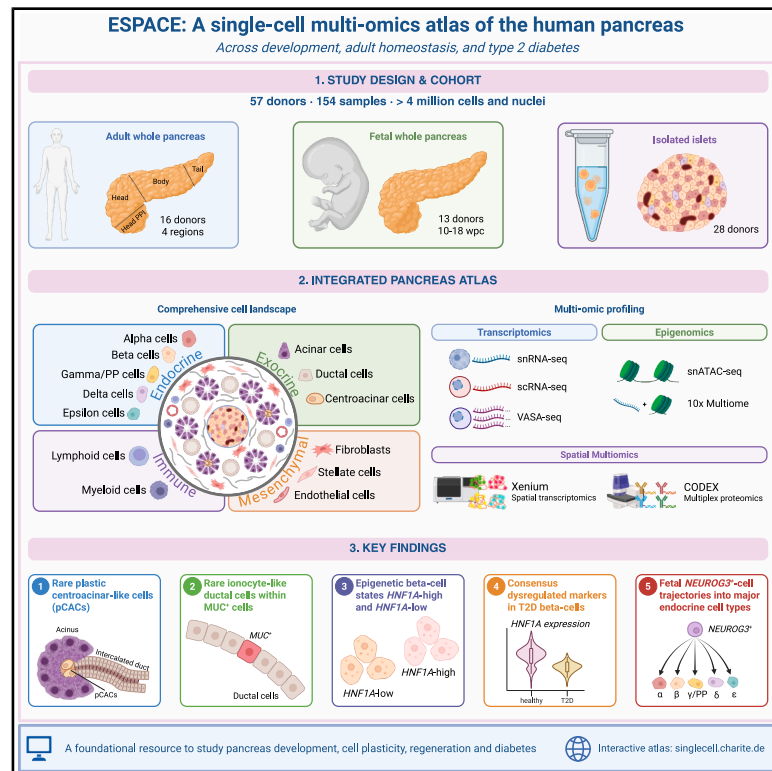


Cell Metabolism

A single-cell multi-omic atlas of the human pancreas reveals potential for cellular plasticity across development, health, and disease

Graphical abstract



Authors

Elisabetta Mereu, Diego Balboa, Johannes Liebig, ..., Wilko Weichert, Christian Conrad, Roland Eils

Correspondence

emereu@carrerasresearch.org (E.M.), roland.eils@bih-charite.de (R.E.)

In brief

Mereu et al. present ESPACE, a single-cell multi-omic atlas of the human pancreas spanning fetal development, adulthood, and type 2 diabetes. Combining transcriptomics, chromatin accessibility, and spatial imaging, they uncover a rare population of plastic centroacinar cells and hidden diversity among hormone-producing islet cells, informing pancreas regeneration and disease.

Highlights

- Multi-omic atlas profiles >4 million human pancreas cells across 57 donors
- Rare plastic centroacinar-like cells (pCACs) share fetal multipotency programs
- Beta cells show HNF1A-high/low chromatin states altered in type 2 diabetes
- Spatial transcriptomics profiling reveals endocrine heterogeneity outside islets



Resource

A single-cell multi-omic atlas of the human pancreas reveals potential for cellular plasticity across development, health, and disease

Elisabetta Mereu,^{1,28,*} Diego Balboa,^{2,3,4,28} Johannes Liebig,^{5,6,28} Aitor Gonzalez-Herrero,^{1,7,28} Anna Martinez Casals,^{8,28} Mariya Mardamshina,^{8,29} Fanny Mollandin,^{2,4,29} Felix Schick Tanz,^{9,29} Alexander Sudy,^{5,6,29} Luca Tosti,^{5,29} Maarten van Agen,^{10,29} Valerie Vandenbempt,^{3,29} Dana Avrahami,¹¹ Frederic Ballllosera Navarro,⁸ Edgar Bernardo,^{2,4,12,13} Frida Björklund,⁸ Robert Lorenz Chua,⁵ Marten Engelse,¹⁰ Javier García-Hurtado,^{2,4} Nathalie Groen,^{10,14} Maaïke Hanegraaf,¹⁰ Pablo Iañez,¹ Katharina Jechow,⁵ Björn Konukiewicz,⁹ Christian Lawrenz,¹⁵

(Author list continued on next page)

¹Josep Carreras Leukemia Research Institute, Barcelona, Spain

²Centre for Genomic Regulation, the Barcelona Institute of Science and Technology, Dr. Aiguader 88, 08003 Barcelona, Spain

³Stem Cells and Metabolism Research Program, Faculty of Medicine, University of Helsinki, Haartmaninkatu 8, 00290 Helsinki, Finland

⁴Centro de Investigación Biomédica en Red Diabetes y Enfermedades Metabólicas Asociadas (CIBERDEM), Barcelona, Spain

⁵Berlin Institute of Health at Charité - Universitätsmedizin Berlin, Digital Health Center, Charitéplatz 1, 10117 Berlin, Germany

⁶Freie Universität Berlin, Berlin, Germany

⁷Universitat de Barcelona, Barcelona, Spain

⁸Department of Bioengineering, Stanford University, Stanford, CA, USA

⁹Institute of Pathology, School of Medicine and Health, Technical University of Munich, Munich, Germany

¹⁰Department of Internal Medicine, Leiden University Medical Center (LUMC), Leiden, the Netherlands

¹¹Department of Developmental Biology and Cancer Research, The Hebrew University-Hadassah Medical School, Jerusalem, Israel

¹²Organ Crosstalk in Metabolic Diseases Group, Tumour Biology Programme, Centro Nacional Investigaciones Oncológicas (CNIO), Madrid, Spain

¹³Centro Nacional de Investigaciones Cardiovasculares Carlos III (CNIC), Madrid, Spain

¹⁴Hubrecht Institute, Royal Netherlands Academy of Arts and Sciences, University Medical Center Utrecht, Utrecht, the Netherlands

¹⁵Steinbeis Transfer GmbH, Center Health Data, Stuttgart, Germany

¹⁶Comparative Experimental Pathology, Institute of Pathology, School of Medicine and Health, Technical University of Munich, Munich, Germany

(Affiliations continued on next page)

SUMMARY

The pancreas plays a central role in major human diseases, yet our understanding of its cellular diversity and plasticity remains incomplete. Here, we present a single-cell multiomics atlas of the human pancreas, profiling over four million cells and nuclei from 57 donors across fetal development, adult homeostasis, and type 2 diabetes (T2D). Integrating single-cell RNA sequencing (scRNA-seq)/single-nucleus RNA sequencing (snRNA-seq), snATAC-seq, VASA-seq, spatial transcriptomics (Xenium), and multiplexed proteomics (CODEX), we resolve gene expression, chromatin accessibility, and spatial organization at high resolution. We identify transcriptionally plastic centroacinar-like cells (pCACs) in adults with fetal-like features, delineate endocrine and exocrine lineage trajectories during development, and define HNF1A-defined beta cell epigenetic states. In T2D, we observe shifts in beta cell subtypes and altered regulatory programs. Glucose perturbation of healthy islets reveals cell-type-specific adaptation and stress responses. This atlas provides a foundational framework to understand pancreas biology and the role of cellular plasticity in regeneration and disease.

INTRODUCTION

The pancreas is a vital organ with both exocrine and endocrine functions, essential for proper digestive processes and glucose regulation in the body. The stability of its functions is sustained

by limited cell turnover, primarily through self-duplication of pancreatic cells.^{1–4} The identity of pancreatic cells is established during embryonic development and maturation and appears to be conserved between mice and humans.^{5–7} Dynamic gene expression allows the pancreas to adapt to changes in its



Nadja Lewandowski-Hoppe,¹⁶ Domenica Marchese,¹⁷ Mauro J. Muraro,¹⁸ Silvia Pellegrini,¹⁹ Valeria Sordi,¹⁹ Ulrike Taron,⁵ Foo Wei Ten,^{5,6} Timo Trefzer,⁵ Sven Twardziok,⁵ Johannes Wirth,¹⁶ Françoise Carlotti,^{10,30} Eelco de Koning,^{10,30} Jorge Ferrer,^{2,4,20,30} Benjamin Glaser,^{21,30} Holger Heyn,^{7,17,22,30} Emma Lundberg,^{8,30} Lorenzo Piemonti,^{18,23,30} Katja Steiger,^{16,24,30} Alexander van Oudenaarden,^{14,25,26,30} Wilko Weichert,^{9,24,30} Christian Conrad,^{5,31} and Roland Eils^{5,27,31,32,*}

¹⁷Centro Nacional de Análisis Genómico (CNAG), Barcelona, Spain

¹⁸Single Cell Discoveries, Utrecht, the Netherlands

¹⁹Diabetes Research Institute, IRCCS Ospedale San Raffaele, Milan, Italy

²⁰Section of Genetics and Genomics, Department of Metabolism, Digestion, and Reproduction, Imperial College London, London, UK

²¹Department of Endocrinology and Metabolism, Hadassah Medical Center, Faculty of Medicine, Hebrew University of Jerusalem, Jerusalem, Israel

²²ICREA, Barcelona, Spain

²³Università Vita-Salute San Raffaele, Milan, Italy

²⁴German Cancer Consortium (DKTK), partner site Munich, a partnership between DKFZ and TUM University Hospital, Munich, Germany

²⁵Oncode Institute, Utrecht, the Netherlands

²⁶University Medical Center Utrecht, Utrecht, the Netherlands

²⁷Intelligent Medicine Institute, Fudan University, Shanghai, China

²⁸These authors contributed equally

²⁹These authors contributed equally

³⁰Senior author

³¹Senior author

³²Lead contact

*Correspondence: emereu@carrerasresearch.org (E.M.), roland.eils@bih-charite.de (R.E.)

<https://doi.org/10.1016/j.cmet.2026.07.023>

functions and preserve homeostasis throughout the lifespan. Studies have shown that adult pancreatic acinar cells can, under certain conditions, undergo epigenomic remodeling and adopt alternative identities.^{8,9} Such plasticity might predispose to neoplastic evolution in the presence of oncogenic drivers.¹⁰ However, the extent and physiological relevance of such processes remain incompletely understood. In this regard, efforts to build a more complete human cell atlas of the pancreas are crucial to delineate the cellular and molecular landscape of the developing, homeostatic, and diseased human pancreas. Single-cell genomics has been applied to investigate the cellular heterogeneity of the human pancreas, particularly within the islets of Langerhans,^{11–14} but these studies have faced important limitations in accurately defining the full spectrum of heterogeneous cell types. Technical challenges include (1) the high enzymatic activity of the pancreas, which makes the isolation of exocrine cells difficult; (2) variability in experimental techniques, resulting in a lack of marker consensus between individual datasets¹⁵; and (3) lack of adjacent cytoarchitecture spatial information. In addition to these limitations, previous studies on the cellular heterogeneity of the human pancreas have been restricted to single-cell or single-nucleus RNA sequencing (sc/nRNA-seq), which provides valuable information on the transcriptomic identity of individual cells but does not capture their epigenetic states or spatial context. As epigenomic mechanisms are essential for maintaining cell identity,^{7,16,17} regulating plasticity, and mediating responses to stress and injury,^{10,18} integrating epigenomic profiling with transcriptomic and spatial approaches facilitates deeper insights into pancreatic cell differentiation, adaptation, and dysfunction.

