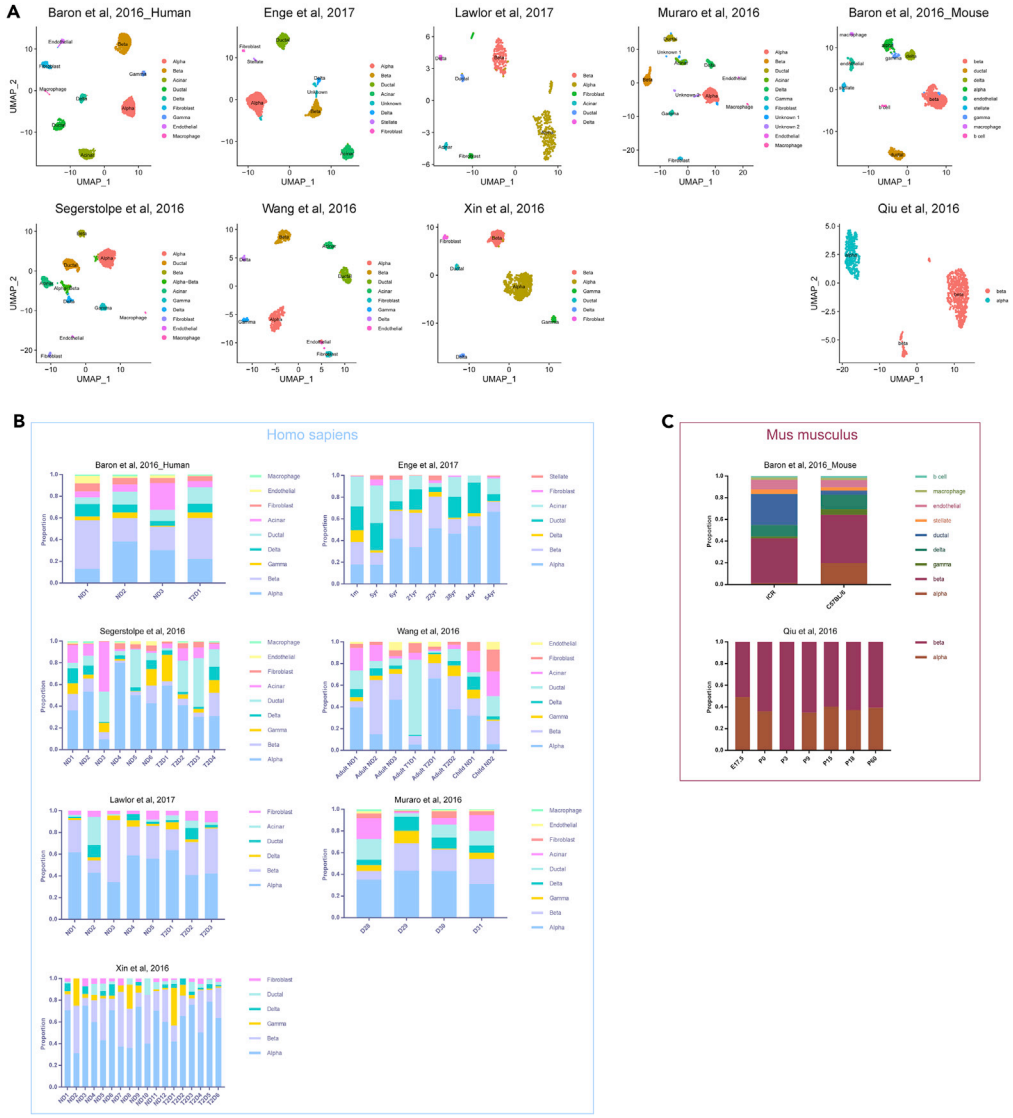


Figure 1. scRNA-seq delineates pancreatic islet cell landscape

(A) uMAP showing main cell types for each dataset. Each dot represents one cell, and cell types are coded with different colors.

(B and C) Proportion of cell populations from different specimens for human datasets (B) and mouse datasets.

common marker genes of alpha, beta, gamma, delta, ductal, and acinar cells across all datasets were obtained, which represented the most reliable markers of these cell populations (Figures 2A–2H). Consistent with previous studies,[5–11,20] well-known marker genes of endocrine and exocrine cells were identified, such as GCG (Gcg), IRX2 (Irx2), LOXL4, and FAP for alpha cells (Figures 2A, 2B, and S5A); INS (Ins1, Ins2), DLK1, and IAPP (Iapp) for beta cells (Figures 2C, 2D, and S5B); PPY (Ppy), SERTM1, and SLC6A4 for gamma dells (Figures 2E and S5C); SST (Sst), LEPR, and BCHE for delta cells (Figures 2F and S6A); KRT19 (Krt19), CFTR (Cftr), MMP7, and AMBP for ductal cells (Figures 2G and S6B); and PRSS1, PNLIP, CTRB2, and REG1A for acinar cells (Figures 2H and S6C). Notably, ADCYAP1 was specifically expressed in beta cells in all human datasets, suggesting its potential as a new marker of beta cells (Figures 2C and S5B). ADCYAP1 gene encodes a pituitary adenylate cyclase activating polypeptide (PACAP), which is a secreted proprotein with the ability to activate adenylyl cyclase. PACAP is a membrane-bound enzyme that increases cyclic adenosine monophosphate levels, resulting in the transcriptional activation of multiple target genes. Other endocrine and exocrine cell-specific genes are summarized in Tables S1 and S2.

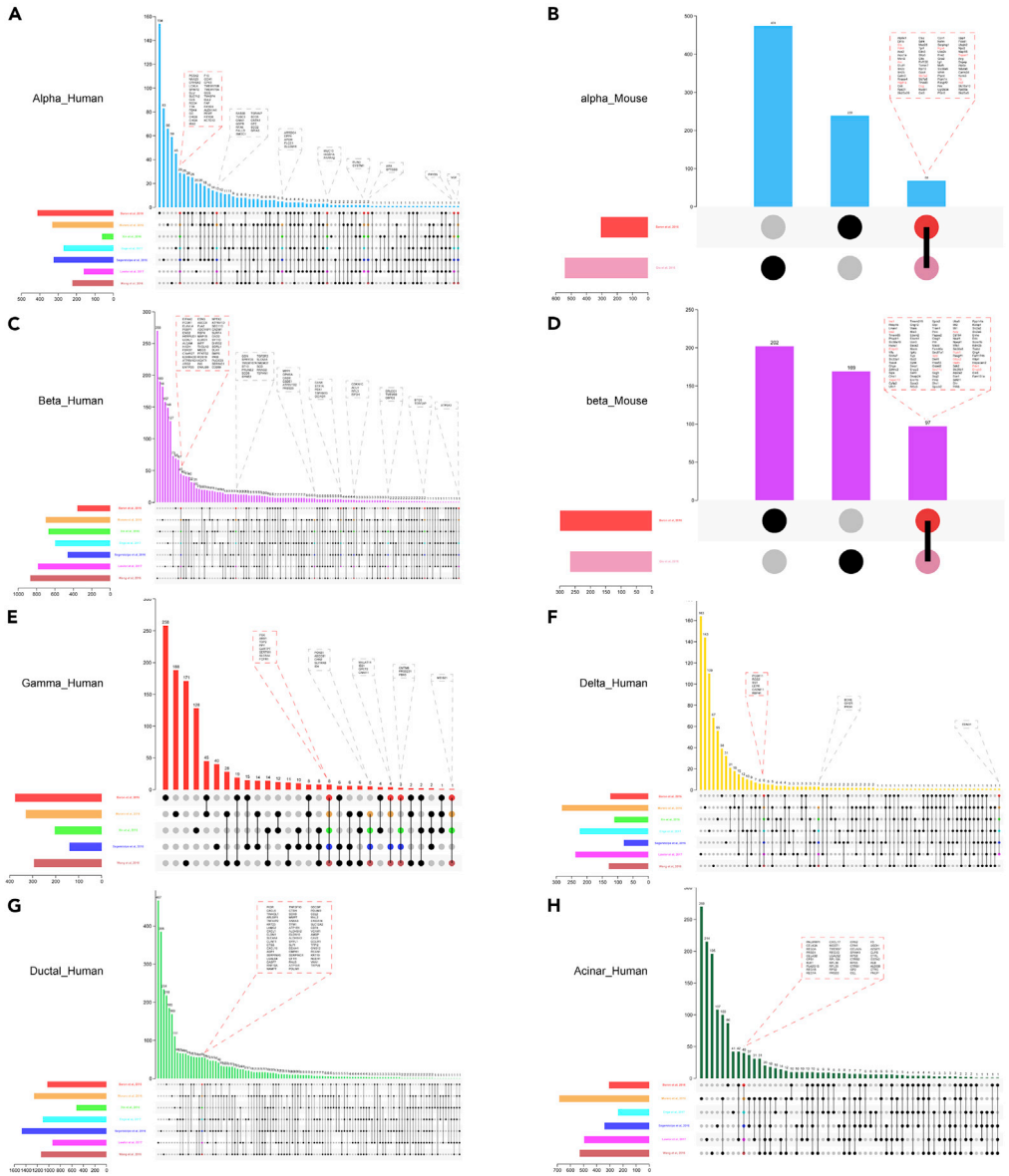


Figure 2. Robust marker genes of endocrine and exocrine cells are identified

(A–H) The upset plots showing common marker genes across all scRNA-seq datasets, which are labeled with different colors. The number of marker genes is shown with bar plots. The common marker genes are summarized in the box, and highlighted with red color when they are conserved across species. Colorful upset plots represent different cell types, including alpha (A and B), beta (C and D), gamma (E), delta (F), ductal (G), and acinar cells (H). The number of common marker genes of each cell type across multiple datasets is shown at the top of the bar chart.

Beta had enriched expression of GLRA1

T2D is a chronic disease caused by relative insulin deficiency and resistance. Beta dysfunction happens long before the blood glucose level was elevated. Accurate function evaluation of beta is crucial for the individual treatment of patients with diabetes.[23] In this regard, it is imperative to develop non-invasive and early diagnostic methods to evaluate the quantity and secretory function of beta cells. Our previous studies showed that a probe of [18F] bor-glycine was specifically enriched in the pancreas for positron emission tomography (PET)/computed tomography (CT) imaging, and the level of this probe was correlated linearly with the quantity of beta cells. The glycine receptors (encoded by GLRA1/2/3 and GLRB) and transporters (encoded by SLC6A5 and SLC6A9) are involved in glycine uptake of beta and maintenance of glycine homeostasis in pancreatic islets.[24] To further elucidate the molecular basis of this probe imaging, we

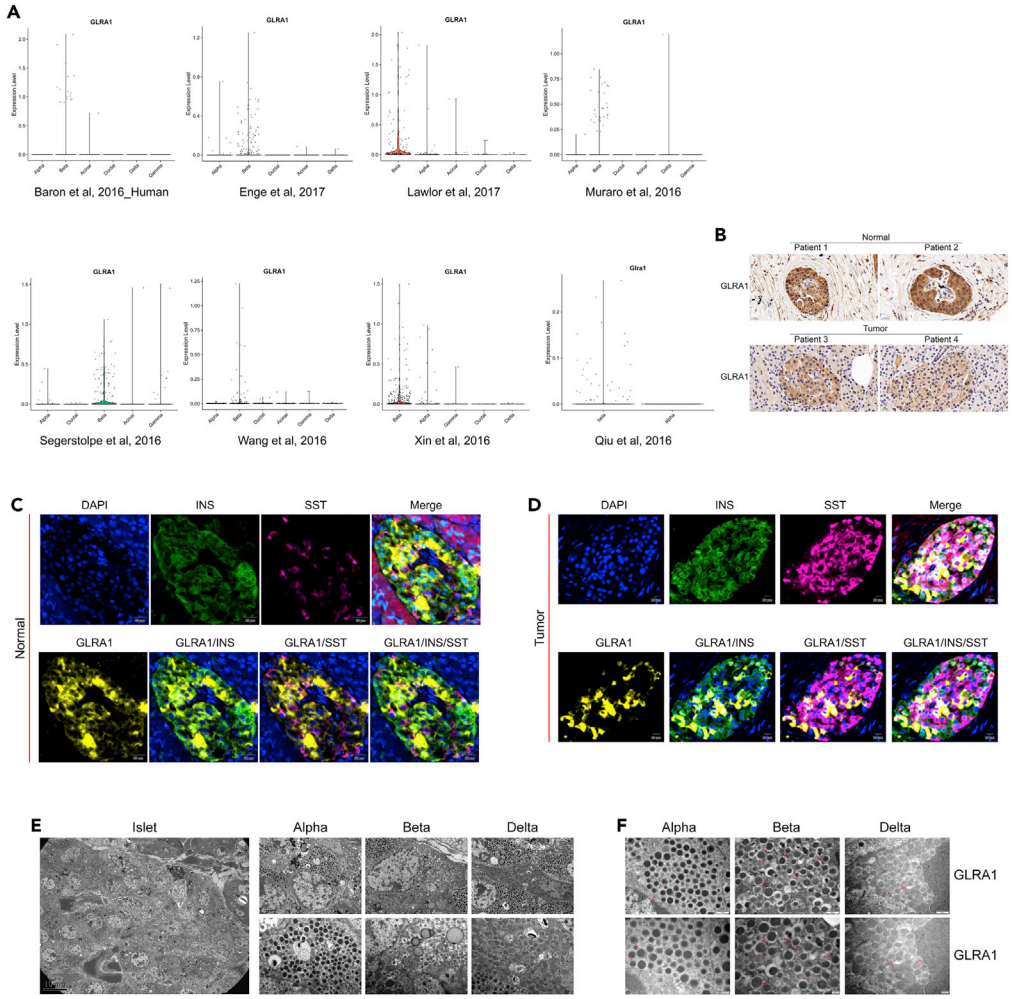


Figure 3. Beta cells have enriched expression of GLRA1 (Glra1)

(A) Violin plots showing the normalized expression of GLRA1 (Glra1) in endocrine and exocrine cell populations for each human and mouse scRNA-seq dataset.