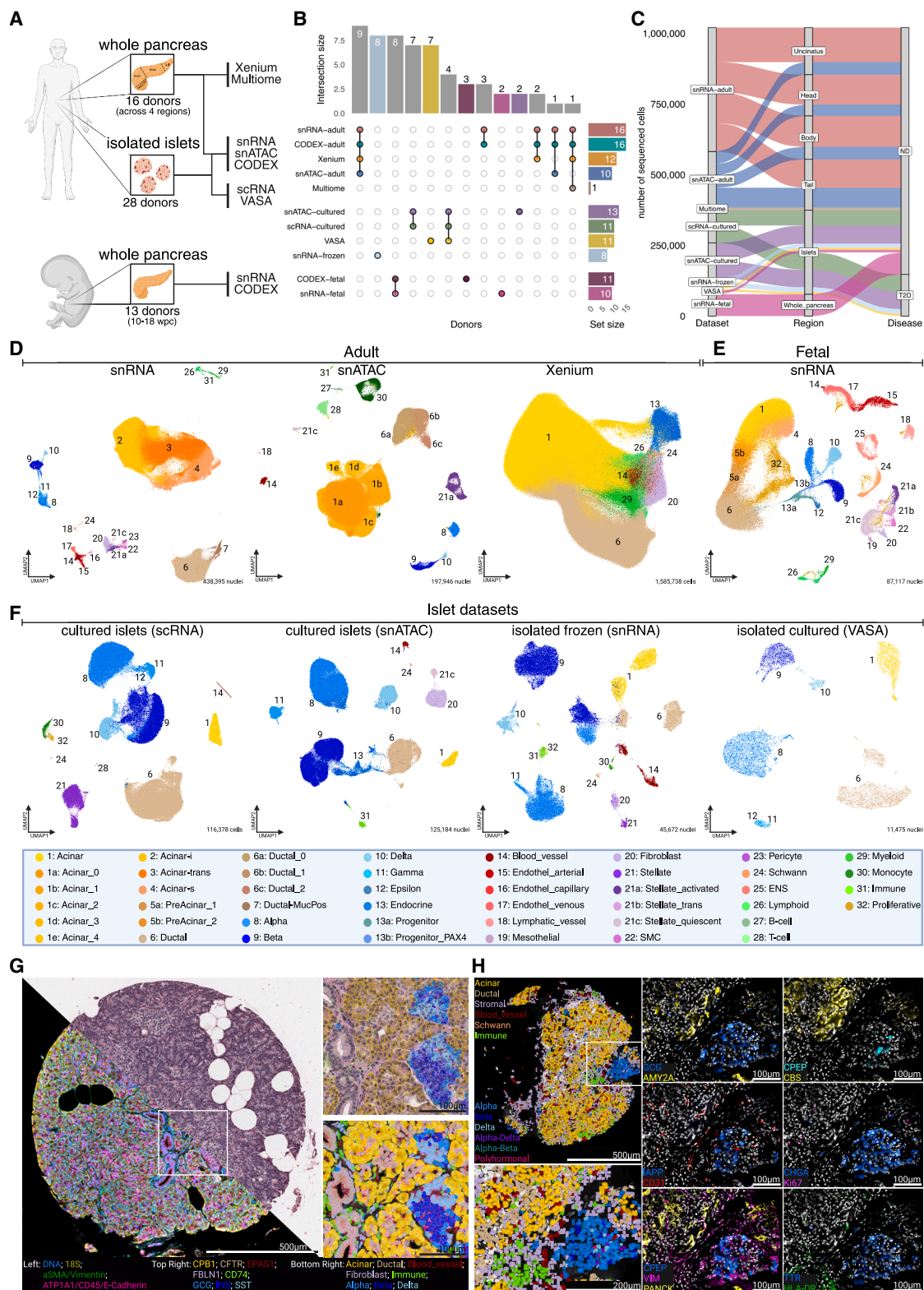
Here, we combined advanced single-cell sequencing techniques, together with spatial transcriptomics (Xenium) and proteomics (CODEX), to create a single-cell multiomics reference atlas of the human pancreas (Figure 1), including early prenatal life, adulthood, and type 2 diabetes (T2D) samples, consisting

of 57 donors. To generate this resource, we processed adult pancreatic tissue from four anatomically distinct regions: head, PPY-rich head, body, and tail. Notably, while many donors displayed histologically normal pancreatic tissue, pathologists also identified features such as fibrosis, lipomatosis, and pancreatic intraepithelial neoplasias (PanINs) in a subset of samples, presenting a view of pancreatic cellular states across physiological and early pathological contexts. Additionally, analysis of human embryonic pancreas samples from 13 to 17 weeks post-conception (wpc) for single-cell transcriptomics and from 10 to 18 wpc for proteomics allowed us to assess cell differentiation during development and comparison with adult pancreas data. The value of this multiomic resource is exemplified in the following key insights: (1) identification of rare centroacinar cells enriched for acinar-to-ductal metaplasia (ADM) and multipotent progenitor markers; (2) identification of very rare subpopulations, such as ionocyte-like cells within the MUC⁺ ductal cells; (3) identification of islet endocrine cell heterogeneity, including two epigenetic beta cell states (*HNF1A*_{high} and *HNF1A*_{low}); (4) identification of consensus dysregulated markers in T2D beta cells, showing altered expression of beta cell regulators; and (5) delineation of *NEUROG3*⁺ endocrine progenitors' differentiation trajectory into major endocrine cell types. An interactive visualization for exploration and investigation of ESPACE datasets is available at singlecell.charite.de.

RESULTS

ESPACE—The Expression and Spatial Analysis Pancreas Atlas in numbers

We integrated data from 154 samples, including adult whole pancreas (64 samples from 16 donors across four anatomical regions), isolated islets (80 samples profiled using four distinct assay types from 28 donors), and fetal pancreas (13 donors from gestational weeks 10–18). We employed ten



(legend on next page)

complementary modalities, encompassing transcriptomic (single-nucleus RNA sequencing [snRNA-seq], single-cell RNA sequencing [scRNA-seq], and VASA-seq), epigenomic (snATAC-seq, 10× Multiome), and spatial platforms (Xenium spatial transcriptomics and CODEX high-plex immunofluorescence), with overlapping donors across multiple datasets (Table S1). Together, these approaches captured over 4 million cells and nuclei, enabling robust cross-platform validation of spatial, proteomic, and transcriptional signatures in both adult and developing pancreatic tissue. To contextualize the scale and quality of ESPACE relative to existing resources, we benchmarked our datasets against previously published pancreatic single-cell atlases, including Tosti et al.¹³ and the HPAP consortium.¹⁹ This comparison demonstrates that ESPACE substantially extends prior resources in cell number, cellular diversity, transcriptome depth, and RNA class coverage across multiple protocols (Figure S1).

Molecular definition of adult human pancreatic cell types using single-nuclei transcriptomics and chromatin accessibility

To comprehensively dissect the cellular heterogeneity of the adult human pancreas, we employed snRNA-seq and snATAC-seq to profile a broad range of tissue samples collected from four pancreas regions (head, PPY-rich head, body, and tail) (Figures 1A–1C and S2A–S2D; Table S1). To ensure adequate representation of acinar and ductal cells, we relied on a custom protocol²⁰ that generates a suspension of intact nuclei from pancreas tissue using citric acid, which inhibits pancreas enzymes' autolytic activity, improving mRNA quality.

Using snRNA-seq, we analyzed 438,395 high-quality nuclei across 64 samples from 16 healthy donors (Figure 1D), identifying major pancreatic cell types using established gene markers^{13,15} (Figures S2E, S2F, and S3A; Table S2), with cell-type assignments further validated by cross-dataset comparison with Tosti et al.,¹³ showing high concordance in marker gene expression profiles (Figure S3B). Next, we conducted a comprehensive analysis of the pancreatic cell compartments, including acinar (Figure S3C), ductal (Figure S3D), and endocrine cells (Figure S3E), uncovering 21 major cell types and 34 additional cell states. This in-depth examination revealed a high level of complexity and granularity within our pancreatic cell atlas, enabling us to capture infrequent cell populations, such as endocrine epsilon cells (Figure S2F), which consist of 70 cells in our dataset (0.015%), and the 21 ionocyte-like cells (Mendeley Data: <http://www.doi.org/10.17632/9vwrk7zt7f.1> Additional Note 1) within the ductal compartment (0.0064%) (Figure S4).

While acinar (*RBPJL* and *PRSS1*) and ductal cells (*CFTR* and *MUC5B*) comprise about 90% of all cells (Figure S3H), followed by the endocrine cells with less than 2% (*INS*, *GCG*, *SST*, and *PPY*), we uncovered a diverse cellular landscape within the pancreatic tissue, including stellate cells (*ACTA2* and *PDZRN4*), Schwann cells (*SCN7A* and *CDH19*), fibroblasts (*COL6A3* and *COL4A1*), and endothelial cells (*FLT1*, *VWF*, and *CDH5*), as well as immune cells, both lymphoid (*CD3E*) and myeloid (*CD163*), which were consistently detected in all the donors (Figures S3H and S3I).

Our findings confirmed previously described acinar cell subtypes,¹³ with secreting acinar (Acinar-s) cells expressing *AMY2A*, *CLPS*, and *CELA2B* and idling acinar (Acinar-i) cells characterized by the expression of *INSR*, *FOXP2*, *RBPJL*, and *MAP3K5* (Figure S3A). Furthermore, we identified a distinct cluster of transitioning cells (referred to as Acinar-trans in Figure 1D), which were marked by the expression of *RBPJL*, *LRIG1*, and *REG3G* (Figure S3A), clearly representing an intermediate state between the Acinar-i and Acinar-s subtypes (Figure S3F). The central cluster, Acinar_7, was particularly interesting, as it contained a distinct subset of cells (Figure S3C), exhibiting a hybrid acinar-ductal phenotype, characterized by high expression of ductal markers such as *CFTR*, *BICC1*, and *SLC4A4*, reminiscent of centroacinar cells (CACs). Similarly, within the ductal population, we identified previously described main groups,¹³ including the conventional Ductal (*CFTR*, *SLC4A4*, *BICC1*, *PDE3A*, and *PKHD1*) and Ductal-Muc⁺ (*KRT19/CK19*, *MUC5B*, and *MUC1*) along with a distinct subpopulation, Ductal_4, that expressed both ductal and acinar markers (Figures S3D and S3G), which could suggest transcriptional transition within the exocrine compartment.

Chromatin accessibility profiling and integration with snRNA-seq data

We analyzed the chromatin accessibility of 197,946 high-quality nuclei from 10 donors previously profiled by snRNA-seq (Figures 1B and 1C). Unsupervised clustering identified 19 cell groups consistently across donors (Figures 1D and S5A). Cell identity was established by integrating gene activity scores derived from chromatin accessibility signals with snRNA-seq-defined markers and further refined using a Multiome dataset as a cross-modal bridge (STAR Methods; Figures S5B–S5D). This integration yielded strong concordance between modalities for major pancreatic cell types, with consistent expression patterns of key acinar and ductal markers (Figures 2A, 2B, and S5E).

While most cell populations were robustly identified, acinar and ductal states showed partial overlap in chromatin

Figure 1. Overview of ESPACE human pancreas cell atlas

(A and B) Number of donors within and shared across different methods.
(C) Cells/nuclei distribution from all experiments across datasets, regions, and disease.
(D–F) Uniform manifold approximation and projection (UMAP) representation of adult, fetal, and islet datasets. Colors/numbers indicate cell identity (see table).
(G) Xenium spatial transcriptomics of a representative adult pancreas core (H&E and immunofluorescence [IF], left). Inset (white square) shows transcript counts on H&E (top right) and cell-type labels on IF (bottom right). Colors denote transcripts/cell types (see legend). Scale bars: 500 μ m for the main image and 100 μ m for the magnified views.
(H) CODEX multiplexed immunohistochemistry of a representative adult pancreas TMA core with cell-type annotations (top left). Inset (white square) shows cell-type labels on nuclei (bottom left) and selected protein stainings (right, see legends). Scale bars: 500 μ m for the main image, 200 μ m for the magnified view (bottom left), and 100 μ m for the right panels.

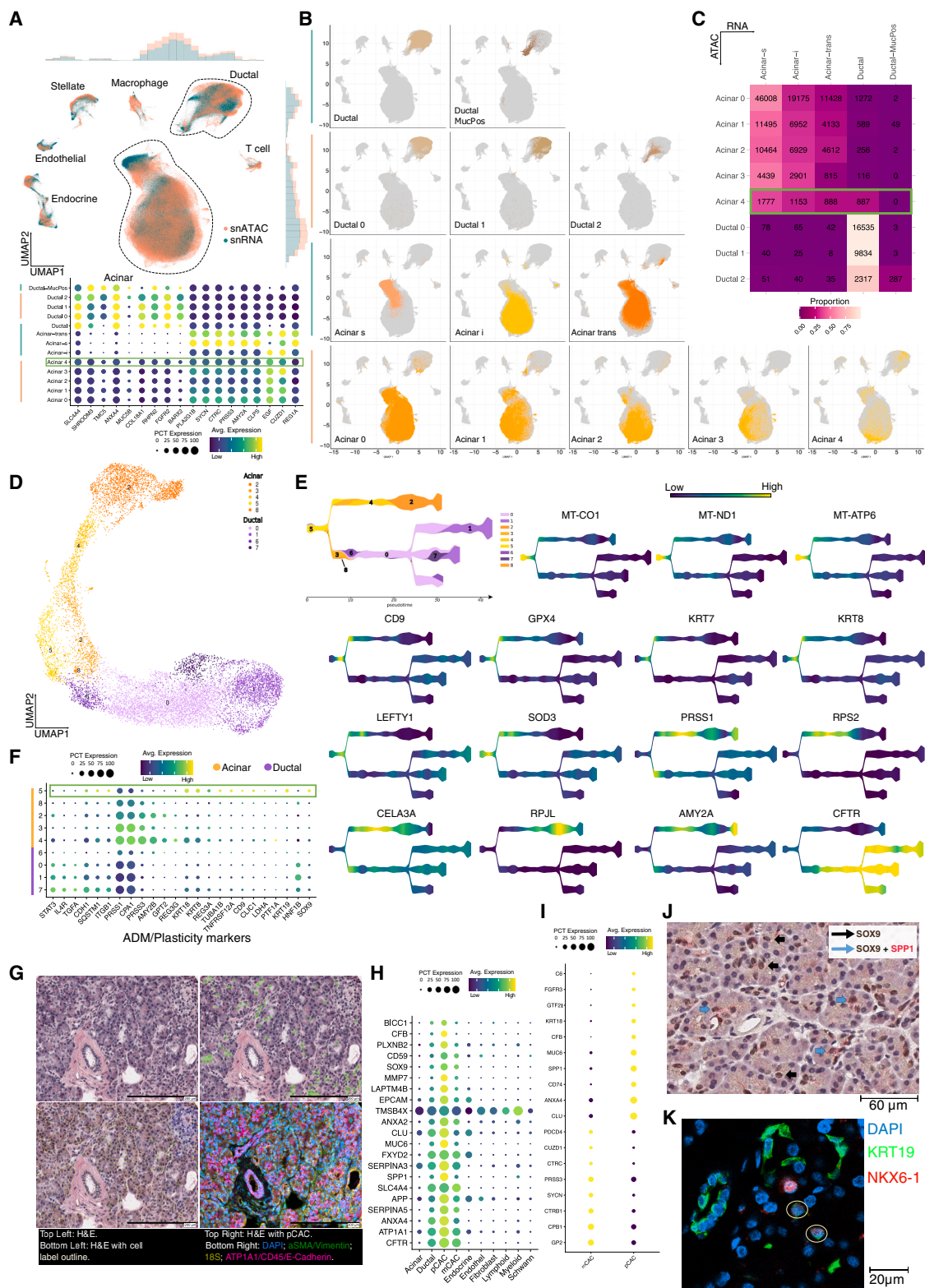


Figure 2. Transcriptional and epigenetic heterogeneity of the human adult pancreas

(A) UMAP of 636,341 integrated snRNA-seq and snATAC-seq cells (bridged via Multiome data), colored by technology. Marginal plots show cell distributions. Dot plots display canonical acinar and ductal markers based on individual dataset annotations.
(B) UMAP showing the location of acinar and ductal subclusters identified by each technology.

(legend continued on next page)

accessibility space, suggesting the presence of intermediate transcriptional states not fully resolved by either modality alone (Figure 2C). Notably, the Acinar_4 cluster in the snATAC-seq data exhibited the strongest correspondence to the ductal phenotype and mapped onto the Acinar_7/Acinar-trans cluster identified in the snRNA-seq analysis, a population co-expressing acinar and ductal markers (Figures 2C, S5F, and S15; Mendeley Data: <http://www.doi.org/10.17632/9vwrk7zt7f.1> Additional Note 3). This convergent finding across both transcriptomic and chromatin accessibility data independently supports the existence of a hybrid acinar-ductal cell state.

High-resolution mapping of acinar and ductal cells uncovers rare pCACs in the adult human pancreas

Intrigued by the presence of cell states with mixed acinar-ductal features, we focused our analysis on cluster 7 from the acinar population and a subset of cluster 4 from the ductal population in the snRNA-seq dataset (Figures S3C and S3D). Acinar-to-ductal plasticity has already been described in disease contexts (e.g., ADM), typically in mouse models^{1,21} but has not previously been characterized at this resolution in healthy human pancreas tissue. Reanalysis of these populations revealed that, rather than forming two distinct groups, they converged into a shared root cluster (cluster 5) with a reduced lineage-committed transcriptional profile (Figure 2D). Using STREAM,²² we reconstructed the complex branching trajectories from both snRNA-seq and snATAC-seq data (Mendeley Data: <http://www.doi.org/10.17632/9vwrk7zt7f.1> Additional Notes 2 and 3), revealing a continuum originating in cluster 5 that progressively diverged toward either a more acinar-like (clusters 2 and 4) or ductal-like identity (clusters 0, 1, 6, and 7) (Figure 2E). The resulting trajectories illustrated this branched transition, beginning with early activation of mitochondrial genes²³ (*MT-CO1*, *MT-ND1*, and *MT-ATP6*); markers of cellular adaptation and plasticity such as *CD9* and *GPX4*^{24,25}; metabolic regulators including *SOD3*; and cytoskeletal genes *KRT7* and *KRT8*, previously associated with ADM and pancreatic remodeling.^{23,26,27} Later stages were marked by expression of mature acinar (*PRSS1* and *AMY2A*) and ductal (*CFTR*) lineage genes.

Additionally, the co-expression of acinar and ductal markers in cluster 5, together with high levels of ADM- and plasticity-associated genes such as *KRT19*, *SOX9*, and *ITGB1*,^{28–30} further supports that cluster 5 represents a plastic cell population resembling centroacinar cells. Here, the term “plastic” is used to denote a trajectory-supported root state within a branched acinar-ductal continuum, characterized by concurrent acinar and ductal transcriptional features together with activation of es-

tablished ADM- and plasticity-associated gene programs, rather than experimentally demonstrated lineage conversion.

We then validated the expression of cluster 5 markers at the protein level using immunohistochemistry (Mendeley Data: <http://www.doi.org/10.17632/9vwrk7zt7f.1> Additional Note 4). *SOX9* expression was confirmed at the protein level in ductal epithelial and centroacinar cells of the adult human pancreas, with a high positivity rate across analyzed tissue cores (Figure S6A). In parallel, we identified *ANXA13* as a marker of this transitional cell population, which presented broad expression in ductal and centroacinar compartments (Figure S6B). Notably, *ANXA13* has not previously been linked to centroacinar cells, suggesting its utility as an indicator of ductal-acinar plasticity. Based on this combined transcriptional and protein-level evidence, we define this population as plastic centroacinar-like cells (pCACs), distinct from canonical centroacinar cells and positioned along the acinar-ductal continuum.

Spatial and protein-level validation of the pCACs

To further validate the existence and spatial localization of pCACs, we performed *in situ* transcriptomic profiling using high-plex spatial *in situ* transcriptomics (Xenium) on pancreatic tissue (Figures 1D and 1G). We applied a module scoring approach (STAR Methods) based on the pCAC gene signature derived from our snRNA-seq data to identify the pCAC population within the spatial dataset (Table S3). The resulting *in situ* projection predominantly highlighted centroacinar cells but also included a subset of acinar and ductal cells. We anticipated that technical differences between the snRNA-seq and the spatial transcriptomics platforms could contribute to discrepancies in cell identification. To refine this classification, we leveraged pathologist-annotated centroacinar cells across the tissue microarray (TMA) sections. Based on these expert annotations, we generated a second gene module score reflecting morphology-defined centroacinar cells (mCAC), which produced a highly specific spatial pattern consistent with their expected anatomical location and morphology. Yet, the snRNA-seq data indicated that the pCAC population does not represent the full spectrum of centroacinar cells but rather a specific subset distinguished by elevated transcriptional plasticity. We therefore focused on the subset of centroacinar-like cells identified by both the single-nuclei pCAC transcriptional signature and the pathologist-derived annotations. This overlap enabled us to derive a refined and robust *in situ* signature for those we identified as the pCACs, which comprised 3.12% of all acinar and ductal cells in the adult snRNA dataset (Figure 2G) vs. the 1.43% found in the Xenium data. Differential gene expression

(C) Heatmap displaying the results of transferring snRNA-derived labels (reference) to snATAC data (query) through scOMM (colors indicate cell proportions). (D) UMAP of acinar or ductal cells (12,396 cells) displaying highly plastic profiles in the snRNA data. Acinar cells exhibiting a ductal-like profile and vice versa were merged and re-clustered for further analysis.

(E) STREAM trajectory inference for cells in (D). Highlighted are acinar-like (orange) and ductal-like (purple) lineages, and labels depict clusters shown in (D). Acinar/ductal identity and plasticity marker expression are displayed in the trajectory.

(F) Dot plot of ADM/plasticity marker expression across clusters from (D). Original cell-type assignments for these clusters are also indicated.

(G) Healthy pancreas H&E staining overlaid with Xenium analysis stainings (bottom right) and identified pCAC. Scale bar: 200 μ m.

(H) pCAC signature gene expression levels in plastic centroacinar cells (pCACs) and broad cell types in the Xenium sample.

(I) pCAC signature expression level comparison between cells identified as pCAC by the joint module score approach vs. morphology-defined centroacinar cells (mCACs).