(B) IHC of GLRA1 in human islets in PDAC and normal pancreatic tissues.

(C and D) Multi-color IHC staining of normal pancreatic (C) and PDAC (D) sections, and representative images of pancreatic islets are obtained. Expressions of INS (Green), SST (Pink), and GLRA1 (Yellow) are marked with different colors. Nuclei are counterstained with DAPI (blue).

(E) Representative images from transmission electron microscopy (TEM) of pancreatic islet cells. Endocrine cell types are distinguished according to their secretory granule morphology. (Left) An overview image. (Right) Magnification of each endocrine cell.

(F) Immunoelectron microscopy showing the expression and distribution of GLRA1 in pancreatic islets. Red arrows are pointing to examples of gold particles labeling GLRA1.

investigated the expression and distribution of glycine receptors and transporters in the pancreas at a single-cell resolution. We found that GLRA1 (Glra1) was restricted to human and mouse beta cells, although with low mRNA levels (Figure 3A). In addition, the beta subpopulation showed low GLRA3 (Glra3) and SLC6A9 (Slc6a9) expression (Figures S7A and S7C). GLRB (Glrb) was expressed across various endocrine and exocrine cells at relatively high mRNA levels (Figure S7B). GLRA2 (Glra2) and SLC6A5 (Slc6a5) were not detected using scRNA-seq analysis.

Next, we conducted immunohistochemistry (IHC) staining and found that human islets had enriched expression of GLRA1 in the normal pancreatic and PDAC tissues (Figure 3B). Multi-color IHC staining was also conducted to explore the distribution of GLRA1 in the pancreatic islets. Some beta cells, instead of all beta cells, had specifically high expression of GLRA1 in human islets in the normal pancreatic and pancreatic ductal

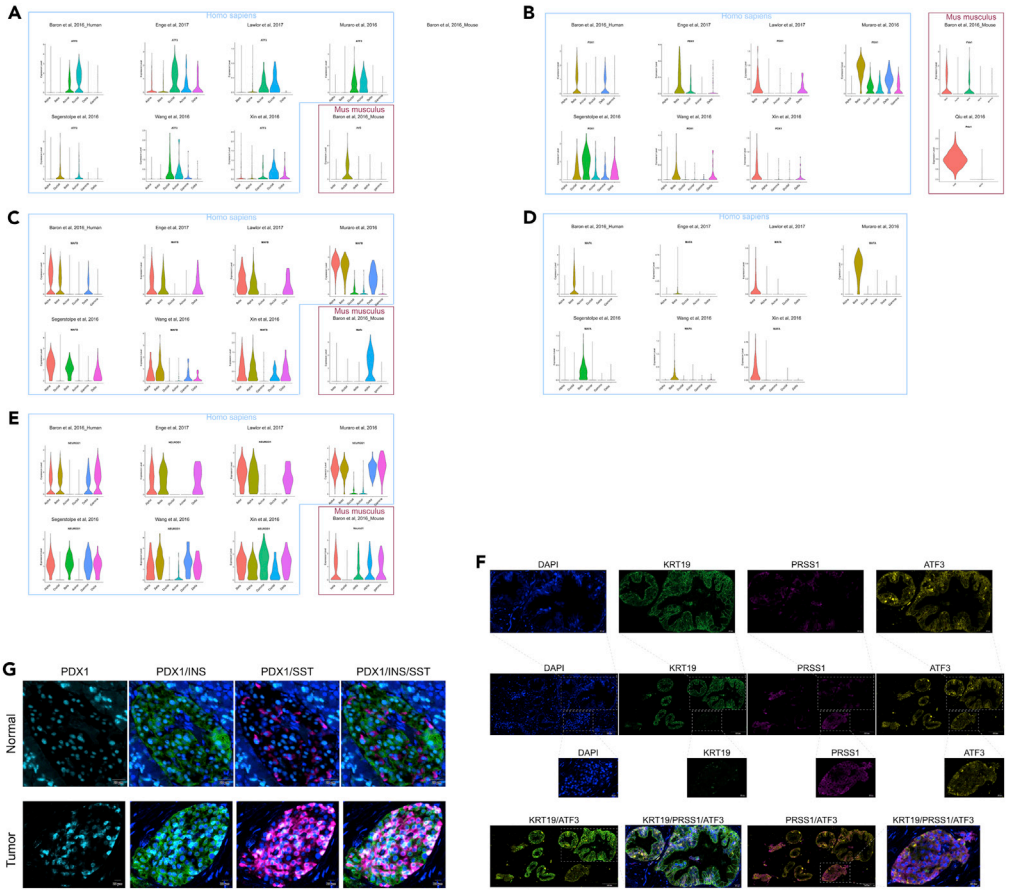


Figure 4. Cell-specific transcription factors (TFs) in endocrine and exocrine cells across species

(A–E) Violin plots showing the expression levels of identified cell-specific TFs using the SCENIC method, including ATF3 (Atf3), PDX1 (Pdx1), MAFB (Mafb), MAFA, and NEUROD1 (Neurod1). (F) A representative image of multi-color IHC staining from human PDAC sections. The depicted area within the section (white rectangle) is zoomed-in. Expressions of KRT19 (green), PRSS1 (pink), and ATF3 (yellow) are marked with different colors. Nuclei are counterstained with DAPI (blue).

(G) Multi-color IHC staining of normal pancreatic and PDAC sections, and representative images of pancreatic islets are the same as those in Figures 3C and 3D. Expressions of PDX1 are marked with azure fluorophore.

adenocarcinoma (PDAC) tissues, suggesting beta subpopulation heterogeneity for glycine-induced insulin secretion (Figures 3C and 3D). In the human islets of Langerhans, different endocrine cell types were distinguished according to their secretory granule morphology under transmission electron microscope. In line with what was previously found,[25,26] glucagon granules of alpha were dark with a thin halo, insulin granules of beta were lighter gray with a thick halo, mature granules had a crystalline core, and somatostatin granules of delta were homogeneously light gray without any halo (Figure 3E). Immunogold electron microscopic investigation revealed that gold particles labeling GLRA1 were mainly localized in the haloes of insulin granules of beta (Figure 3F). Therefore, beta cells specifically express GLRA1, which might demonstrate the possibility of evaluating the function of beta cells using the glycine-specific probe for PET/CT imaging.

Transcription factor expression in endocrine and exocrine cells

TFs are a family of DNA-binding proteins that are of vital importance for defining the molecular state of cells.[27] Therefore, we investigated the cell type-enriched TFs using the SCENIC method. A conservative and divergent TF pattern between human and mouse pancreas was found (Figures 4A–4E, S8A, and S8B). Multiple datasets supported that ATF3 (Atf3) was chiefly expressed in exocrine cells, including ductal and acinar cells, whereas NEUROD1 (Neurod1) was expressed in endocrine cells, including alpha, beta, gamma, and delta cells. Consistent with previous studies,[5,7,28] MAFA was restricted to beta cells. MAFB was found in both alpha and beta cells in human islets, whereas Mafb was restricted to alpha cells in mouse

islets. In addition, we found that both beta and delta cells had enriched expression of PDX1 (Pdx1) in human and mouse islets. Multi-color IHC staining was performed to validate ATF3 and PDX1 expression patterns in human pancreatic tissue. In panel 1, we simultaneously detected ATF3, KRT19 (ductal cell), and PRSS1 (acinar cell) expression and found that cells in the ductal and acinar regions coexpressed ATF3, KRT19 and ATF3, and PRSS1, respectively (Figure 4F). In panel 2, we simultaneously detected INS (beta), SST (delta), and PDX1 expression, and found that PDX1 was localized in the nucleus of beta and delta cells in normal pancreatic and PDAC tissues (Figure 4G).

scRNA-seq uncovers cell atlas of islet of Langerhans in type 2 diabetes mice

To delineate the cell atlas of islet of Langerhans in T2D mice, the Lepr knockout model was used. Pancreatic islets were isolated from three control (BKS_WT; m/m) and three T2D (BKS_DB; db/db) mice at the ages of 9 weeks. The basic characteristics of mice are summarized in Figure 5A. Mice in the BKS_DB group had significantly higher blood glucose and GHbA1c levels than those in the BKS_WT group. Genotype identification was performed to validate the Lepr knockout (Figures S9A and S9B). We also compared the vital organs between the two groups using hematoxylin-eosin staining (Figure S9C). No significant changes were observed in the spleen, heart, lung, and kidney, whereas BKS_DB mice had a disorganized pancreas and droplet formation in the liver compared with BKS_WT mice. The isolated islets were cultured overnight and enzymatically digested to generate single-cell suspensions. We conducted scRNA-seq and stringent quality control, and acquired 6,898 cells from the BKS_DB mice and 6,729 cells from the BKS_WT mice. The workflow of scRNA-seq analysis is shown in Figure 5B.

A total of 16 original clusters were obtained using uMAP and integrated into 11 known cell types according to the well-known markers of cell populations mentioned earlier (Figures 5C–5F). We delineated a complete pancreatic islet cell atlas in T2D mice, including alpha, beta, gamma, delta, ductal, acinar cells, macrophages, fibroblasts, endothelial cells, b cells, and t cells. Remarkably, the beta subpopulation was composed of five clusters, whereas each of the other cell subpopulations was composed of only one cluster, which demonstrated that beta cell had more obvious heterogeneity. Endocrine-specific genes, including Gcg, Ins1, Ppy, and Sst, were almost exclusively expressed in the corresponding cell subpopulations, indicating less ambient RNA contamination in our dataset. The marker genes of each cell type were identified using FindMakers function (Figure 5G, Tables S1 and S2). Consistent with the results obtained using open datasets, our dataset also supported the robust marker genes of endocrine and exocrine cells identified earlier. However, we did not detect Adcyap1 in any cell types. Multi-color IHC showed ADCYAP1 was expressed in various cell types, in addition to beta cells in human islets (Figure S9D). Immunogold electron microscope assay indicated delta cells had enriched expression of ADCYAP1, which was inconsistent with scRNA-seq analysis (Figure S9E). Beta cells were the predominant cell type in the islets of BKS_WT and BKS_DB mice (Figure 5H). The glycine receptor (Glra1, Glrb) and transporter (Slc6a9) were expressed in beta cells with low mRNA levels, whereas Glra2 and Slc6a5 were not detected (Figure S10A).

The SCENIC method was used to identify cell-specific TFs in our dataset (Figure 6A). In line with previous findings obtained using open datasets, Irx2 and Mafb were restricted to alpha cells, whereas Mafa was restricted to beta cells (Figure 6B). Both beta and delta cells showed enriched expression of Pdx1. Furthermore, Neurod1 was a marker of pan-endocrine cells. However, Atf3 was expressed in delta and gamma cells, in addition to ductal cells.

Pathological alterations of beta subpopulation in type 2 diabetes

Decreased beta mass and function is the hallmark of T2D. To further investigate the pathological alterations of beta cells in T2D, we first isolated beta subpopulation and conducted a separate clustering analysis to allow for subtle differences between the BKS_WT and BKS_DB groups. A total of six distinct beta subclusters were identified (Figures 6C and 6D). All subclusters expressed Ins1 and Ins2, confirming their beta cell identity (Figure 6E). The marker genes of each beta subcluster were obtained (Figure 6F and Table S3). Intriguingly, subclusters 0 and 1 were mainly derived from the BKS_DB and BKS_WT mice, respectively, which represents beta cells in the T2D and control groups (Figure 6G). We compared subclusters 0 and 1 and found many differentially expressed genes (DEGs) (Figure 6H). Subcluster 0 had higher expression of Fos and Jun and lower expression of Ins1 and Mafa than subcluster 1. GO and KEGG analyses showed that DEGs were involved in transcription regulator activity, ubiquitin protein ligase binding, and MAPK and TNF signaling pathways (Figure 6I). All beta subclusters showed low proliferation activity when examining the expression levels of Mki67 and Pcna (Figures S10B and S10C).