(J) Simultaneous immunohistochemical staining of *SOX9* and *SPP1* in an adult pancreas tissue core section. Scale bar: 60 μ m.

(K) Immunofluorescence for *KRT19* and *NKX6-1* in pancreas exocrine and endocrine portions of two different adult pancreas donors. Scale bar: 20 μ m.

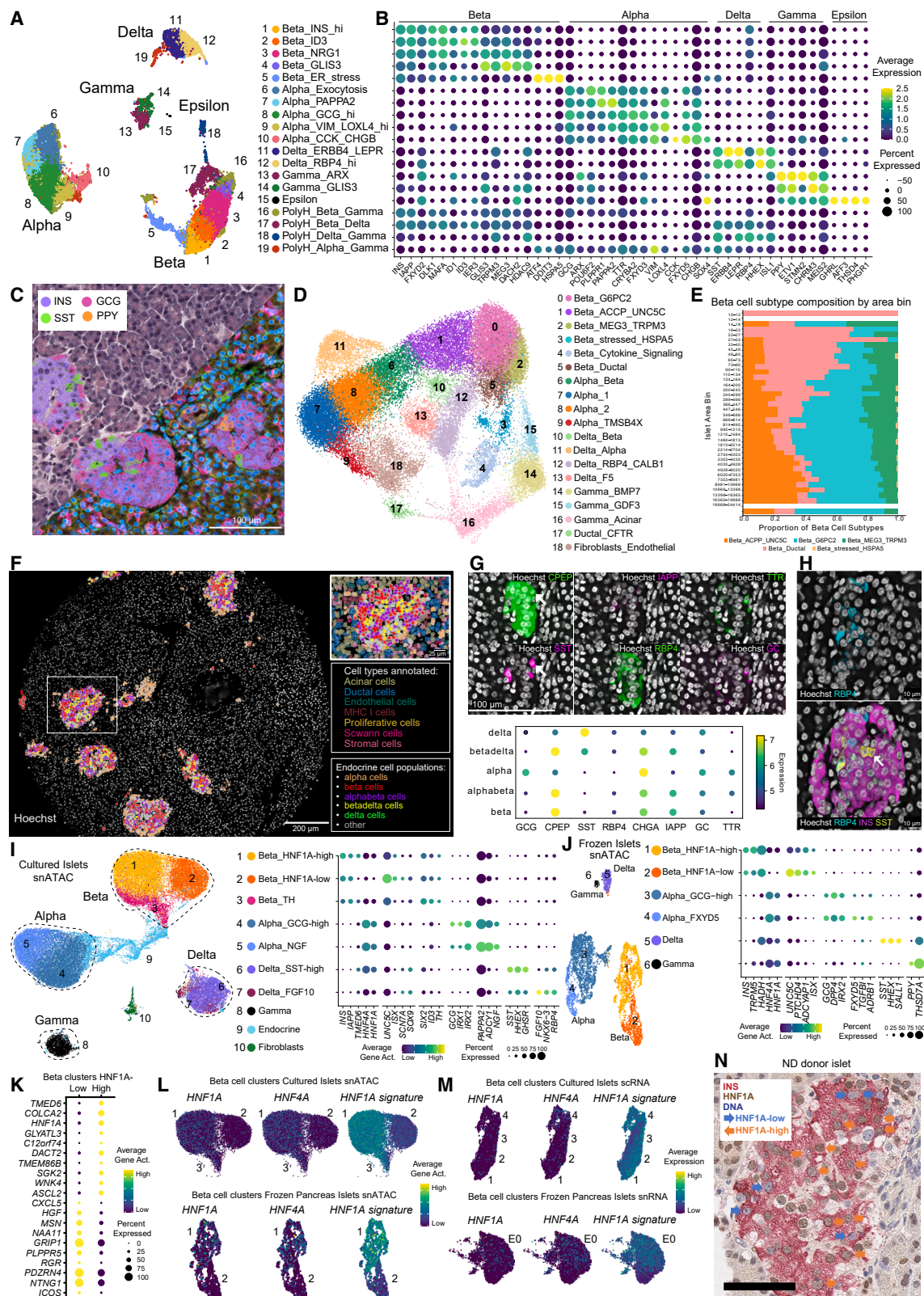


Figure 3. Heterogeneity of adult pancreas islet endocrine cells

(A and B) scRNA-seq UMAP projection of cultured islet endocrine cells ($n = 27,476$), colored according to cell-type subpopulations, along with dot plot showing expression of cell-type subpopulation-specific gene markers.

(legend continued on next page)

analysis (Table S4) showed that pCACs present high co-expression of markers such as *CFTR*, *SLC4A4*, *BICCT1*, and *SOX9*, consistent with a ductal and centroacinar identity (Figure 2H). By contrast, mCACs exhibited a more acinar-like transcriptional profile, with top differentially expressed genes including *CPB1*, *CTRB1*, *PRSS3*, *CTRC*, *SYCN*, and *CUZD1* (Figure 2I). Notably, the pCACs also expressed a distinct set of genes associated with plasticity, stress response, and differentiation, such as *CLU*, *ANXA4*, *KRT18*, and *SPP1* (Figure 2I), highlighting their transcriptional divergence and potential progenitor-like state.

Next, we systematically mapped the spatial distribution of pCACs (Figure S6C) and interrogated their surrounding microenvironment. Neighborhood enrichment analysis (STAR Methods) revealed that pCACs were consistently found in close proximity to myeloid cells (Figure S6D), predominantly macrophages, beyond the expected proximity to acinar and ductal cells. This pattern was robust across all TMA cores and donors (Figure S6E). While other stromal populations, such as lymphoid, fibroblasts, and endothelial cells, were occasionally nearby, their proximity varied depending on the specific core or sample. Interestingly, quantitative correlation analysis showed that the presence of myeloid cells had the strongest association with pCAC abundance (Pearson $r > 0.8$), whether measured across donors, across TMAs, or within individual cores.

Further protein-level validation included the identification of a subset of centroacinar-located cells co-expressing *SOX9* with the plasticity-associated marker *SPP1*, distinguishing pCACs from *SOX9*⁺/*SPP1*⁺ canonical centroacinar cells (Figure 2J). In addition, rare *KRT19*⁺/*NKX6-1*⁺ cells were detected in the exocrine pancreas outside islet regions, supporting the presence of progenitor-like states consistent with the transcriptional landscape inferred from snRNA-seq data (Figures 2K, S6F, and S6G). Together, these protein-level co-stainings support the transcriptionally inferred progenitor-like state of pCACs and are consistent with their low frequency identified in the snRNA-seq datasets (Figures 2D and 2E).

Transcriptional heterogeneity of adult pancreatic endocrine islet cells

Several studies have revealed substantial transcriptional heterogeneity among pancreatic islet endocrine cells, along multiple

non-overlapping axes.^{14,31–37} This diversity is thought to play a critical role in fine-tuning glucose homeostasis in health and disease.³⁸ To address this, we compared four complementary scRNA-seq datasets generated using different technologies: (1) snRNA-seq of islet endocrine cells from frozen adult pancreas sections, (2) scRNA-seq of isolated cultured islets, (3) snRNA-seq of isolated frozen islets, and (4) single-cell full-length mRNA VASA-seq of isolated cultured islets (Figures 1D, 1F, and 3A). This integrative approach enabled a systematic analysis of transcriptional diversity while pinpointing method-specific differences introduced by tissue processing, sequencing methodologies, and regional identity^{37,39} (Figure S7) (Mendeley Data: <http://www.doi.org/10.17632/9vwrk7zt7f.1> Additional Note 5).

We first focused our analysis on the five major human islet endocrine cell types (alpha, beta, gamma, delta, and epsilon cells) that were consistently detected across all datasets. We also observed subpopulations of endocrine cells with a polyhormonal identity (expressing more than one canonical hormone of the major endocrine groups: *INS*, *GCG*, *SST*, and *PPY*) (Figure S8B; Table S5). To identify robust marker genes of the endocrine populations, we performed a comprehensive analysis throughout the different datasets (STAR Methods), comparing the genes differentially expressed across these cell types, resulting in a consensus gene set panel for each population (Figure S8B; Table S6).

Detailed analysis of the isolated cultured islets scRNA-seq dataset (27,476 endocrine cells) unveiled various subpopulations that heterogeneously expressed genes associated with endocrine cell functions (Figures 3A and 3B; Table S7). Beta cells presented five distinct subpopulations with an expression gradient of beta cell canonical markers, specifically *INS* and *IAPP* (Figure 3B). The *INS*_high cluster 1 displayed higher expression of *INS* and mature beta cell marker *MAFA*, together with the highest expression levels of *IAPP*, *FXYD2*, *RBP4*, and *CDKN1C*, suggesting beta cells are highly active in insulin biosynthesis^{32,37} (Figure 3B). Cluster 2 presented increased expression of transcription factors *ID1* and *ID3*, a subpopulation previously described.³⁵ Cluster 3 presented an intermediate phenotype with intermediate levels of *NRG1* and *TRPM3*, a calcium channel involved in beta cell function.⁴⁰ Cluster 4 displayed lower *INS* levels together with increased expression of *GLIS3*, an important

(C) Xenium spatial transcriptomics of a representative islet-containing tissue core (H&E and IF staining, markers as Figure 1G), with endocrine hormone transcript counts overlaid. Scale bar: 100 μ m.

(D) UMAP projection of Xenium dataset endocrine cells (40,714) colored by cell-type subpopulations.

(E) Beta cell subtype distribution (Xenium dataset) across endocrine clusters of different size (binned by area, groups of 10–24,500 μ m²).

(F) Visualization of the CODEX dataset on islet cells. Representative tissue core containing islet cells segmented and annotated (see legends). Scale bars: 200 μ m in main image and 25 μ m in inset.

(G) Immunostainings (top) of key endocrine protein markers: CPEP, TTR, IAPP, RBP4, and SST. Arrows indicate hybrid cells co-expressing CPEP (beta) and SST (delta). Dot plot (bottom) shows protein levels across main endocrine populations. Scale bar: 100 μ m.

(H) Tissue core from non-diseased pancreas tail showing heterogeneous expression of pancreatic islet markers. Arrow indicates an example of an SST⁺ positive cell that is RBP4 negative. Scale bar: 10 μ m.

(I) Cultured islets snATAC-seq UMAP projection, colored by cell-type subpopulations. Dot plot depicting the gene body chromatin accessibility (gene activity) of cell-type-subpopulation markers.

(J) Adult pancreas snRNA-seq dataset endocrine cells UMAP projection, colored by cell-type subpopulations. Dot plot depicting gene activity of cell-type-subpopulation markers.

(K) Dot plot depicting gene activity of markers for beta cell subpopulations high and low in HNF1A activity (cultured islets snATAC-seq dataset).

(L and M) Feature plots representing the gene activity of *HNF1A*, *HNF4A*, and *HNF1A* signature genes in cultured and frozen beta cell subpopulations from snATAC-seq and sc/snRNA-seq datasets.

(N) Representative HNF1A and INS protein immunohistochemistry in islets of an adult pancreas from a donor with ND donor. Scale bar: 60 μ m.

transcription factor for beta cell development and function,⁴¹ *HDAC9*, and *ZNF385D*.⁴² Finally, beta cell cluster 5 expressed endoplasmic reticulum (ER) stress markers *ATF4*, *DDIT3*, and *HSPA5* (Figure 3B; Table S7).

Alpha cell subpopulations also presented significant transcriptional heterogeneity. These included clusters with varying levels of *GCG* and canonical alpha cell markers like *TTR* and *CRYBA2* (Figure 3B): cluster 6, with increased expression of genes involved in exocytosis, including *POU6F2*, *DCC*, and *NPAS3*; cluster 7, with elevated *PAPPA2*, *G6PC2*, *ASPH*, and genes involved in microtubule formation; cluster 8, with high levels of *LOXL4* and *VIM*, a profile associated with dysfunctional alpha cells⁴³; and cluster 9, with expression of *CCK* and *CHGB*. Similarly, delta cells presented two clear subpopulations, one with lower *SST* and increased *ERBB4* and *LEPR* expression (cluster 11), and one with increased *SST*, *HHEX*, *ISL1*, and *RBP4* (cluster 12) (Table S7). Next, we explored and validated the presence of these endocrine cell subpopulations in the complementary islet datasets (Figure S7), revealing distinct subpopulations for beta, alpha, and delta cells in all datasets and with similar markers as those seen in cultured islets (Figures S8C–S8F; Table S7).

Spatial transcriptomic and proteomic investigation of adult pancreatic endocrine cells

We leveraged spatial transcriptomics (Xenium) and proteomics (CODEX) to orthogonally validate the existence of endocrine subpopulations. Analysis of the Xenium dataset (Figure 3C) identified 40,714 endocrine cells with distinct endocrine cell types and subpopulations (Figure 3D; Table S5) that recapitulated the sn/scRNA-seq findings: beta cell subpopulations with increased expression of *INS* (cluster 1); elevated expression of genes involved in beta cell maturation and function: *G6PC2*, *MAFA*, and *NTN1* (cluster 0); smaller subpopulations with increased ER-stress markers *HSPA5* (cluster 3); cytokine signaling genes (cluster 4); and pancreatic ductal cell marker genes *CFTR* and *SERPINA3* (cluster 5).

We then explored the spatial distribution of these endocrine subpopulations using unsupervised cell neighborhood analysis, identifying endocrine cell clusters that varied in size over a wide range (10 to 25,000 μm^2 in area), from individual endocrine cells to well-defined islets (Figure 3E; Table S8). We identified numerous individual cells expressing endocrine hormone transcripts (*INS*, *GCG*, *SST*, *PPY*, and *GHRL*) outside islet structures, scattered through the pancreas parenchyma and within ductal structures (Figure S9). Distribution of beta cell subpopulations across the endocrine clusters of different sizes was not homogeneous, with beta cell cluster 5 (Beta_Ductal) being more abundant in smaller endocrine clusters (10–500 μm^2 in area), while beta cell clusters 0 and 1 mostly composing the bigger endocrine clusters/islets (2,000–20,000 μm^2 in area).

Xenium spatial transcriptomics uniquely preserves spatial context lost in dissociation-based approaches, revealing small endocrine cell populations dispersed outside canonical islet structures, within the parenchyma and ductal regions, with distinct transcriptional profiles (Table S8). Previously identified by immunostaining,⁴⁴ these populations have unknown physiological roles and may represent localized hormone sources, specialized metabolic functions,⁴⁵ endocrine regenerative potential,^{46,47} or cells at higher risk of autoimmune targeting.⁴⁸

Highly multiplexed single-cell imaging on the same donor tissue sections (Figure 3F) enabled protein expression analysis while preserving spatial context. Following unbiased cell segmentation, quantification, and clustering (Figures S10A–S10C), we annotated a total of 65,548 islet cells using an eight-marker antibody panel: CHGA, GCG, GC, and TTR (alpha cell-associated); CPEP and IAPP (beta cell-associated); and RBP4 and SST (delta cell-associated). This imaging-based approach provided a complementary view of islet cell heterogeneity and confirmed the presence of several transcriptionally defined subpopulations. Marker expression patterns across alpha, beta, and delta cells demonstrated considerable complexity, including cells with co-expression of multiple hormones, potentially reflecting transitional or polyhormonal states, and individual or small clusters of endocrine cells outside islet structures (Figure 3G). While the classical beta cell population expresses both CPEP and IAPP markers (Figure S10B), other beta cell populations, which are still defined by CPEP positivity, exhibited heterogeneous expression of RBP4 (Figure S10C). This variation is consistent with previous studies showing that RBP4-negative beta cells display enhanced electrophysiological function compared with their RBP4-positive counterparts,³⁷ suggesting that this proteomic variation may correspond to functional differences. A further example of intrapopulation heterogeneity was observed in delta cells co-expressing CPEP, IAPP, and varying levels of SST (Figure 3G, top row), mirroring transcriptional subtypes identified in the Xenium dataset. Another example consistent with transcriptomic findings is that delta cells were found to be either RBP4-positive or RBP4-negative (Figure 3H). Proteomic heterogeneity and complex expression patterns were not limited to beta and delta cells but extended to every identified population, as demonstrated in Figure 3G (bottom row).

These proteomic results confirm that endocrine cell heterogeneity is preserved at the protein level and within the spatial architecture of the pancreas, including the presence of transitional or hybrid polyhormonal phenotypes.⁴⁹ Notably, these hybrid cells were not randomly distributed. Beta cells frequently neighbored other hybrid beta cells co-expressing delta or alpha identity markers, suggesting spatially restricted microenvironments that may support phenotypic plasticity (Figure S10D), while more canonical alpha and delta cell populations were spatially distant from polyhormonal states (Figure S10E).

Together, these findings demonstrate that integrating spatial transcriptomics and multiplexed proteomics enables identification and localization of rare and transitional endocrine states within native pancreatic tissue, with implications for endocrine function, regeneration, and disease susceptibility.

We took advantage of snATAC-seq generated on the same isolated culture islet cells to investigate chromatin accessibility heterogeneity (Figure 3I). We identified subpopulations of beta, alpha, and delta cell types with distinct chromatin accessibility profiles, in line with the previous transcriptomic and proteomic observations (Table S9).

Notably, we also identified an additional level of heterogeneity in beta cells, related to high and low *HNF1A* transcription factor activity (Figure 3I). These two distinct beta cell clusters presented distinct levels of accessibility at the *HNF1A* gene body and its direct target genes (e.g., *TMED6*, *SGK2*, *A1CF*, and *HNF4A*) (Figure 3K; Table S9).⁵⁰ Similar beta cell clusters

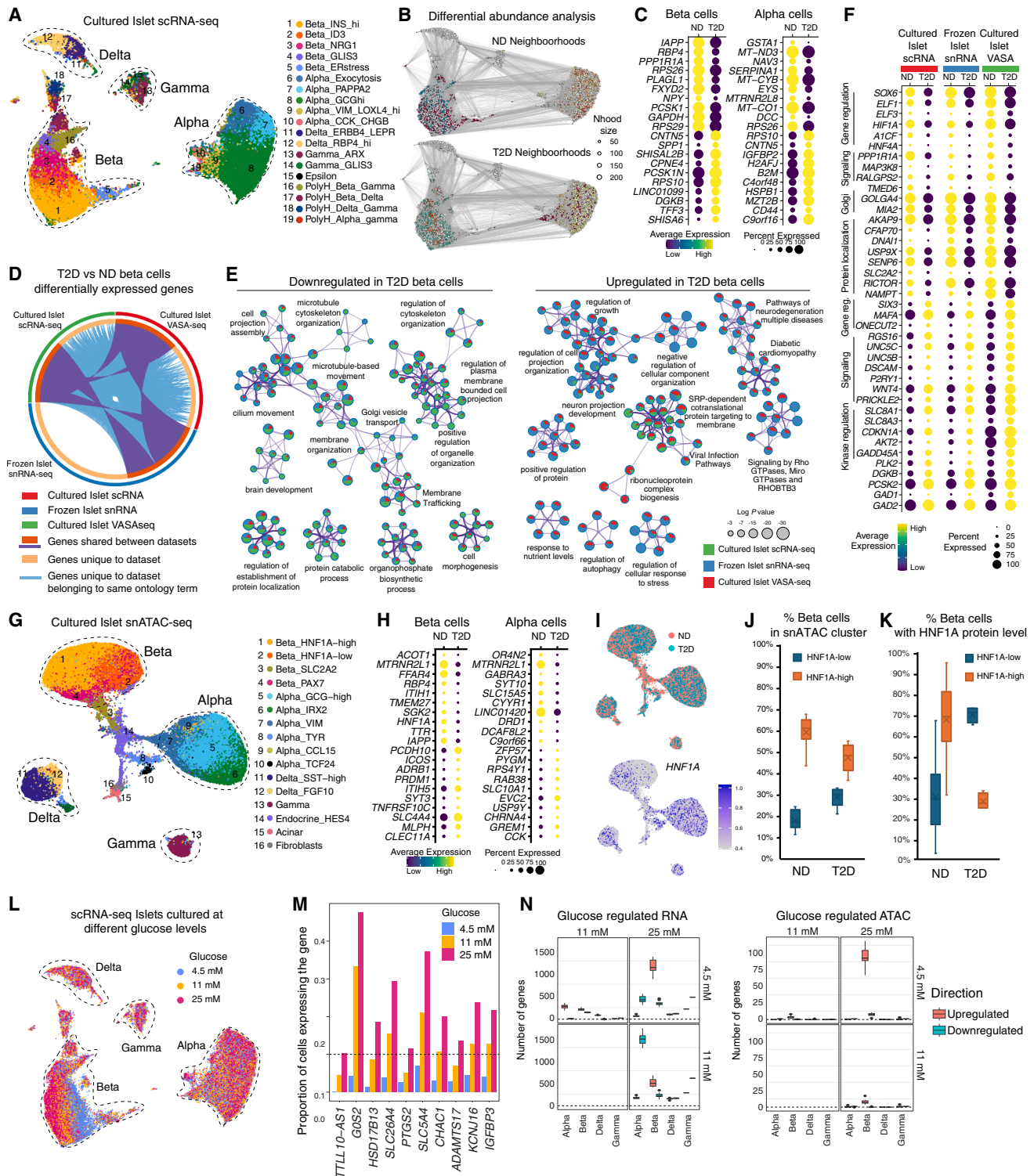


Figure 4. Regulatory landscape of human islets from individuals with T2D

(A) Cultured islets scRNA-seq endocrine cells UMAP projection integrating donors with type 2 diabetes (T2D) and no diabetes (ND) donors colored by cell-type subpopulations.

(B and C) Differential cell abundance analysis of the integrated T2D/ND dataset, with dot plots showing differentially expressed genes in beta and alpha cells.

(D) Differentially expressed genes in T2D beta cells shared across three datasets. Arc colors indicate genes shared between datasets (orange/purple), dataset-unique genes (beige), and genes sharing ontology terms (blue).

(E) Metascape networks of enriched ontology terms in T2D vs. ND beta cells downregulated and upregulated gene sets, integrating the three different datasets.