A

Group	Age (week)	Sex	Genotype (Lepr)	Weight (g)	Abdominal circumference (cm)	Glucose (nmol/L)	GHbA1c (%)
BKS_WT_1	9	male	WT/WT	23.96	7.8	6.2	4
BKS_WT_2	9	male	WT/WT	21.96	6.8	7.1	4.1
BKS_WT_3	9	male	WT/WT	22.41	7.2	6.8	4.1
BKS_DB_1	9	male	KO/KO	46.22	10.2	27	8
BKS_DB_2	9	male	KO/KO	46.56	11.1	28	7.1
BKS_DB_3	9	male	KO/KO	52.29	10.6	27.9	7.4

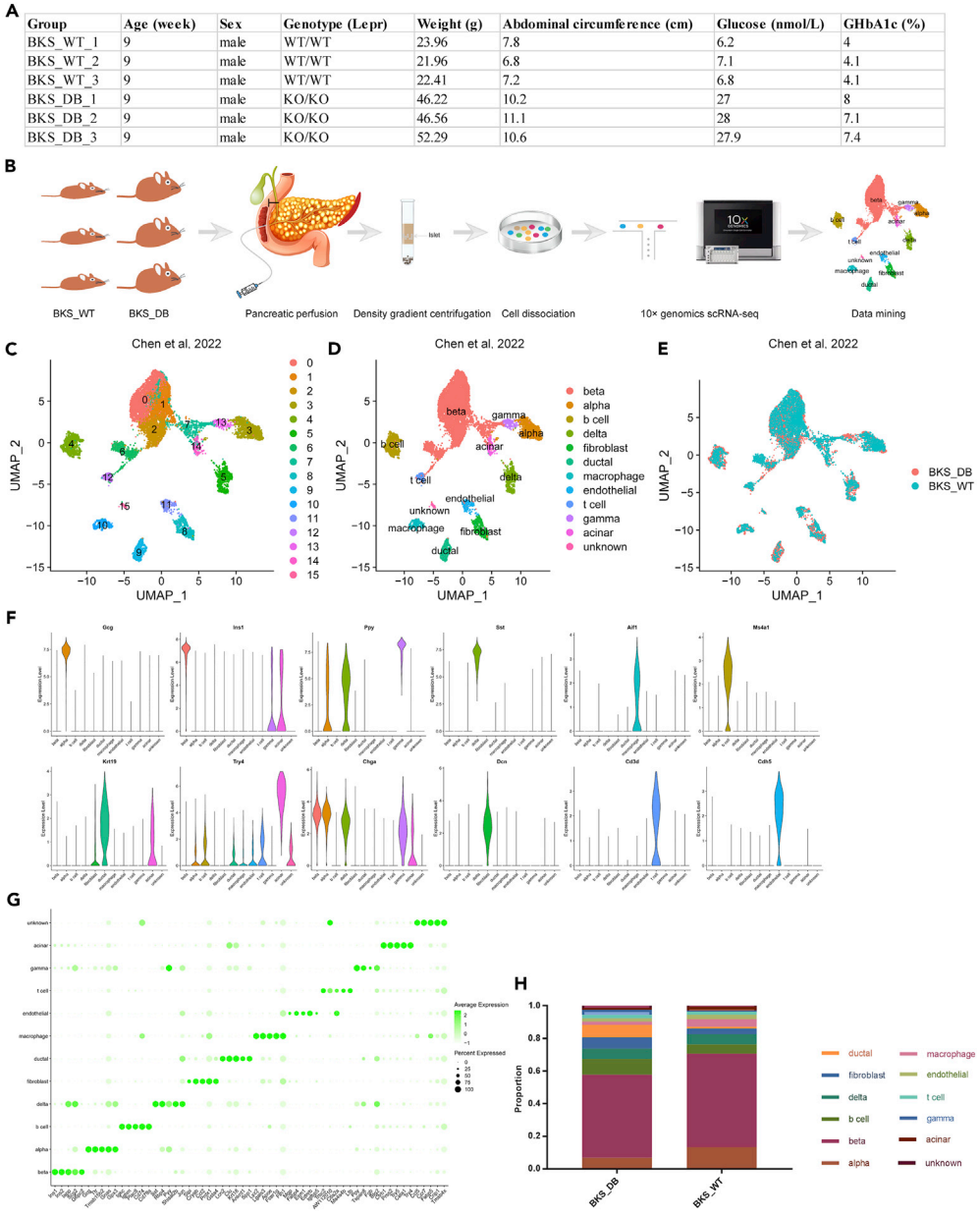


Figure 5. scRNA-seq delineates the cellular landscape of pancreatic islets in type 2 diabetes mice

(A) General characteristics of mouse specimens in scRNA-seq.

(B) Graphical scheme describing the workflow of scRNA-seq analysis.

(C–E) uMAP showing the original clusters (C), identified cell types (D), and group information (E) with different color code.

(F) Violin plots showing the expression levels of known marker genes of cell populations among identified cell types.

(G) Dot plot showing the top 5 cell type-specific genes. Size of dots indicates the proportion of cells expressing a selected marker, and intensity of color indicates the average expression level.

(H) The proportion of cell populations in the control (BKS_WT) and type 2 diabetic (BKS_DB) mice.

Beta dysfunction is characterized by abnormal glucose-stimulated insulin secretion (GSIS), which is dependent on glucose uptake and glycolysis.[29] To gain insight into the pathological alterations related to GSIS in T2D, we examined the expression levels of relevant genes among the beta subclusters in our dataset (Figure 7A). Subcluster 0 had significantly lower expression levels of Slc2a2 and G6pc2 than subcluster 1. Slc2a2 encodes the glucose transporter, Glut2, which is required for glucose sensing. Therefore, the fact that beta cells in T2D became insensitive to blood glucose levels might be attributed to decreased Glut2 levels.

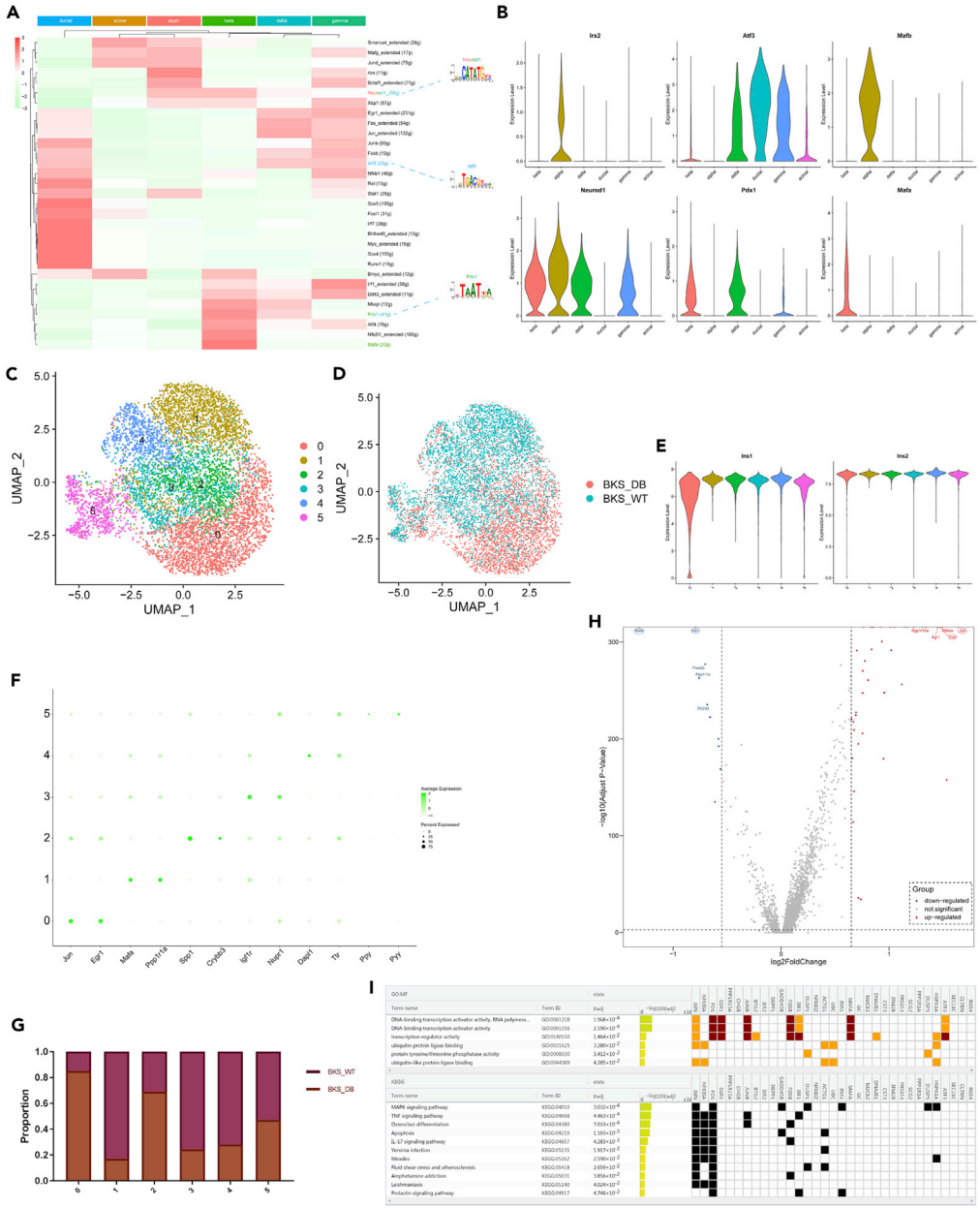


Figure 6. The identification of beta cell heterogeneity

(A and B) (A) Heatmap of the activation scores of each endocrine and exocrine cell type for expression regulated by transcription factors (TFs) using the SENIC method. Cell types are indicated on top. The inferred motif of the selected TF is shown on the right. (B) Violin plots showing the expression levels of the selected TFs in endocrine and exocrine cell types in our scRNA-seq dataset.

(C and D) uMAP showing the original beta subclusters (C) and group information (D).

(E) Violin plots showing the expression levels of Ins1 and Ins2 across beta subclusters.

(F) Dot plot showing the top 2 cell type-specific genes in beta subclusters. Size of dots indicates the proportion of cells expressing a selected marker, and intensity of color indicates the average expression level.

(G) The proportion of beta subclusters between the BKS_WT and BKS_DB groups. Legends are shown on right.

(H) Volcano plot showing the differentially expressed genes (DEGs) between subclusters 0 and 1. Red and blue dots represent the upregulated genes in subclusters 0 and 1, respectively.

(I) Top 50 DEGs between subclusters 0 and 1 were used to conduct GO and KEGG analysis (logfc.threshold = 0.5, q-value < 0.05). Top GO and KEGG terms are shown. Bar plots are colored according to their -log10p values. Heatmaps showing the expression levels of DEGs.

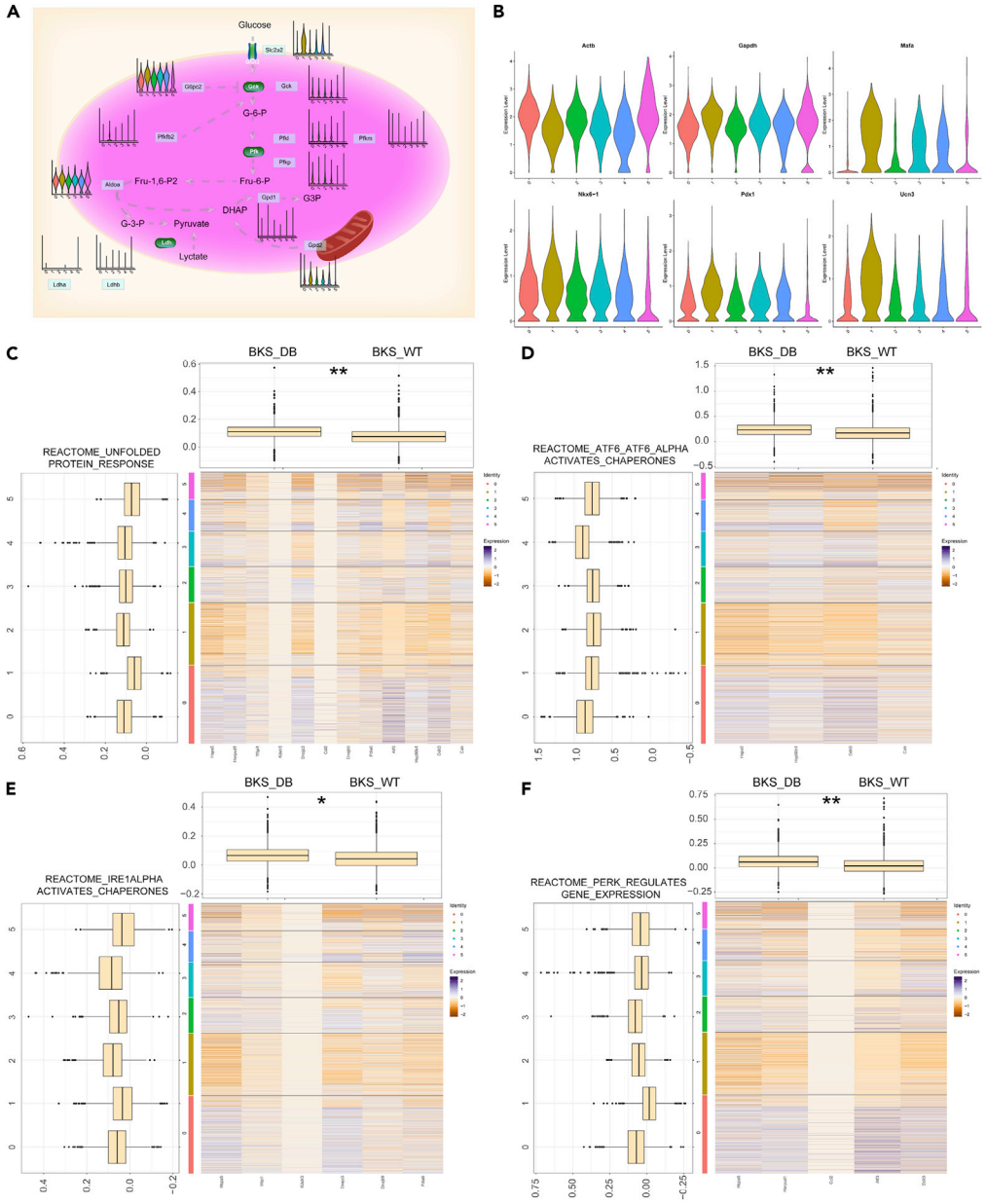


Figure 7. Glycolysis, dedifferentiation, and unfolded protein response of the beta subclusters

(A) Schematic diagram of glycolysis. Violin plots showing the expression levels of the key components of glycolysis across beta subclusters.

(B) Violin plots showing the expression levels of the mature beta cell markers across beta subclusters. Both Actb and Gapdh are the reference genes.

(C–F) Unfolded protein response- (C) and ATF6 (D), IRE1 (E), and PERK (F) pathways-related scores using the AddModuleScore function. (Left) Boxplots of expression of reactome pathways-related signatures (defined as the mean normalized expression of genes in the signatures) across beta subclusters. (Right) Heatmap of Z-scored mean expression of genes in signatures. (Top) Boxplots of total expression of signatures across specimens. *$p < 0.05$ and **$p < 0.01$, respectively (independent samples t test).

G6pc2 encodes a subunit of glucose-6-phosphatase that catalyzes the hydrolysis of glucose-6-phosphate. Aldoa (encoding a member of the class I fructose-bisphosphate aldolase protein family) was not significantly differentially expressed among the beta subclusters. Beta cells had relatively low expression levels of glycolytic enzymes from glucose-6-phosphate to lactate, including Gck, Pfkfb2, Pfkl, Pfkm, Gpd1, Ldha,

and Ldhb. In addition, higher Pfkp and lower Gpd2 expression levels were detected in subcluster 0 compared with subcluster 1. Gpd2 (encoding glycerol-3-phosphate dehydrogenase 2) is involved in NAD+ regeneration essential for GSIS. Taken together, beta cells in T2D exhibited impaired GSIS.

We also compared the marker genes of beta cell maturation across beta subclusters (Figure 7B). The expression levels of Mafa, Nkx6-1, Pdx1, and Ucn3 were significantly downregulated in subcluster 0 compared with subcluster 1, whereas reference genes, Actb and Gapdh, were comparable across these beta subclusters. In line with previous studies,[30,31] these results supported that beta cells might undergo dedifferentiation under constant stimulation of high blood glucose levels in T2D mice, worsening beta dysfunction. Thus, beta cell redifferentiation might be a promising strategy for diabetes treatment.

ER stress and hypoxia in beta cells in type 2 diabetes

The unfolded protein response (UPR) is the main ER stress response system and is triggered by high levels of insulin synthesis in the beta cells.[32] The UPR system consists of ATF6-, IRE1-, and PERK-related pathways, which play a significant role in maintaining healthy beta function.[33] We calculated the UPR score based on pathway-related gene sets using the AddModuleScore function. Of note, subcluster 0 had significantly higher scores and gene expression (Hspa5, Herpud1, Dnajc3, and Ddit3) for the UPR pathway than subcluster 1, and beta cells in the BKS_DB group had significantly higher scores than those in the BKS_WT group (Figure 7C). Similar trends were observed in the ATF6-, IRE1-, and PERK-related pathways (Figures 7D–7F). Our findings proved that UPR pathways were activated in beta cells in mice with T2D.

In agreement with recent findings that exposure to high glucose levels induced hypoxia-related pathways in beta cell lines[34] and islets from mice suffering from diabetes were hypoxic[35–37] we found that subcluster 0 significantly upregulated hypoxic signatures, such as Jun, Fos, and Ppp1r15a, and had higher hypoxic scores than subcluster 1. Higher hypoxic scores were observed in beta cells in the BKS_DB group (Figure S10D). These results prompted further investigation into the mechanism of activation of hypoxic pathway in beta cells in diabetes. In addition, all beta subclusters expressed a similar average degree of genes in glycolysis and fatty acid metabolism pathways, and there were no significant differences between the BKS_WT and BKS_DB groups.

To explore the transcriptional changes in nonbeta cells in UPR pathways, we isolated each cell type and conducted clustering analysis (Figure S11A). Overall, there were no significant differences in UPR pathway scores between the BKS_WT and BKS_DB groups for most nonbeta cell subpopulations (Figure S11B). Gamma cells from BKS_DB mice had higher ATF6- and PERK-related pathways scores than BKS_WT mice. Thus, both beta and gamma underwent ER stress under high blood glucose levels in T2D mice.

Key modules and TFs across beta subclusters

To further figure out transcriptional profiles of beta cells in T2D mice, we randomly picked out 200 cells from each beta subcluster to construct a new gene-cell matrix for weighted gene co-expression network analysis (WGCNA). Cluster analysis for the samples was performed (data not shown), which was consistent with previous unsupervised clustering results in Figure 6C. After selecting the appropriate soft-thresholding power (n = 3), gene network construction was conducted (Figure 8A). We observed significant correlation between the blue module and subcluster 0 (cor = 0.81; p = 1.5e-178) (Figure 8B). Then, gene list of the blue module was exported for visualization of network connections among these genes using Cytoscape software (Figure 8C). The hub genes in the blue module included Jun, Junb, Fos, and Fosb. Therefore, key modules of beta subclusters were identified.

The SCENIC method was used to identify beta subcluster-specific TFs. Notably, subcluster 0, representing beta cells in T2D, had enriched expression of Jun, Junb, Fos, and Fosb, whereas downregulated the expression of Pdx1 and Nkx6-1, consistent with the WGCNA analysis of the present study (Figures 8D and 8E). These results suggested that pancreatic beta cells in T2D exhibited new transcriptional profiles and lost the features of mature beta cells during dedifferentiation.

Pseudotime analysis uncovers gradual changes of beta cells in normal and diabetic mice

Recently, pseudotime analysis of scRNA-seq has been widely used to infer cellular trajectories and observe gradual changes in cell subpopulations.[12,18] We isolated beta cells in subcluster 0 and 1 to perform pseudotime analysis (Figures 8F, 8G, S10G, and S10H). The tree structure in two dimensions began with beta

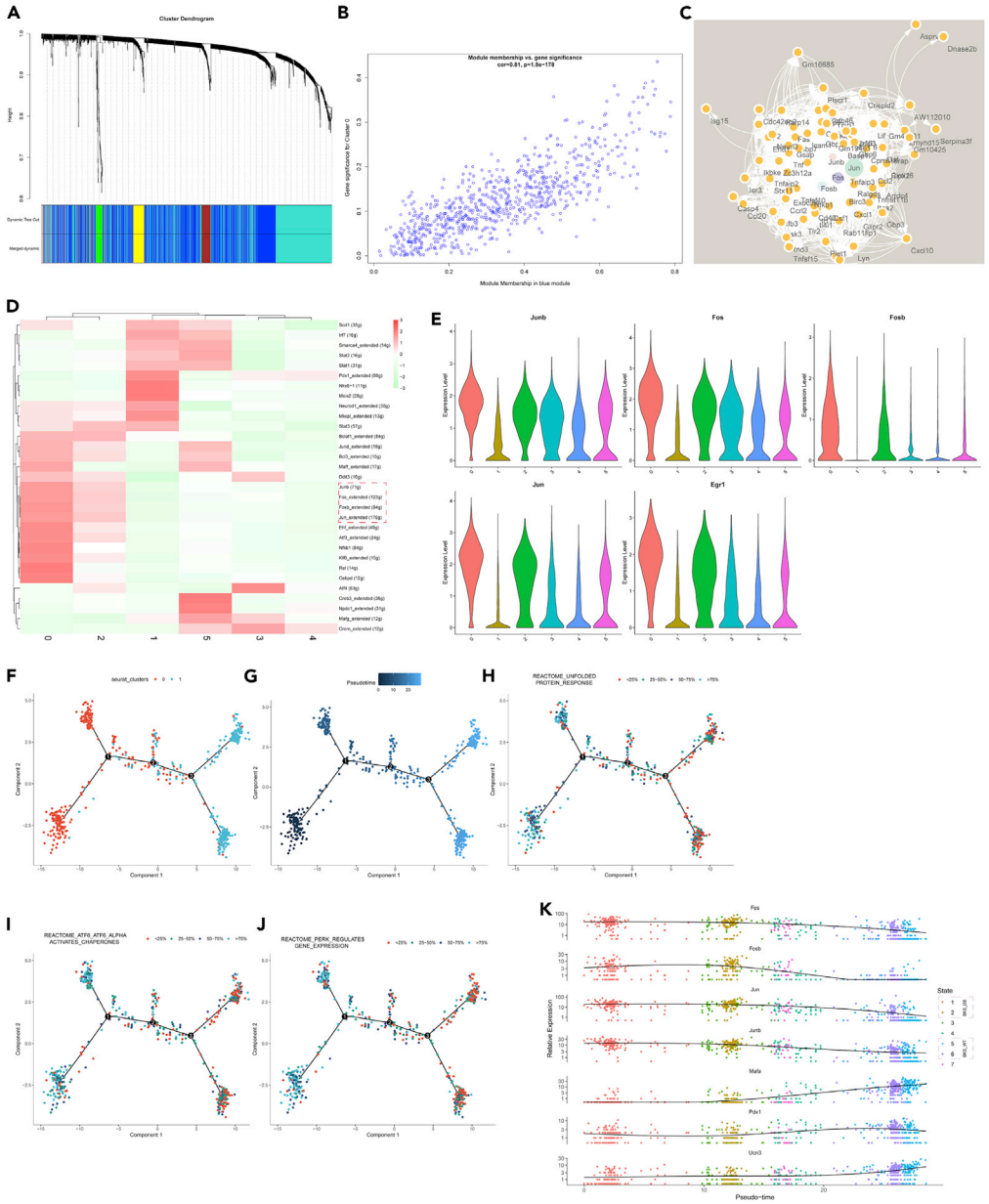


Figure 8. WGCNA and pseudotime analysis of the beta subclusters

(A) Cluster dendrogram of genes, with dissimilarity according to topological overlap using the Dynamic Tree Cut algorithm, together with assigned merged and original module colors.

(B) Scatterplot showing gene significance (GS) for subcluster 0 versus module membership (MM) in the blue module. There is a significant correlation between subcluster 0 and blue module.

(C) Visualization of network connections among the most connected genes in the blue module representing subcluster 0 using the Cytoscape software.

(D) Heatmap of the activation scores of each beta subcluster for expression regulated by TFs using the SENIC method. Cell types are indicated below.

(E) Violin plots showing the expression levels of the key transcription regulators across beta subclusters.

(F–J) Pseudotime analysis exploring the cellular trajectories of beta subclusters with high variable genes. Each dot represents one cell, which is color coded according to their original cluster (F), pseudotime score (G), unfolded protein response- (H) and ATF6 (I) and PERK (J) pathways-related scores.

(K) Relative expression of selected genes among different states in the cellular trajectories.