(legend continued on next page)

HNF1A high- and low-chromatin accessibility were detected in frozen pancreas snATAC-seq, confirming the presence of this heterogeneity axis in non-cultured, minimally perturbed, adult islet beta cells (Figure 3J; Table S9). We selected the top 25 genes with differential accessibility (i.e., gene activity) between these two beta cell populations in cultured and frozen beta cells (Figure 3K; Table S9) and utilized them to investigate in the snATAC and scRNA datasets the presence of cells with an enhanced *HNF1A* transcriptional program (Figures 3L and 3M). Furthermore, we conducted immunohistochemistry for *HNF1A* and *INS* on adult pancreas sections and observed that ~30% of *INS*⁺ beta cells presented lower levels of *HNF1A* protein in their nuclei, validating the heterogeneous expression of *HNF1A* in adult beta cells also at the protein level (Figure 3N). Thus, our studies support *HNF1A*-high and low beta cell states at protein, transcript, and transcription factor activity levels and demonstrate that this heterogeneity exists in frozen tissue that has not undergone islet culture or isolation processes.^{50–52}

Regulatory landscape of islets from individuals with T2D

We investigated the transcriptional and chromatin accessibility changes associated with T2D. Initially, we integrated scRNA-seq data from isolated cultured islets obtained from both donors with no diabetes (ND) and T2D (Figures 4A, S11A, and S11B). Leveraging scCOMM,⁵³ we compared the heterogeneity in endocrine cells observed in the ND dataset to that of the T2D dataset, identifying similar populations in both datasets (Figure S11C). We then performed differential cell-type abundance analysis in the integrated dataset using MILO,⁵⁴ which assigns cells to partially overlapping neighborhoods on a k-nearest neighbor graph. This unveiled various cell neighborhoods particularly enriched in ND and T2D (Figures 4B and S11D). Beta cells in ND neighborhoods presented increased expression of *IAPP*, *RBP4*, *PCSK1*, and *FXYD2*, together with ribosomal protein genes *RPS26* and *RPS29*, suggesting a beta cell subpopulation with enhanced insulin biosynthesis.⁵⁵ T2D beta cell neighborhoods presented increased expression of *SPP1*, *DGKB*, and *TFF3*, genes previously linked with beta cell dysfunction and development^{56–58} (Figure 4C; Table S10). Alpha cells in ND neighborhoods showed higher expression of mitochondrial genes involved in oxidative phosphorylation (*MT-ND3*, *MT-CYB*, and *MT-CO1*), while the T2D alpha cells presented increased expression of *IGFBP2*, *B2M*, and *CD44* (Figure 4C).

To investigate the genes consistently dysregulated in beta cells from T2D individuals, we compared the differentially expressed genes across three different datasets: (1) scRNA-seq of isolated cultured islets (5 ND and 6 T2D donors); (2) snRNA-

seq of isolated frozen islets (4 ND and 4 T2D donors); and (3) VASA-seq of isolated cultured islets (6 ND and 5 T2D donors). Despite major differences in tissue processing and sequencing technologies, we identified 87 genes that were downregulated on the T2D beta cells common to the 3 datasets, including transcription factors (*SOX6*, *ELF1*, *ELF3*, and *HIF1A*), direct *HNF1A* targets (*A1CF*, *HNF4A*, and *TMED6*),⁵⁰ regulators of signaling pathways (*MAP3K8* and *RICTOR*), and cilia-related genes (*CFAP70* and *DNAI1*) (Figure 4D–4F). In the intersection of T2D beta cells' upregulated genes (64 genes), we found transcription factors related to beta cell maturation (*MAFA*, *SIX3*, and *ONECUT2*),⁵⁸ suggesting a compensatory response to improve beta cell function. Also, we observed an upregulation of genes associated with islet immune response (*GAD1*, *GAD2*, and *HLA-A*), netrin receptors (*UNC5C*, *UNC5B*, and *DSCAM*), previously linked to beta cell survival and insulin secretion,⁵⁹ and signaling pathways (*P2RY1*, *AKT2*, *PLK2*, *WNT4*, *PRICKLE2*, *CDKN1A*, and *GADD45A*) (Figure 4F; Table S11). The overlap at the level of biological pathways in which these genes are involved was considerable: cilium and microtubule-based movement, membrane trafficking, regulation of organelle organization, cytoskeleton, protein localization, or cell projection were among the most enriched biological pathways in the genes that are downregulated in T2D beta cells across datasets. Conversely, regulation of growth, response to nutrients and stress, signaling by GTPases, and SRP-dependent cotranslational protein targeting to membranes were the biological pathways enriched in T2D beta cell-upregulated genes. These parallel analyses highlight the common dysregulated pathways across datasets.

We also explored the scATAC dataset of T2D and ND human islets and observed a reduced proportion of beta cells presenting the *HNF1A* high program in T2D donors, aligning with the downregulation of *HNF1A* targets observed in T2D scRNA-seq datasets (Figures 4G–4K; Table S11).

Islets cultured in different levels of glucose reveal cell-type-specific transcriptional programs

Pancreatic islets are exposed to a range of glucose concentrations in health and disease. We explored the impact of culturing in different glucose concentrations (4.5, 11, and 25 mM glucose, fasting, mild, and severe hyperglycemic levels, respectively) on the regulatory dynamics of endocrine cells using scRNA-seq and snATAC-seq and compared these effects with alterations induced by long-term *in vivo* hyperglycemia characteristic of T2D. Culture in 11 or 25 mM glucose significantly reduced insulin secretion capacity in a glucose-concentration-dependent

(F) Dot plot representing the expression of differentially regulated genes in T2D vs. ND beta cells common to the three datasets.

(G and H) Cultured islets snATAC-seq UMAP integrating T2D and ND donors, with dot plots showing differential chromatin accessibility in beta and alpha cells (Table S11).

(I) UMAP projection as in (G), colored by disease status (ND or T2D, top) and *HNF1A* gene body chromatin accessibility levels (bottom).

(J) Proportion of beta cells in *HNF1A*-high or *HNF1A*-low snATAC-seq clusters in samples from ND and T2D donors. Data presented in Box-Whisker plots, where horizontal line denotes the median, cross indicates the mean value, whiskers the maximum and minimum values, and box ranges from the 25% to the 75% percentiles. Same for data presented in (K).

(K) Proportion of *INS*⁺ beta cells with high or low *HNF1A* protein levels in ND and T2D donors, following immunohistochemistry as shown in Figure 3N.

(L) UMAP projection of endocrine cells from islets cultured at different levels of glucose.

(M) Proportion of beta cells expressing glucose-induced genes at different glucose concentrations.

(N) Number of glucose-regulated genes in different islet cell types and glucose concentrations at RNA and ATAC levels. Data presented in Box-Whisker plots, where horizontal line denotes the median, whiskers the maximum and minimum values, and box ranges from the 25% to the 75% percentiles.

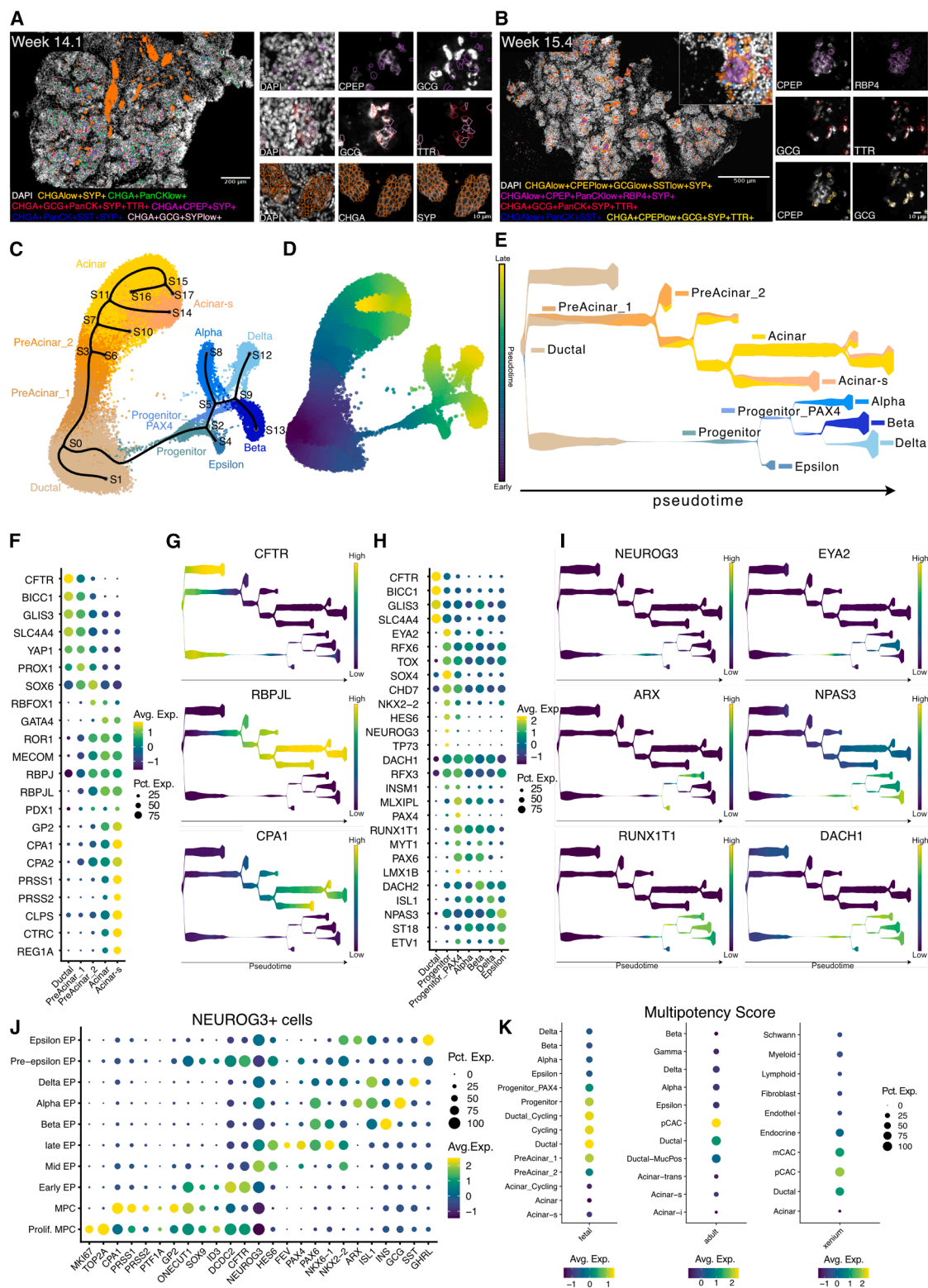


Figure 5. Transcriptional heterogeneity of human fetal pancreas

(A and B) Multiplexed imaging of endocrine populations at 14.1 (A) and 15.4 (B) wpc, with insets showing beta (purple), alpha (red/pink), CHGA+SYP+ (orange), and alpha-beta (yellow) cells. Scale bars: 200µm (A), 500µm (B), and 10 µm for all insets.

(C–E) STREAM trajectory analysis of the non-cycling fetal epithelial subset: UMAP with inferred branches (C), pseudotime visualization (D), and cell density stream plots colored by cell type (E).

(legend continued on next page)

manner (Figures S12A–S12F), without affecting total insulin content (Figure S12G). These glucose concentrations induced different condition-specific transcriptional profiles (Figures 4L–4N; Table S11) between cell types: alpha and beta cells showed increased transcriptional responses to glucose variation compared with delta or gamma cells. Using snATAC-seq, beta cells presented increased gene responsiveness to elevated glucose compared with other cell types (Figure 4N; Table S11). Altogether, this suggests endocrine cell-type-specific patterns of responses to increasing glucose concentrations (Mendeley Data: <http://www.doi.org/10.17632/9vwrk7zt7f.1> Additional Note 5).

Developmental trajectories of human fetal pancreas

Human pancreas development remains poorly understood due to limited access to fetal tissue. We performed snRNA-seq (frozen material) and high-plex spatial proteomic imaging (FFPE) from dissected fetal pancreases, ranging from 10 to 18 wpc (Figures 5, S13, and S17–S20; Table S12).

Following QC and ambient RNA correction (STAR Methods), the fetal snRNA-seq dataset comprised 87,117 high-quality nuclei across 44 transcriptionally distinct populations, including acinar, ductal, endocrine, mesenchymal, endothelial, neuronal, and immune cells. Rare and transitional populations, such as pre-acinar and *NEUROG3*⁺ endocrine progenitors, were identified alongside canonical alpha, beta, delta, and epsilon endocrine subsets (Figures 1E, S13B–S13D, and S14). Complementary spatial proteomics from 224,713 FFPE-embedded cells validated these populations across developmental stages (Figures 5A, 5B, and S17–S20) (see also Mendeley Data: <http://www.doi.org/10.17632/9vwrk7zt7f.1> Additional Note 6 and 7).

We inferred developmental trajectories in the fetal pancreas using STREAM,²² reconstructing three main branching cellular trajectories (Figures 5C–5E): (1) A short ductal differentiation trajectory, likely representing trunk progenitors development into mature ductal cells; (2) a ductal to acinar differentiation trajectory, characterized by cells expressing both ductal (*CFTR*, *BICC1*, and *GLIS3*) and acinar-related genes (*GP2*, *CPA2*, and *CLPS*), shifting toward acinar markers (Figures 5F and 5G); and (3) an endocrine differentiation trajectory, starting with cells expressing *NEUROG3* transitioning from ductal to endocrine lineage (Figures 5H, 5I, and S14).^{59,60} Subsequent endocrine progenitors in the trajectory expressed transcription factors important for lineage allocation, such as *RFX6*, *PAX4*, *ARX*, and *RUNX1T1*.^{61–63} Pseudotime analysis pinpointed other transcription factors that could potentially play a role in endocrine cell development (*EYA2*, *NPAS3*, *DACH1*, and *DACH2*) (Figures 5I and S13E). The endocrine progenitor trajectory branches into developing epsilon cells (*GHRL* and *GHRLOS*) and a branch transcriptionally similar to endocrine progenitors but with reduced *NEUROG3* and increased *PAX4* expression (“Progenitor_PAX4” cluster). This bifurcated into the alpha cell cluster (*GCG* and *TTR*), with PPY-positive cells (i.e., early gamma

cells) (Figure S13D), then into beta (*INS* and *MAFA*) and delta cell (*SST* and *RBP4*) branches, covering all major islet cell types.

NEUROG3 expression during fetal pancreas development is critical for the differentiation of endocrine cells.^{59,64} We observed the highest *NEUROG3* expression in arising endocrine cells but also at a lower proportion in the clusters annotated as acinar and ductal cells. Investigation of *NEUROG3*-expressing cells identified 10 subpopulations (Figures 5J and S14A–S14E).^{65,66} We identified early endocrine progenitor (EP) cells, expressing ductal markers *DCDC2*, *CFTR*, *SOX9*, *ID3*, and highly *ONECUT1*; mid EP cells, with the highest expression of *NEUROG3*; and late EP cells, expressing *FEV*, *PAX4*, *PAX6*, *NKX2-2*, and *NKX6-1*. These EP stages presented reduced expression of endocrine hormone genes. Additionally, clusters with clear endocrine hormone expression profiles were identified: pre-epsilon and epsilon EP cells expressed *GHRL* and *NKX2-2* and were transcriptionally close to mid EP, while beta EP (*INS*⁺), alpha EP (*GCG*⁺), and delta EP (*SST*⁺) cells were more similar to late EP cells (Figures S14C–S14E). Cells with a profile that is reminiscent of mouse multipotent progenitor cells (MPCs) expressed markers such as *GP2*, *PTF1A*, *ONECUT1*, and acinar-specific genes *CPA1*, *PRSS1*, and *PRSS2*. While MPCs expressing *CPA1* and *GP2*^{8,67} are described in fetal pancreas, the presence of *NEUROG3* in these cells has not been previously reported. A subset of MPCs presented active proliferation markers (*MKI67* and *TOP2A*), a cell cycle gene signature that decreased in more differentiated EP cell types (Figure S14D). We observed a lower but detectable *NEUROG3* expression in multipotent and proliferative cells when compared with fetal acinar cells and mid EP cells (Figure S14E). This suggests a *NEUROG3*-low MPC during endocrine cell differentiation, similar to what has been described in mice.^{64,68} We examined Notch-signaling pathway gene expression, crucial for pancreatic development^{69,70} (Figure S14F), observing high *DLL1*, *YAP1*, *HES1*, and *JAG1* expression in MPCs and early EP, with a consistent decrease in mid EP and the specified endocrine cell EP populations that coincides with the increased *NEUROG3*.⁷¹

Comparison of transcriptional programs in adult and fetal pancreatic cells

This ESPACE atlas enables comprehensive comparison of fetal and adult snRNA-seq datasets generated identically, minimizing batch effects. We subset both datasets into four compartments (acinar, ductal, endocrine, and mesenchyme), integrated them, and calculated differentially expressed and conserved genes between adult and fetal datasets (STAR Methods; Mendeley Data: <http://www.doi.org/10.17632/9vwrk7zt7f.1> Additional Note 8; Figure S20; Table S13), generating a resource to evaluate the transcriptional difference between immature fetal and mature adult pancreatic cells.

Connecting developmental insights with our findings in the adult pancreas, we revisited the plastic centroacinar cells

(F and G) Exocrine differentiation. (F) Dot plot of ductal to acinar genes and (G) stream plots of *CFTR*, *RBPJL*, and *CPA1* expression over pseudotime.

(H and I) Endocrine differentiation. (H) Dot plot of ductal to endocrine genes, and (I) stream plots of *NEUROG3* and differentiation transcription factors over pseudotime.

(J) Dot plot of *NEUROG3*⁺ cells showing markers of multipotency and cell identity.

(K) Multipotency module score across fetal, adult, and Xenium datasets.

(pCACs) previously identified (Figures 2G–2K). We observed preferential expression of MPC markers in ductal and pCAC populations in adult pancreas snRNA-seq and Xenium datasets (Figures S14G–S14I) and that adult pCACs share transcriptional features with fetal acinar-precursor populations, particularly Pre-Acinar_1 (Figure S15G), suggesting a conserved or reactivated gene program. For further investigation, we curated a set of multipotency-associated genes from fetal pancreas studies^{72–74} and computed a module score across all cell types (STAR Methods). Among adult cells, pCACs displayed the highest multipotency scores, with expression levels comparable to the fetal progenitor-like populations (Figure 5K). To validate this observation in an independent spatial dataset, we applied the same module score to the Xenium adult pancreas data and again found that pCACs had the highest multipotency score among all cell types (Figure 5K), reinforcing their unique transcriptional identity and suggesting persistence of developmental programs.

DISCUSSION

The ESPACE atlas provides a foundational single-cell multiomics reference of the human pancreas, resolving cellular diversity and plasticity at a scale and resolution not previously achieved. A key strength of ESPACE lies in its integration of diverse single-cell technologies, offering a unified framework to understand pancreas biology in unprecedented detail. We leveraged this unique dataset to uncover rare but transcriptionally distinct pCACs, an intermediate state along the acinar-to-ductal axis. Identified in snRNA-seq and snATAC-seq, these cells co-express acinar and ductal markers and display elevated expression of ADM- and plasticity-associated genes. Spatial transcriptomics revealed that these cells are consistently located near myeloid and also lymphoid immune cells, suggesting an association with reactive or inflammatory environments. Proteomic validation confirmed their centroacinar identity via SOX9 and identified ANXA13 as a marker. Further protein-level validation through multiplex immunohistochemical co-staining demonstrated the existence of SOX9⁺/SPP1⁺ double-positive centroacinar cells, supporting the presence of a plastic subset within the broader CAC population, and identified rare KRT19⁺/NKX6.1⁺ double-positive cells in the exocrine compartment, consistent with the progenitor-like transcriptional state of pCACs. Integration with fetal pancreas data revealed that pCACs share transcriptional features with fetal acinar progenitors, indicating that these adult cells retain or reactivate a developmental gene program associated with cellular plasticity. Altogether, these findings suggest that the adult human pancreas retains a transcriptionally plastic niche potentially primed for regeneration or metaplasia, an observation with significant implications for pancreatic cancer initiation and repair processes.

In the endocrine compartment, multimodal profiling revealed substantial heterogeneity across all five major islet cell types. Notably, beta cells exhibited epigenetic divergence at the HNF1A locus, giving rise to HNF1A-high and HNF1A-low chromatin accessibility states, an axis of regulatory heterogeneity that has been linked to distinct electrophysiological and exocytosis traits.^{50–52} We validated this heterogeneity across technologies and sample sources, including ex vivo snap-frozen tissue,

with immunostaining further confirming this subtype-specific marker expression at the protein level.

The vast majority of pancreatic endocrine cell omics studies isolated islets or sorted cells,^{33–37} thus introducing at least 2 potential confounders: the stress applied to isolate and culture the islets and the inability to analyze endocrine cells outside well-developed islets. Our study addresses these limitations by profiling nuclei from frozen adult pancreas sections and performing spatial transcriptomics, enabling the unbiased investigation of all endocrine cells present in the pancreas with reduced isolation-induced artifacts. We uncovered endocrine-like cells situated outside conventional islet structures, often in close proximity to pancreatic ducts, expressing a set of genes associated with progenitor-like or plastic states (e.g., CD44, CD24, and CD9).⁷⁵ The proximity of these cells to ductal regions has previously been interpreted as evidence of a ductal reservoir of facultative progenitors.^{44,75–77} However, extensive genetic lineage-tracing and cell-ablation studies in mice have not supported a contribution of adult ducts to the beta cell pool, and lineage tracing has not yet been performed in the human pancreas.^{4,78–81} Our data therefore identify a distinct transcriptomic signature whose developmental or dynamic significance remains to be established.

From a disease perspective, our atlas provides new insights into T2D-associated alterations. In T2D donors, we observed changes in beta cell subpopulation abundance, with a notable reduction in the HNF1A-high epigenetic state that cannot be explained by isolation or cell culture artifacts. Transcriptionally, T2D islets displayed putative compensatory (e.g., increased MAFA, SIX3, and ONECUT2) and dysfunctional (e.g., down-regulated HNF1A targets) programs, suggesting concurrent adaptation and failure. Exposure of ND islets to varying glucose concentrations recapitulated some of these patterns and revealed cell-type-specific transcriptional and epigenetic responses, particularly pronounced in beta cells. These data underscore the dynamic regulatory landscape of islet cells and suggest that metabolic context can influence cell identity and function in a manner akin to chronic disease.