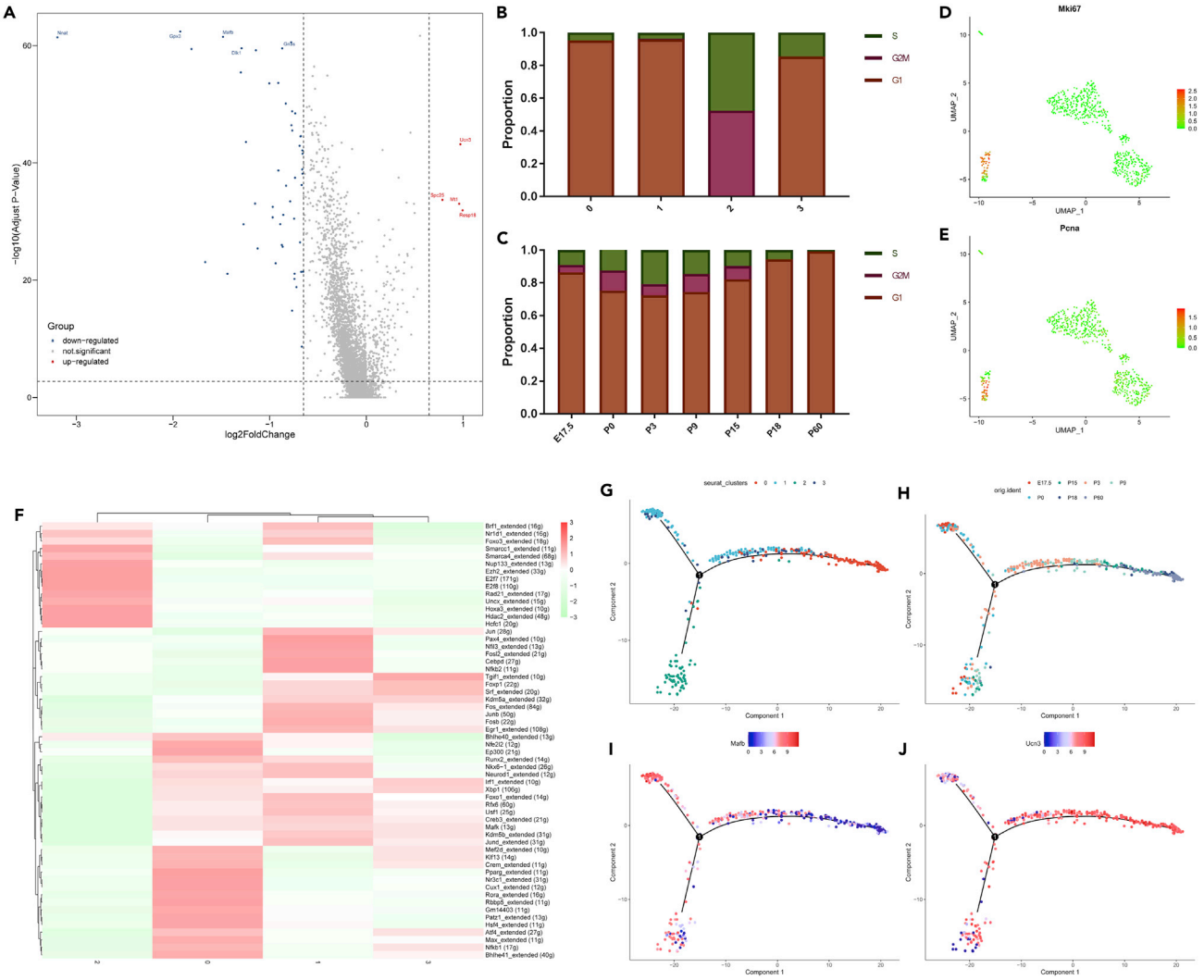


Figure 9. scRNA-seq deciphering beta maturation features

(A) Volcano plot showing the differentially expression genes (DEGs) between beta subclusters 0 and 1 in Qiu's dataset, related to Figure S12C. Red and blue dots represent the upregulated genes in subclusters 0 and 1, respectively.

(B and C) Proportion of cell cycle phases in the beta subclusters (B) and specimens (C).

(D and E) uMAP plots showing expression of selected proliferation marker genes in the beta subclusters. Legend shows a color gradient of normalized expression. Red dots represent the beta cells with high expression levels of genes.

(F) Heatmap of the activation scores of each beta subcluster for expression regulated by TFs using the SENIC method. Cell types are indicated below.

(G–J) Pseudotime analysis exploring the cellular trajectories of beta subclusters with high variable genes. Each dot represents one cell, which is color coded according to their original cluster (G), group information (H), expression levels of Mafb (I), and Ucn3 (J).

cells in subcluster 0, followed by mixed beta cells, and finally with beta cells in subcluster 1. To gain insight into the gradual changes of beta cells for UPR pathways, we converted UPR scores to categorical variables and found that UPR scores gradually decreased from the left to the right side of the tree structure (Figures 8H–8J). This result showed that the UPR pathways were gradually activated during the progression of diabetes. In addition, RNA velocity was performed to verify the dynamic change from subcluster 0 to 1 (Figures S12A–S12C). Intriguingly, we also observed changes in mature beta markers (Mafa, Pdx1, and Ucn3) and Jun, Junb, Fos, and Fosb from subcluster 0 to subcluster 1 (Figures 8K and S13A).

scRNA-seq deciphering beta maturation features

Beta cells in human and mouse datasets retrieved from open databases were isolated for separate clustering analysis. Multiple beta subclusters were identified in all datasets, except for Lawlor's dataset

(Figure S13B). We also found a larger degree of variation in the beta subcluster composition of each specimen (Figure S13C). These results revealed beta heterogeneity under different physiological and pathological conditions, including islet development, diabetes, and aging.

We noticed that the beta subcluster composition changed regularly from the fetal to adult stage in Qiu's dataset, including embryonic day 17.5 (E17.5), postnatal day 0 (P0), P3, P9, P15, P18, and P60. Proportion of subcluster 0 increased gradually from E17.5 to P60, whereas subcluster 1 changed with almost the opposite tendency (Figure S12C). Therefore, we compared subclusters 0 and 1. The volcano plot showed that Ucn3, Spc25, Mt1, and Resp18 were upregulated in subcluster 0, whereas Gpx3, Mafb, Dlk1, and Gnas were upregulated in subcluster 1 (Figure 9A). Ucn3 was a mature beta cell,[38] and Mafb was an immature beta cell marker that was highly expressed in beta cells with early stage.[39] Thus, subclusters 0 and 1 represented beta cells with late and early stages. We then inferred the cell cycling composition of each beta subcluster (Figures 9B and 9C). Beta cells with S and G2M phases were predominant in subcluster 2, plus they had enriched expression of Mki67 and Pcna, therefore identified as proliferating beta cells (Figures 9D and 9E). In addition, we observed that beta cells with G1 phase were predominant in P60, suggesting that beta cells with late stage were not proliferating actively. The SCENIC method was used to investigate cell cluster-specific TFs (Figure 9F). Subcluster 0 specifically expressed many TFs, such as Pparg, Nr3c1, Nfkb1, and Max. Pseudotime analysis revealed developmental trajectory of beta cells from the early stage (E17.5, P0) to the late stage (P18, P60) (Figures 9G and 9H). Mafb and Ucn3 exhibited dynamic changes during beta cells development (Figures 9I and 9J). The expression of Mafb decreased dramatically from P3 to P9.

Web-based interactive analysis for pancreatic islet scRNA-seq datasets

In recent years, advances in scRNA-seq have allowed for acquiring transcriptional profile of each single cell, instead of the average expression level of the whole tissue. Such a paradigm shift is expected to deepen our understanding of cell heterogeneity and the pathological alterations underlying diseases. However, scRNA-seq analysis based on R and Python tools was user unfriendly for experimental biologists and doctors. Thus, we developed a Web-based interactive tool for the data mining of pancreatic islet scRNA-seq datasets (Figure S14). We collected seven human and three mouse scRNA-seq datasets for pancreatic islets. It is easy to delineate a complete pancreatic islet cell atlas and randomly search for any gene expression level in a specific cell type. This Web-based interactive tool is a good resource for studying pancreatic islet development, aging, and diabetes, and is freely available at http://pica.novelbrain.com.

DISCUSSION

In this study, we have delineated the cellular landscape of islets of Langerhans and identified the robust marker genes of each endocrine and exocrine cell type through cross-species scRNA-seq analysis. We identified almost all endocrine cell types, including alpha, beta, gamma, and delta cells in both human and mouse islets. In particular, gamma cells account for only a small percentage of islets, whose transcriptional profile has not been elucidated before. A large number of gamma cells and their marker genes and TFs were identified in this study. However, we did not identify any epsilon cells, which might be attributed to strict quality control to filter out latent epsilon cells. In addition, we also identified ductal cells, acinar cells, endothelial cells, fibroblasts, macrophages, stellate cells, and b cells, which were not removed during pancreatic islet isolation. In previous studies, the marker genes of cell populations varied greatly owing to different scRNA-seq platforms and a limited sample size. Here, more reliable marker genes were obtained by taking the common marker genes of cell populations across all scRNA-seq datasets. Intriguingly, ADCYAP1 was restricted to beta cells in all human datasets using scRNA-seq analysis (Figure S5B). However, we observed different results using multi-color IHC and immunogold electron microscopy. ADCYAP was expressed at the protein level in various cell types, in addition to beta cells. Thus, ADCYAP is a beta cell-specific biomarker only at the mRNA level.

Droplet-based methods are widely used to perform scRNA-seq. Nevertheless, ambient RNA in the cell suspension will be counted along with a cell's native mRNA, resulting in cross-contamination between different cell populations.[40,41] We observed obvious ambient RNA contamination in endocrine cells, instead of in other cell types, in scRNA-seq datasets from open databases. In contrast, only slight ambient RNA contamination was found in our scRNA-seq dataset of the control and T2D mice. After strict quality control, we acquired single-cell transcriptomes of 13,627 cells (6,729 cells from the control mice versus 6,898 cells from the T2D mice) using 10X chromium microfluidic chips, showing a huge advantage in the

number of captured cells compared with previous studies.[5–11,20] There was great variation in the human islet composition, and human donors had different ages and therapeutic regimens. These confounding factors may cover up the underlying differences between ND and T2D donors, especially in scRNA-seq datasets with small sample sizes. However, mouse donors are homogeneous with the same age, sex, and lifestyle and therefore are more likely to reveal subtle pathological alterations in diabetes. We provide a valuable resource for deciphering the pathological alterations of T2D by conducting scRNA-seq for the control and T2D mice.

Pancreatic islets are vital micro-organs involved in maintaining blood glucose homeostasis, which are mainly composed of various hormone-secreting cell types, including alpha cells (glucagon), beta cells (insulin), gamma cells (pancreatic-polypeptide), delta cells (somatostatin), and epsilon cells (ghrelin).[42] Many studies have shown that islet composition varies not only between species but also within species.[21,22,43] Mouse islets have a well-defined structure with a central core of beta cells (60%–80%), surrounded by a layer of other endocrine cells, including alpha cells (15%–20%), delta cells (<10%), and gamma cells (<1%).[44] In this study, we also found that mouse islets consist mostly of beta cells (>60%), and a small percentage of alpha, gamma, and delta cells. Compared with mouse islets, human islets tend to have more alpha and fewer beta cells, and a more scattered organization of endocrine cells. In line with previous findings, islet composition heterogeneity across human donors was identified, and human islets were mainly composed of alpha and beta cells, and few delta and gamma cells in this study. Furthermore, we observed that the proportion of beta cells decreased with age, and the number of alpha and beta cells dramatically decreased in a T1D donor. Overall, the composition of human islets changes according to individual variation, physiological, and pathological conditions.

The direct determination of pancreatic beta cell mass is of great importance in the evaluation of the diabetes status. Previous studies have developed non-invasive imaging methods to accurately measure beta cell mass based on optical and PET platforms. Both Newport green and dithizone are fluorescent probes, and they have been used for *ex vivo* imaging of islets to facilitate pancreatic islet isolation and transplantation.[45] In 2000, 18F-FDG (2-deoxy-2-18F-fluoro-D-glucose) was first used for PET imaging to evaluate the endocrine pancreas in rats with diabetes.[46,47] Subsequently, many beta cell-specific probes were introduced, such as C-DTBZ,[48] 18F-AV-133,[49] 68Ga-NOTA-exendin-4,[50] and PiF.[51] However, there is still a dearth of an ideal probe to specifically bind with beta cells. Our previous study developed a probe of [18F] bor-glycine, which showed good performance in PET/CT imaging of the pancreas. The level of this probe correlated linearly with beta cell mass. Glycine receptors and transporters are involved in the glycine uptake of beta and maintenance of glycine homeostasis in pancreatic islets.[24] To further investigate the underlying molecular mechanism of imaging, we found that GLRA1 was restricted to beta cells using scRNA-seq, multi-color IHC, and immunogold electron microscopy. Therefore, we speculate that GLRA1 plays a crucial role in [18F] bor-glycine based PET/CT imaging, which remains to be explored.

SLC2A2 encodes the glucose transporter isoform GLUT2. The encoded protein mediates facilitated bidirectional glucose transport in the liver, intestine, kidney and pancreatic beta cells.[52] GLUT2 is required for GSIS as a key glucose sensor in beta cells.[53] In diabetic mice and rats, impaired GSIS is associated with significantly reduced expression of Glut2.[54] Transgenic expression of Glut2 in beta cells restored normal GSIS.[55] In this study, six beta subclusters were identified, with subclusters 0 and 1 representing beta cells in T2D (BKS_DB) and control (BKS_WT) mice, respectively. Consistent with these studies, we revealed that subcluster 0 significantly reduced Slc2a2 expression compared with subcluster 1 through scRNA-seq analysis. In addition, G6pc2 encoding a rate-controlling enzyme of glycolysis was downregulated in subcluster 0. We demonstrated that beta cells in T2D had impaired GSIS by downregulating key components of glycolysis. The marker genes involved in depolarization and calcium levels that are related to GSIS remain to be investigated.

Beta cells synthesize insulin with great efficiency, approximately half of the total amount of protein synthesized. The well-developed and normal ER is essential for insulin secretion of beta cells in response to high glucose stimulation.[56] A growing body of research has demonstrated a correlation between ER malfunction and T2D.[57,58] Beta cells compensate for the increase in blood glucose by increasing insulin secretion capacity, resulting in the accumulation of misfolded proteins and the activation of UPR pathways. In this study, we calculated the UPR scores of each beta cell using the AddModuleScore function and found that beta cells in T2D (BKS_DB) mice showed significantly higher UPR scores than control (BKS_WT) mice, which

indicated the ER stress in T2D. However, prolonged and unmitigated ER stress will activate the apoptosis-related pathways and cause cell death.[59] The molecular mechanism underlying the loss of ER homeostasis, which might be a key drug target, remains to be elucidated.

Overall, our work uncovers the cell landscape of the islets of Langerhans, identifies the robust marker genes and TFs of each endocrine and exocrine cell type, and deciphers the pathological alterations of T2D through scRNA-seq analysis. In addition, we developed a Web-based interactive tool, creating new opportunities for the data mining of pancreatic islet scRNA-seq datasets.

Limitations of the study

This study had several limitations. First, droplet-based scRNA-seq datasets exhibited the ambient RNA contamination, especially in endocrine cells, which might confound the biological interpretation. Second, this study had a limited sample size for scRNA-seq of the control and T2D mice. Third, the proportion of delta, gamma, exocrine (ductal and acinar cells), and mesenchymal cells are relatively lower, and their heterogeneities were hard to be identified in this study. Fourth, mice used for scRNA-seq were 9 weeks old, thus the pathological alterations on other stages remain unclear. Finally, Epsilon cells were not identified in our study. The implementation of large-scale scRNA-seq cohorts and advances in bioinformatics methods will hopefully identify the characteristics of epsilon cells, the minority cell population in mice and human islets.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Human subjects
  - Animals
- METHOD DETAILS
  - scRNA-seq datasets from open databases
  - Hematoxylin-eosin (HE) and IHC stating
  - Pancreatic islet isolation
  - Pancreatic islet dissociation and cell purification
  - Cell capture and cDNA library preparation
  - scRNA-seq data processing, quality control, and analysis
  - Identification of robust marker genes
  - Multi-color immunohistochemistry
  - Immunoelectron microscopy
  - Single-cell regulatory network inference and clustering (SCENIC) analysis
  - GO and KEGG analysis
  - Gene sets scoring
  - Weighted gene co-expression network (WGCNA) analysis
  - Pseudotime analysis
  - RNA velocity
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Statistical analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2022.105366.

ACKNOWLEDGMENTS

We thank all staffs from the CapitalBio Technology in Beijing for the cDNA library preparation. We are also grateful to the authors of R packages we used. This study was supported by the Beijing Natural Science Foundation (No. Z200018), Natural Science Foundation of China (No. 82171722 and 81871954), and Clinical

Medicine Plus X - Young Scholars Project, Peking University, the Fundamental Research Funds for the Central Universities (No. PKU2020LCXQ026).

AUTHOR CONTRIBUTIONS

Conceptualization, K.C., A.D., X.T.; literature search, K.C., Y.H.; data collection, K.C., J.Z.; formal analysis, K.C.; validation, K.C., J.Z., Y.H., Y.Y., and X.T.; investigation and visualization, K.C.; methodology, K.C. and X.T.; writing – original draft, K.C.; project administration, A.D. and Y.Y.; writing – review & editing, A.D. and X.T.; supervision, Y.Y. and A.D. All authors read and approved the final version of the manuscript.

DECLARATION OF INTERESTS

We declare there are no any competing financial interests in relation to this work.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: May 25, 2022
Revised: September 13, 2022
Accepted: October 12, 2022
Published: November 18, 2022

REFERENCES

1. Wang, Y.J., and Kaestner, K.H. (2019). Single-cell RNA-seq of the pancreatic islets–a promise not yet fulfilled? Cell Metabol. *29*, 539–544.

2. Bru-Tari, E., Oropeza, D., and Herrera, P.L. (2020). Cell heterogeneity and paracrine interactions in human islet function, a perspective focused in beta-cell regeneration strategies. Front. Endocrinol. *11*, 619150.

3. Choudhury, A.A., and Devi Rajeswari, V. (2021). Gestational diabetes mellitus - a metabolic and reproductive disorder. Biomed. Pharmacother. *143*, 112183.

4. 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. Metabol. *6*, 974–990.

5. Baron, M., Veres, A., Wolock, S.L., Faust, A.L., Gaujoux, R., Vetere, A., Ryu, J.H., Wagner, B.K., Shen-Orr, S.S., Klein, A.M., et al. (2016). A single-cell transcriptomic map of the human and mouse pancreas reveals inter- and intra-cell population structure. Cell Syst. *3*, 346–360.e4.

6. Enge, M., Arda, H.E., Mignardi, M., Beausang, J., Bottino, R., Kim, S.K., and Quake, S.R. (2017). Single-cell analysis of human pancreas reveals transcriptional signatures of aging and somatic mutation patterns. Cell *171*, 321–330.e14.

7. 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.

8. Muraro, M.J., Dharmadhikari, G., Grün, D., Groen, N., Dielen, T., Jansen, E., van Gurp, L., Engelse, M.A., Carlotti, F., de Koning, E.J.P., et al. (2016). A single-cell transcriptome atlas of the human pancreas. Cell Syst. *3*, 385–394.e3.

9. 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 Metabol. *24*, 593–607.

10. Wang, Y.J., Schug, J., Won, K.J., Liu, C., Naji, A., Avrahami, D., Golson, M.L., and Kaestner, K.H. (2016). Single-cell Transcriptomics of the human endocrine pancreas. Diabetes *65*, 3028–3038.

11. 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 Metabol. *24*, 608–615.

12. Chen, K., Wang, Q., Li, M., Guo, H., Liu, W., Wang, F., Tian, X., and Yang, Y. (2021). Single-cell RNA-seq reveals dynamic change in tumor microenvironment during pancreatic ductal adenocarcinoma malignant progression. EBioMedicine *66*, 103315.

13. Elyada, E., Bolisetty, M., Laise, P., Flynn, W.F., Courtois, E.T., Burkhart, R.A., Teinor, J.A., Belleau, P., Biffi, G., Lucito, M.S., et al. (2019). Cross-species single-cell analysis of pancreatic ductal adenocarcinoma reveals antigen-presenting cancer-associated fibroblasts. Cancer Discov. *9*, 1102–1123.

14. Chen, K., Liu, X., Liu, W., Wang, F., Tian, X., and Yang, Y. (2022). Development and validation of prognostic and diagnostic model for pancreatic ductal adenocarcinoma based on scRNA-seq and bulk-seq datasets. Hum. Mol. Genet. *31*, 1705–1719.

15. Chen, Z., Zhou, L., Liu, L., Hou, Y., Xiong, M., Yang, Y., Hu, J., and Chen, K. (2020). Single-cell RNA sequencing highlights the role of inflammatory cancer-associated fibroblasts in bladder urothelial carcinoma. Nat. Commun. *11*, 5077.

16. Chen, K., Wang, Q., Liu, X., Wang, F., Yang, Y., and Tian, X. (2022). Hypoxic pancreatic cancer derived exosomal miR-30b-5p promotes tumor angiogenesis by inhibiting GJA1 expression. Int. J. Biol. Sci. *18*, 1220–1237.

17. Zhang, M., Yang, H., Wan, L., Wang, Z., Wang, H., Ge, C., Liu, Y., Hao, Y., Zhang, D., Shi, G., et al. (2020). Single-cell transcriptomic architecture and intercellular crosstalk of human intrahepatic cholangiocarcinoma. J. Hepatol. *73*, 1118–1130.

18. Zheng, C., Zheng, L., Yoo, J.K., Guo, H., Zhang, Y., Guo, X., Kang, B., Hu, R., Huang, J.Y., Zhang, Q., et al. (2017). Landscape of infiltrating T cells in liver cancer revealed by single-cell sequencing. Cell *169*, 1342–1356.e16.

19. Guo, X., Zhang, Y., Zheng, L., Zheng, C., Song, J., Zhang, Q., Kang, B., Liu, Z., Jin, L., Xing, R., et al. (2018). Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. Nat. Med. *24*, 978–985.

20. Qiu, W.L., Zhang, Y.W., Feng, Y., Li, L.C., Yang, L., and Xu, C.R. (2018). Deciphering pancreatic islet beta cell and alpha cell maturation pathways and characteristic features at the single-cell level. Cell Metabol. *27*, 702.

21. Steiner, D.J., Kim, A., Miller, K., and Hara, M. (2010). Pancreatic islet plasticity: interspecies

comparison of islet architecture and composition. Islets *2*, 135–145.

22. Kim, A., Miller, K., Jo, J., Kilimnik, G., Wojcik, P., and Hara, M. (2009). Islet architecture: a comparative study. Islets *1*, 129–136.

23. Eizirik, D.L., Pasquali, L., and Cnop, M. (2020). Pancreatic beta-cells in type 1 and type 2 diabetes mellitus: different pathways to failure. Nat. Rev. Endocrinol. *16*, 349–362.

24. Yan-Do, R., Duong, E., Manning Fox, J.E., Dai, X., Suzuki, K., Khan, S., Bautista, A., Ferdaoussi, M., Lyon, J., Wu, X., et al. (2016). A glycine-insulin autocrine feedback loop enhances insulin secretion from human beta-cells and is impaired in type 2 diabetes. Diabetes *65*, 2311–2321.

25. Momose, K., Nunomiya, S., Nakata, M., Yada, T., Kikuchi, M., and Yashiro, T. (2006). Immunohistochemical and electron-microscopic observation of beta-cells in pancreatic islets of spontaneously diabetic Goto-Kakizaki rats. Med. Mol. Morphol. *39*, 146–153.

26. de Boer, P., and Giepmans, B.N. (2021). State-of-the-art microscopy to understand islets of Langerhans: what to expect next? Immunol. Cell Biol. *99*, 509–520.

27. Simicevic, J., and Deplancke, B. (2017). Transcription factor proteomics-Tools, applications, and challenges. Proteomics *17*, 1600317.

28. Dai, C., Brissova, M., Hang, Y., Thompson, C., Poffenberger, G., Shostak, A., Chen, Z., Stein, R., and Powers, A.C. (2012). Islet-enriched gene expression and glucose-induced insulin secretion in human and mouse islets. Diabetologia *55*, 707–718.

29. Jensen, M.V., Joseph, J.W., Ronnebaum, S.M., Burgess, S.C., Sherry, A.D., and Newgard, C.B. (2008). Metabolic cycling in control of glucose-stimulated insulin secretion. Am. J. Physiol. Endocrinol. Metab. *295*, E1287–E1297.

30. Wang, Z., York, N.W., Nichols, C.G., and Remedi, M.S. (2014). Pancreatic beta cell dedifferentiation in diabetes and redifferentiation following insulin therapy. Cell Metabol. *19*, 872–882.

31. Weir, G.C., Aguayo-Mazzucato, C., and Bonner-Weir, S. (2013). beta-cell dedifferentiation in diabetes is important, but what is it? Islets *5*, 233–237.

32. Sun, J., Cui, J., He, Q., Chen, Z., Arvan, P., and Liu, M. (2015). Proinsulin misfolding and endoplasmic reticulum stress during the development and progression of diabetes. Mol. Aspects Med. *42*, 105–118.

33. Zhang, I.X., Raghavan, M., and Satin, L.S. (2020). The endoplasmic reticulum and calcium homeostasis in pancreatic beta cells. Endocrinology *161*, bqz028.

34. Bensellam, M., Duvillié, B., Rybachuk, G., Laybutt, D.R., Magnan, C., Guiot, Y., Pouysségur, J., and Jonas, J.C. (2012). Glucose-induced O(2) consumption activates hypoxia inducible factors 1 and 2 in rat insulin-secreting pancreatic beta-cells. PLoS One *7*, e29807.

35. Zheng, X., Zheng, X., Wang, X., Ma, Z., Gupta Sunkari, V., Botusan, I., Takeda, T., Björklund, A., Inoue, M., Catrina, S.B., et al. (2012). Acute hypoxia induces apoptosis of pancreatic beta-cell by activation of the unfolded protein response and upregulation of CHOP. Cell Death Dis. *3*, e322.

36. Sato, Y., Endo, H., Okuyama, H., Takeda, T., Iwahashi, H., Imagawa, A., Yamagata, K., Shimomura, I., and Inoue, M. (2011). Cellular hypoxia of pancreatic beta-cells due to high levels of oxygen consumption for insulin secretion in vitro. J. Biol. Chem. *286*, 12524–12532.

37. Li, X., Zhang, L., Meshinchi, S., Dias-Leme, C., Raffin, D., Johnson, J.D., Treutelaar, M.K., and Burant, C.F. (2006). Islet microvasculature in islet hyperplasia and failure in a model of type 2 diabetes. Diabetes *55*, 2965–2973.

38. Blum, B., Hrvatin, S., Schuetz, C., Bonal, C., Rezania, A., and Melton, D.A. (2012). Functional beta-cell maturation is marked by an increased glucose threshold and by expression of urocortin 3. Nat. Biotechnol. *30*, 261–264.

39. Artner, I., Le Lay, J., Hang, Y., Elghazi, L., Schisler, J.C., Henderson, E., Sosa-Pineda, B., and Stein, R. (2006). MafB: an activator of the glucagon gene expressed in developing islet alpha- and beta-cells. Diabetes *55*, 297–304.

40. Young, M.D., and Behjati, S. (2020). SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. Gigascience *9*, giaa151.

41. Yang, S., Corbett, S.E., Koga, Y., Wang, Z., Johnson, W.E., Yajima, M., and Campbell, J.D. (2020). Decontamination of ambient RNA in single-cell RNA-seq with DecontX. Genome Biol. *21*, 57.

42. Huang, H.H., Harrington, S., and Stehno-Bittel, L. (2018). The flaws and future of islet volume measurements. Cell Transplant. *27*, 1017–1026.

43. Cottle, L., Gan, W.J., Gilroy, I., Samra, J.S., Gill, A.J., Loudovaris, T., Thomas, H.E., Hawthorne, W.J., Kebede, M.A., and Thorn, P. (2021). Structural and functional polarisation of human pancreatic beta cells in islets from organ donors with and without type 2 diabetes. Diabetologia *64*, 618–629.

44. Quesada, I., Tudurí, E., Ripoll, C., and Nadal, A. (2008). Physiology of the pancreatic alpha-cell and glucagon secretion, role in glucose homeostasis and diabetes. J. Endocrinol. *199*, 5–19.

45. Leibiger, I.B., Caicedo, A., and Berggren, P.O. (2012). Non-invasive in vivo imaging of pancreatic beta-cell function and survival - a perspective. Acta Physiol. *204*, 178–185.

46. Malaisse, W.J., Damhaut, P., Ladriere, L., and Goldman, S. (2000). Fate of 2-deoxy-2-[18F] fluoro-D-glucose in hyperglycemic rats. Int. J. Mol. Med. *6*, 549–552.

47. Malaisse, W.J., Damhaut, P., Malaisse-Lagae, F., Ladriere, L., Olivares, E., and Goldman, S. (2000). Fate of 2-deoxy-2-[18F]fluoro-D-glucose in control and diabetic rats. Int. J. Mol. Med. *5*, 525–532.

48. Goland, R., Freeby, M., Parsey, R., Saisho, Y., Kumar, D., Simpson, N., Hirsch, J., Prince, M., Maffei, A., Mann, J.J., et al. (2009). 11C-dihydrotetrabenazine PET of the pancreas in subjects with long-standing type 1 diabetes and in healthy controls. J. Nucl. Med. *50*, 382–389.

49. Normandin, M.D., Petersen, K.F., Ding, Y.S., Lin, S.F., Naik, S., Fowles, K., Skovronsky, D.M., Herold, K.C., McCarthy, T.J., Calle, R.A., et al. (2012). In vivo imaging of endogenous pancreatic beta-cell mass in healthy and type 1 diabetic subjects using 18F-fluoropropyl-dihydrotetrabenazine and PET. J. Nucl. Med. *53*, 908–916.

50. Luo, Y., Pan, Q., Yao, S., Yu, M., Wu, W., Xue, H., Kiesewetter, D.O., Zhu, Z., Li, F., Zhao, Y., et al. (2016). Glucagon-like peptide-1 receptor PET/CT with 68Ga-NOTA-Exendin-4 for detecting localized insulinoma: a prospective cohort study. J. Nucl. Med. *57*, 715–720.

51. Kang, N.Y., Lee, J.Y., Lee, S.H., Song, I.H., Hwang, Y.H., Kim, M.J., Phue, W.H., Agrawalla, B.K., Wan, S.Y.D., Lalic, J., et al. (2020). Multimodal imaging probe development for pancreatic beta cells: from fluorescence to PET. J. Am. Chem. Soc. *142*, 3430–3439.

52. Leturque, A., Brot-Laroche, E., and Le Gall, M. (2009). GLUT2 mutations, translocation, and receptor function in diet sugar managing. Am. J. Physiol. Endocrinol. Metab. *296*, E985–E992.

53. Thorens, B. (2015). GLUT2, glucose sensing and glucose homeostasis. Diabetologia *58*, 221–232.

54. Thorens, B., Weir, G.C., Leahy, J.L., Lodish, H.F., and Bonner-Weir, S. (1990). Reduced expression of the liver/beta-cell glucose transporter isoform in glucose-insensitive pancreatic beta cells of diabetic rats. Proc. Natl. Acad. Sci. USA *87*, 6492–6496.

55. Guillam, M.T., Hümmler, E., Schaerer, E., Yeh, J.I., Birnbaum, M.J., Beermann, F., Schmidt, A., Dériaz, N., Thorens, B., and Wu, J.Y. (1997). Early diabetes and abnormal postnatal pancreatic islet development in mice lacking Glut-2. Nat. Genet. *17*, 327–330.

56. Fonseca, S.G., Gromada, J., and Urano, F. (2011). Endoplasmic reticulum stress and pancreatic beta-cell death. Trends Endocrinol. Metabol. *22*, 266–274.

57. Sabatini, P.V., Speckmann, T., and Lynn, F.C. (2019). Friend and foe: beta-cell Ca(2+) signaling and the development of diabetes. Mol. Metabol. *21*, 1–12.

58. Meyerovich, K., Ortis, F., Allagnat, F., and Cardozo, A.K. (2016). Endoplasmic reticulum stress and the unfolded protein response in pancreatic islet inflammation. J. Mol. Endocrinol. *57*, R1–R17.

59. Woehlbier, U., and Hetz, C. (2011). Modulating stress responses by the UPRosome: a matter of life and death. Trends Biochem. Sci. *36*, 329–337.

60. Tuch, B.E., Szymanska, B., Yao, M., Tabiin, M.T., Gross, D.J., Holman, S., Swan, M.A., Humphrey, R.K.B., Marshall, G.M., and Simpson, A.M. (2003). Function of a genetically modified human liver cell line that stores, processes and secretes insulin. Gene Ther. *10*, 490–503.

61. Aibar, S., González-Blas, C.B., Moerman, T., Huynh-Thu, V.A., Imrichova, H., Hulselmans, G., Rambow, F., Marine, J.C., Geurts, P., Aerts, J., et al. (2017). SCENIC: single-cell regulatory network inference and clustering. Nat. Methods *14*, 1083–1086.

62. 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*, 1888–1902.e21.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal to GLRA1	Abcam	Cat#ab228541
Rabbit monoclonal to ATF3	Abcam	Cat#ab254268; RRID: AB_2910214
Rabbit monoclonal to KRT19	Abcam	Cat#ab52625; RRID: AB_2281020
Rabbit monoclonal to PRSS1	Abcam	Cat#ab200997
Rabbit monoclonal to INS	Abcam	Cat#ab181547; RRID: AB_2716761
Rabbit monoclonal to SST	Abcam	Cat#ab111912; RRID: AB_10903864
Rabbit monoclonal to PDX1	Abcam	Cat#ab134150; RRID: AB_2631338
Rabbit monoclonal to ADCYAP1	Abcam	Cat#ab181205
Deposited data		
ScRNA-seq dataset for the control and type 2 diabetes mice	This paper	GSE200531
Baron et al.[5]	GEO	GSE84133
Enge et al.[6]	GEO	GSE81547
Lawlor et al.[7]	GEO	GSE86473
Muraro et al.[8]	GEO	GSE85241
Segerstolpe et al.[9]	ArrayExpress	E-MTAB-5061
Wang et al.[10]	GEO	GSE83139
Xin et al.[11]	GEO	GSE81608
Qiu et al.[20]	GEO	GSE87375
Experimental models: Organisms/strains		
C57BLKS-Leprdb/db (BKS_DB)	GemPharmatech. Co,.LTD company	N/A
wt/wt mice (BKS_WT)	GemPharmatech. Co,.LTD company	N/A
Software and algorithms		
Cell Ranger	10x Genomics	https://support.10xgenomics.com/single-cell-gene-expression/software/downloads/latest
Seurat	Stuart et al.[62]	https://satijalab.org/seurat/
Harmony	N/A	https://github.com/immunogenomics/harmony
SCENIC	Aibar et al.[61]	N/A

RESOURCE AVAILABILITY

Lead contact

Further information and requests for datasets and reagents should be directed to and will be fulfilled by the lead contact, Kai Chen (Drchenkai@pku.edu.cn).

Materials availability

This study did not generate new unique reagents.

Data and code availability

Single-cell RNA-seq data have been deposited at GEO and are publicly available. Accession numbers are listed in the key resources table.

Any additional information required to reanalyze the data reported in this study will be provided upon request to the lead contact.

The accession number for the scRNA-seq data reported in this paper is GSE200531.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human subjects

Human primary pancreatic ductal adenocarcinoma (PDAC) and matched adjacent normal pancreatic resection specimens were retrieved from the Department of General Surgery of Peking University First Hospital. Two male and two female patients aged over 18 did not receive any treatments before specimen collections for IHC. Gender had no effect on this study. This study was approved by Ethics Committee of Peking University First Hospital (Approval No. 2019-147) and was conducted in accordance with ethical guidelines (Declaration of Helsinki). Written informed consent was obtained from all participants.

Animals

Male homozygous C57BLKS-Leprdb/db (BKS_DB) and wt/wt mice (BKS_WT) were purchased from the GemPharmatech. Co,.LTD company and housed in a specific pathogen-free (SPF) environment with free access to food and water. All animal procedures and studies were conducted in accordance with the Medicine Animal Care and Use Committee at Peking University First Hospital. Genotyping assay was conducted to verify Lepr knockout using the Quick Genotyping Assay Kit for Mouse Tail according to the manufacturer's protocol (Beyotime, Cat. no. D7283M). The following primers were designed: Lepr-F1: CCAAGAAC CAAGTGTTTGCAGTCAC; Lepr-R1: GTTCTCACCAGAGGTCCCTAAACTC; Lepr-F2: TTTTACCAAG CATGCAGAATCAGTG; Lepr-R2: ACCCTTGCTCTTCATCAGTTTCCAC. Body weights and blood glucose levels of mice were monitored regularly.

METHOD DETAILS

scRNA-seq datasets from open databases

The pancreatic islet scRNA-seq datasets were downloaded from open databases, including Gene Expression Omnibus (GEO) and ArrayExpress (EBI). Accession numbers were summarized in Figure S1A. Pancreatic islets were isolated from human and mouse donors of different ages, developmental stages, and diabetes types. Raw gene-cell matrix from each dataset was imported into R software and conducted the initial quality control to filter out low-quality cells. The general characteristics of specimens were also collected according to information provided by research articles.

Hematoxylin-eosin (HE) and IHC stating

HE and IHC staining was performed on freshly prepared 4% paraformaldehyde fixed paraffin-embedded tissue sections. The 5-μm sections were cut and mounted on slides. Slides were deparaffinized in xylene, rehydrated in decreasing concentration of ethanol, then washed with PBS twice. For HE staining, slides were subjected to staining using the Hematoxylin and Eosin Staining Kit (Beyotime, Cat. no. C0105M). For IHC staining, slides were added 3% hydrogen peroxide to remove endogenous peroxidase activity. Antigen retrieval was performed in 10 mmol/L citrate buffer, in a pressure cooker for 10 minutes. Tissues were blocked with 10% goat serum at room temperature for 60 minutes and incubated with primary antibody for GLRA1 (Abcam, Cat. no. ab228541, 1:100) at 4°C overnight in a humidified chamber. Finally, each slide was treated with HRP conjugated secondary antibody at room temperature for 60 minutes, and fresh diaminobenzidine (DAB) (Beyotime, Cat. no. P0203) for 5–10 minutes. The cell nucleus was counterstained with Mayer's hematoxylin solution (Biodee, Cat. no. 131-09665). The pictures of pancreatic islet were obtained for each section.

Pancreatic islet isolation

Both BKS_DB (n = 3) and BKS_WT (n = 3) male mice were euthanized with cervical dislocation at 9 weeks of age. Mice were fed a regular diet (ambient fed) before islets isolation. Tape the limbs of the mouse in supine position, spray the skin of the abdominal area with 75% ethanol. Make a vertical incision to fully expose organs in abdominal cavity. Find the bile duct and ampulla of Vater, which is located at the duodenal papilla, under a dissection microscope. Carefully insert the syringe with 10 mL of the collagenase Type V solution (Sigma-Aldrich, Cat. no. C9263, 0.5 mg/mL) into the ampulla of Vater, and steadily inject

2–3 mL of the collagenase solution when the common bile duct was clamped. Stop injection until the head, neck, body and tail region of the pancreas are all fully inflated. Carefully dissect out the inflated pancreas and place it in a 50 mL tube containing 2 mL of ice-cold collagenase solution. After completing pancreas isolation of six mice, place the tube in a rack in 37°C water bath for about 15 minutes. Shake gently the tube every five minutes. Take out tube from water bath and shake vigorously for 30 seconds. Then, place the tubes on ice and add 45 mL of ice-cold Hank's Balanced Salt Solution (HBSS) supplemented with 5% fetal bovine serum (Gibco, Cat. no. 10099-141) to stop the enzymatic digestion. Filter out undigested tissues with the mesh sieve. Centrifuge the tube in a swinging-bucket centrifuge at 300g for 2 minutes. Discard the supernatant and rinse twice with HBSS. After the last centrifugation remove all HBSS and add 10 mL room temperature Histopaque-1077 solution (Sigma-Aldrich, Cat. no. 10771). Vortex briefly and slowly add another 10 mL Histopaque-1077 and 20 mL room temperature RPMI-1640 (Gibco, Cat. no. C11875500BT). Centrifuge the tube in a swinging-bucket centrifuge at 900g with no brake for 20 minutes. Pipette out the layer of islets (5–10 mL) into a new 50 mL collection tube. Add 20 mL of ice cold HBSS and repeat washing and centrifuging three times. Pipette out the complete RPMI-1640 containing pancreatic islets into a Petri dish (NEST Biotechnology). Incubate the islets in a sterile incubator at 37°C with 5% CO2 overnight.

Pancreatic islet dissociation and cell purification

Transfer cultured islets overnight to a new 15 mL tube and centrifuge at 200g for 3 minutes. Discard the supernatant and rinse twice with PBS. After the last centrifugation, remove all PBS and add 3 mL room temperature Accutase solution (Gibco, Cat. no. A1110501). Incubate the tube in 37°C water bath for 20 minutes. Pipette up and down every 2 minutes. Terminate the digestion by adding 5 mL complete RPMI-1640. Centrifuge at 200g for 3 minutes and discard the supernatant. The pellet was resuspended with 2 mL RPMI-1640 and filtered with 35 um cell strainer (Falcon, Cat. no. 352235).

Cell capture and cDNA library preparation

The concentration of single cell suspension was determined using the Count Star instrument and adjusted to 1000 cells/ul. Using single cell 3′ Library and Gel Bead Kit v3.1 (10x Genomics, Cat. No. 1000121) and Chromium Next GEM Chip G Single Cell Kit (10x Genomics, Cat. No. 1000120), the cell suspension was loaded onto the Chromium single cell controller (10x Genomics) to generate single-cell gel beads in the emulsion according to the manufacturer's protocol. In brief, single cells were suspended in PBS containing 0.04% BSA. About 6,000 cells were added to each channel, and the target cell will be recovered (about 3000 cells). Captured cells were lysed and the released RNA were barcoded through reverse transcription in individual GEMs. Reverse transcription was performed in 200 ul tubes (NEST Biotechnology, Cat. No. 401001) on a S1000TM Touch Thermal Cycler (Bio Rad) at 53°C for 45 minutes, followed by 85°C for 5 min, and hold at 4°C. The cDNA was generated and then amplified, and quality assessed using the Agilent 4200. The cDNA libraries were finally sequenced using an Illumina Novaseq6000 sequencer with a sequencing depth of at least 100,000 reads per cell with paired-end 150 bp (PE150) reading strategy.

scRNA-seq data processing, quality control, and analysis

The Cell Ranger software was downloaded from 10x Genomics website (https://support.10xgenomics.com/single-cell-gene-expression/software/downloads/latest). Alignment (mm10 reference for mouse samples), filtering, barcode counting, and UMI counting were performed with cell ranger count module to generate gene-cell matrix (v3.0.1, 10x Genomics). The raw gene-cell matrixes from BKS_DB and BKS_WT specimens were merged and loaded into R package Seurat (v3.2.3). Then, gene-cell matrixes were filtered out low-quality cells (<200 genes/cell, > 15% mitochondria genes, < 1000 transcripts/cell) and genes (<10 cells/gene). Gene expression levels were normalized (LogNormalize). Total 2000 highly variable genes were acquired and used to conduct PCA reduction dimension. Uniform manifold approximation and projection (u-MAP) was performed. R package DoubletFinder (v2.0.3) was used to identify doublets using the same PCs in PCA above, assuming around 5% doublet formation rate to the loaded cells for each specimen in a droplet channel. R package Harmony (https://github.com/immunogenomics/harmony) was used to the integration of cells from various specimens. Signature genes of each cluster matched with marker genes of known cell populations reported in previous literatures and CellMarker database. Violin plot showed the expression levels of selected signature genes in each cluster.

Identification of robust marker genes

After identifying endocrine and exocrine cell types, we obtained cell-specific marker genes for each scRNA-seq dataset using the FindMakers function in Seurat. Wilcoxon test was used to find differential expression marker genes. Significance was considered as an average natural logarithm (fold change) of at least 0.25 and an adjusted p-value lower than 0.05. Common marker genes across all datasets were obtained. The upset plots were used to show the intersection of marker genes.

Multi-color immunohistochemistry

To explore the expression and distribution of ATF3, KRT19, and PRSS1 in exocrine cells, and INS, SST, GLRA1, PDX1, and ADCYAP in pancreatic islet cells, formalin-fixed paraffin-embedded (FFPE) sections from PDAC and adjacent normal tissues were subjected to multiple IHC using PANO Multiplex IHC kit (Panovue, Cat. No. 0079100100). Briefly, sections were deparaffinized in fresh xylene for 10 minutes three times, rehydrated in decreasing concentration of ethanol (100%–95% – 70%), then washed 3 times with PBS. The antigen retrieval was performed with microwave heating method and cooled down for at least 10–15 minutes in the ice-water bath. Blocking was performed with blocking solution (Panovue, Cat. No. 0018001120) for 10 minutes at room temperature, followed by incubation with primary antibody for 30 minutes, secondary antibody for 10 minutes, and TSA Opal fluorophores for 10 minutes. Repeated antigen retrieval, blocking, primary and antibody incubation, TSA Opal fluorophores staining for each marker. Finally, all sections were counterstained with DAPI (Sigma-Aldrich, Cat. No. D9542) for 5 minutes and mounted. Sections were scanned using the panoVIEW VS200 (china). Images capture was conducted using the HALO software. The following primary antibodies were used. Panel 1: ATF3 (Abcam, Cat. no. ab254268, 1:500), KRT19 (Abcam, Cat. no. ab52625, 1:500), PRSS1 (Abcam, Cat. no. ab200997, 1:500); Panel 2: INS (Abcam, Cat. no. ab181547, 1:800), SST (Abcam, Cat. no. ab111912, 1:100), GLRA1 (Abcam, Cat. no. ab228541, 1:200), PDX1 (Abcam, Cat. no. ab134150, 1:200), ADCYAP1 (Abcam, Cat. no. ab181205, 1:500).

Immunoelectron microscopy

Immunoelectron microscopy was performed.[60] In brief, fresh pancreatic resection specimens were fixed with 2% glutaraldehyde and 1% paraformaldehyde. Tissue was dehydrated in solutions of graded alcohol and embedded in Durcupan. Ultra-thin sections were cut on a Reichert Ultracut-S. Sections were incubated with the primary antibody, GLRA1 (Abcam, Cat. no. ab228541) and ADCYAP1 (Abcam, Cat. no. ab181205) at 37°C for 60 minutes. Washed the sections with PBS three times, then the sections were transferred to the secondary antibody solution (IgG-colloidal gold) and incubated at 37°C for 60 minutes. Sections were taken out of the solution and were rinsed in PBS three times. Ultra-thin sections were counterstained with lead citrate and observed under a JEM-1010 transmission electron microscope.

Single-cell regulatory network inference and clustering (SCENIC) analysis

We conducted SCENIC analysis using R package SCENIC (v1.1.3).[61] In short, we randomly picked out 200–300 cells from each cell subpopulation to construct new gene-cell matrix from the Seurat object. Regions for TF searching were restricted to 10 k distance centered the transcriptional start site (TSS) or 500 bp upstream of the TSSs. Transcription factor binding motifs (TFBS) overrepresented on a gene list and networks inferring were done using R package Rcis-Target (v1.6.0) and GENIE3 (v1.8.0), with the 20-thousand motifs database. The cluster-specific TFs were defined as the top 10 or 15 highly enriched TFs according to a decrease in fold change compared with all the other cell clusters using a Wilcoxon rank-sum test. Visualization of TFs network was performed using the heatmap.

GO and KEGG analysis

FindMarkers function in Seurat was applied to identify differential expression genes (DEGs) between two beta subclusters (logfc.threshold = 0.5, q-value < 0.05). Volcano plot was used to visualize the result. The online tool g:Profiler (https://biit.cs.ut.ee/gprofiler/gost) was used to conduct GO and KEGG analysis for top 30–100 DEGs.

Gene sets scoring

Gene sets from UPR, hypoxia, glycolysis, and fatty acid metabolism pathways were downloaded from the MSigDB (http://www.gsea-msigdb.org/gsea/) and Reactome databases (http://www.reactome.org). The pathways-related scores were calculated using the AddModuleScore function in Seurat.

Weighted gene co-expression network (WGCNA) analysis

We randomly picked out 200–300 cells from each beta subcluster to construct a new gene-cell matrix for WGCNA analysis using R package WGCNA (v1.69). A signed network was constructed using any genes that was expressed at UMI value of 1 or higher at least 5 samples. After constructing the adjacency matrix and selecting the optimal soft-power (n = 3), we calculated the topological overlaps matrix (TOM). Genes with high similar co-expression relationship were grouped together using average linkage hierarchical clustering upon TOM. Through the Dynamic Tree Cut method, we identified the cluster-specific gene modules. Subsequently, we proved blue module was significantly associated with subcluster 0 (type 2 diabetes mice) by the scatter plot of Gene Significance (GS) vs Module Membership (MM). Finally, the hub genes representing subcluster 0 in blue module were exported to perform visualization by Cytoscape software (v 3.6.0).

Pseudotime analysis

R package Monocle 2 (v2.18.0) was applied to conduct cellular trajectory inference with the assumption that one-dimensional 'time' can describe the high-dimensional expression values, so called pseudotime analysis of single cells. General pipeline was implemented. Gene-cell matrix of the beta subclusters was loaded into R environment. And we presented cell trajectory and position with tree structure in two-dimension space after log normalization and DDDTree reduction dimension. Then we set the pattern of each cell in the plot according to the expression levels of selected signatures and pathways-related scores.

RNA velocity

RNA velocity analysis was performed to further validate Beta cell developmet trajectory using the velocyto (v0.17.17) and scvelo (v0.0.4) python packages. In brief, the spliced and unspliced matrixes were counted and bam files were converted to loom files. Convert data from Seurat to anndata in python. Next, RNA velocity was calculated using the steady-state model (stochastic option). The velocity graph was used to embed RNA velocities into the u-MAP plots. The downstream visualization was completed with the scv.pl.velocity_embedding_grid and scv.pl.heatmap functions.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis

All statistical analyses were conducted using the SPSS (v22.0) and R software. For continuous variable, the independent-samples t test and Mann-Whitney U test was performed to compare means between two groups. For categorical variable, the chi-square test or rank sum test was performed. Coefficients of Spearman's rank correlation or Pearson's correlation were calculated to describe the correlation of two variables. Statistical significance was defined as *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.