Our profiling of the fetal pancreas provides a high-resolution view of human pancreatic development. Pseudotime analyses reconstructed the three major lineage trajectories, acinar, ductal, and endocrine, and identified transcription factors, such as EYA2, NPAS3, DACH1, and DACH2, which are not yet described in the context of endocrine lineage commitment. The fine-grained dissection of NEUROG3⁺ cell states revealed intermediate populations poised for commitment to all endocrine lineages and expands our understanding of human islet ontogeny. Comparison to adult data revealed that fetal and adult cell types, while sharing classical cell-type identity markers, remain transcriptionally distinct. In the endocrine compartment, alpha, delta, and epsilon cells exhibit continuum-like transitions toward their mature states, whereas fetal and adult beta cells remain transcriptionally separated, highlighting the postnatal maturation of beta cells.

Ultimately, ESPACE provides the community with the tools needed to interrogate fundamental questions in pancreatic function, regeneration, and pathology and serves as a blueprint for future multimodal organ atlases illuminating human biology at single-cell resolution.

Limitations of the study

Despite its comprehensive and multimodal scope, the ESPACE atlas has several limitations. Although we included samples from 57 individuals and spanned diverse anatomical regions and disease contexts, the availability of matched data across all modalities (e.g., spatial, transcriptomic, and epigenetic) for every donor was limited, potentially constraining some integrative analyses. We showed that enzymatic dissociation and *in vitro* culturing of islets, although controlled, introduce stress-related transcriptional artifacts, particularly in endocrine cells. Similarly, islet functional assays may be influenced by isolation, transport stress, and partial dedifferentiation, which should be considered when interpreting the results. While we addressed this by including frozen nuclei-based assays, certain dynamic cellular states may still be under- or over-represented. snA-TAC-seq remains at lower resolution compared with RNA-based modalities, limiting fine-grained identification of rare regulatory and transitional states. The use of multimodal (RNA+ATAC) assays would further refine the correspondence between transcriptional and epigenetic states, both in exocrine and endocrine populations, improving the cross-modal label transfer and deep-learning-based integration used here. Furthermore, ambient RNA contamination is especially pronounced in pancreatic tissue due to the exceptionally high transcriptional and ribosomal activity of endocrine and acinar cells. We applied CellBender⁸² to mitigate ambient RNA contamination across samples, as it has shown strong performance in multiple benchmarking studies.^{83,84} However, ambient RNA could not be fully eliminated, as suggested by the spurious detection of established marker genes in unexpected cell populations. To ensure transparency and flexibility for downstream analyses, we provide the raw, uncorrected count matrices, allowing users to apply alternative ambient RNA correction strategies according to their preferences.

Single-cell and spatial omics provide powerful but complementary tools for tissue profiling, each with distinct technical limitations. In our fetal pancreas snRNA-seq dataset, mesenchymal cells were markedly underrepresented compared with spatial proteomics, likely reflecting cell-type-specific sensitivity to nuclear isolation, a known recovery bias that underscores the value of multimodal integration.^{85,86}

Additionally, although spatial transcriptomics avoids dissociation bias, accurate cell segmentation remains a major challenge. Imperfect segmentation can distort cell identities and hinder cross-platform integration, a problem exacerbated by the sparse nature of spatial transcript counts. While segmentation-free approaches reduce some bias, they lack single-cell resolution and compatibility with atlas-level analyses.^{87,88} We adopted a segmentation-based strategy to enable direct cell-level comparison with matched snRNA-seq datasets, facilitating cell-type annotation, label transfer, and atlas integration despite these limitations. Spatial proteomics faced analogous challenges due to heterogeneous cell morphology, particularly in densely packed fetal tissue, requiring customized segmentation parameters. Additionally, two-dimensional tissue sectioning may conflate transcripts from vertically adjacent cells, producing hybrid profiles, especially in densely packed structures such as islets, a limitation that emerging 3D spatial reconstruction approaches (e.g., CODA⁸⁹) may help address in future studies.

Finally, while our cross-sectional design supports inference of developmental trajectories, direct lineage tracing and functional validation remain important directions for future work.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Roland Eils (roland.eils@bih-charite.de).

Materials availability

This study did not generate any new, unique reagents.

Data and code availability

- Data have been deposited at the European Genome-phenome Archive (EGA), which is hosted by the EBI and the CRG, under accession number EGAS50000001872. Further information about EGA can be found at <https://ega-archive.org> and “The European Genome-phenome Archive of human data consented for biomedical research.” An interactive visualization for exploration and investigation of the data is available at the <https://singlecell.charite.de> website as well as downloadable data objects.
- The analysis pipelines used for spatial proteomics (CODEX) data processing are available on GitHub (https://github.com/CellProfiling/ell_espace). Raw and processed spatial proteomics datasets, including microscopy images, cell segmentation masks, single-cell quantifications, and cell-type annotations, have been deposited in the Biolineage Archive under accession S-BIAD3557 (<https://www.ebi.ac.uk/biostudies/bioimages/studies/S-BIAD3557>).
- Raw data used to create graphs in Figures 4J, 4K, and S12A–S12G are presented in Data S1. Additional information related to this manuscript can be found in the Mendeley repository: Mereu, Elisabetta; Liebig, Johannes; Balboa, Diego; and Gonzalez, Aitor (2026), *A single-cell multi-omic atlas of the human pancreas reveals potential for cellular plasticity across development, health, and disease*, Mendeley Data, V1, doi: [10.17632/9vwrk7zt7f.1](https://doi.org/10.17632/9vwrk7zt7f.1).

CONSORTIA

Dana Avrahami, Diego Balboa, Frederic B. Navarro, Edgar Bernardo, Frida Björklund, Françoise Carlotti, Robert L. Chua, Christian Conrad, Eelco de Koning, Roland Eils, Marten Engelse, Jorge Ferrer, Javier García-Hurtado, Benjamin Glaser, Aitor Gonzalez-Herrero, Nathalie Groen, Maaike Hanegraaf, Bianca Henning, Holger Heyn, Pablo Iañez, Katharina Jechow, Björn Konukiewicz, Christian Lawrenz, Nadja Lewandowski-Hoppe, Johannes Liebig, Emma Lundberg, Domenica Marchese, Mariya Mardamshina, Anna M. Casals, Elisabetta Mereu, Fanny Mollandin, Mauro Muraro, Silvia Pellegrini, Lorenzo Piemonti, Felix Schickentanz, Inna Sitnik, Valeria Sordi, Rebecca Maltez de Sousa, Katja Steiger, Alexander Sudy, Ulrike Taron, Foo Wei Ten, Luca Tosti, Timo Trefzer, Sven Twardziok, Maarten van Aken, Alexander van Oudenaarden, Valerie Vandenbempt, Wilko Weichert, Johannes Wirth.

ACKNOWLEDGMENTS

This study was supported by the European Commission (ESPACE, 874710, Horizon 2020). This publication is part of the Human Cell Atlas—www.humancellatlas.org/publications. We further acknowledge the following: the Spatial Proteomics Facility, Royal Institute of Technology, funded by Science for Life Laboratory and National Microscopy Infrastructure, NMI (VR-RFI 2016-00968) for the PhenoCycler-Fusion experiments on fetal pancreas; Human pancreatic islets were provided by the NIDDK-Funded Integrated Islet

Distribution Program (IIDP) (RRID: SCR_014387) City of Hope, NIH grant #2UC4DK098085; the Scientific Computing of IT Division, Charité – Universitätsmedizin Berlin for providing computational resources; support by de.NBI Cloud, German Network for Bioinformatics Infrastructure (de.NBI), and ELIXIR-DE (Forschungszentrum Jülich and W-de.NBI-001, W-de.NBI-004, W-de.NBI-008, W-de.NBI-010, W-de.NBI-013, W-de.NBI-014, W-de.NBI-016, and W-de.NBI-022).

We thank the Sequencing Facility, Max Planck Institute for Molecular Genetics (MPIMG), for sample sequencing and Charité – Universitätsmedizin Berlin's Research Data Management (Forschungsdatenmanagement) for providing and supporting the OMERO platform.

We are grateful to Prof. Wilko Weichert, who sadly passed away during the preparation of this manuscript, for his support and contributions.

Some figures and panels were created using BioRender.

Funding: this work was funded by the European Commission (EU Horizon2020 874710) and supported by additional funding sources as follows. E.M.: Ramón y Cajal fellowship (RYC2021-032359-I), the Spanish Ministry of Science, and the Catalan Agency for Management of University and Research Grants (AGAUR, 2021 SGR 01586). D.B.: EMBO Long-Term Fellowship (ALT295-2019); 2022 EFSD Rising Star Programme; Ramón y Cajal fellowship (RYC2021-033131-I), Spanish Ministry of Science; Early Career Researcher grant, Sigrid Jusélius Foundation (003701165704); Academy Research Fellowship 2024, Research Council of Finland (361593); and Excellence Emerging Investigator Grant, Novo Nordisk Foundation (NNF24OC0089232). J.F. was funded by the AXA Chair in risk prediction in age-related diseases; the European Research Council Advanced Grant (789055); the Spanish Ministry of Science Innovation and Universities (PID2021-122522OB-I00 and PID2024-157322OB-I00); and the Merck Sharpe & Dohme (MSD) European Foundation for the Study of Diabetes Award. F.M. is funded by a Marie Skłodowska-Curie postdoctoral fellowship (n101152820).

AUTHOR CONTRIBUTIONS

Conceptualization: E.M., D.B., L.T., F.C., E.d.K., J.F., B.G., H.H., E.L., L.P., K.S., A.v.O., W.W., C.C., and R.E.; software: J.L., F.W.T., S.T., and A.S.; formal analysis: E.M., D.B., J.L., A.G.-H., A.M.C., Ma.Ma, F.M., F.S., L.T., V.V., E.B., N.G., P.I., K.J., A.S., F.W.T., and T.T.; investigation: E.M., D.B., J.L., A.G.-H., A.M.C., Ma.Ma, F.S., L.T., E.B., F.B., R.L.C., M.E., J.G.-H., M.H., K.J., B.K., A.S., T.T., M.v.A., D.M., and F.B.N.; resources: E.M., D.B., D.A., C.L., Ma.Mu, S.P., V.S., F.C., E.d.K., J.F., B.G., H.H., E.L., L.P., K.S., A.v.O., W.W., C.C., and R.E.; data curation: E.M., D.B., J.L., A.G.-H., A.M.C., U.T., and S.T.; visualization: E.M., D.B., J.L., A.G.-H., A.M.C., and Ma.Ma; supervision: E.M., D.B., F.C., E.d.K., J.F., B.G., H.H., E.L., L.P., K.S., A.v.O., W.W., C.C., and R.E.; writing – original draft: E.M., D.B., J.L., A.G.-H., A.M.C., and Ma. Ma; writing – review & editing: E.M., D.B., J.L., A.G.-H., A.M.C., Ma. Ma, F.M., F.S., L.T., V.V., K.J., A.S., F.W.T., R.L.C., F.C., J.F., B.G., L.P., A.v.O., C.C., and R.E.; project administration: C.C. and R.E.; funding acquisition: F.C., E.d.K., J.F., B.G., H.H., E.L., L.P., K.S., A.v.O., W.W., C.C., and R.E.

DECLARATION OF INTERESTS

H.H. is a co-founder and CSO of Omniscope; a scientific advisory board member of Nanostring, Bruker, and MiRXES; and a consultant at Moderna and Singularity. H.H. also received an honorarium from Genentech. M.M. is a co-founder and CEO of Single Cell Discoveries. A.v.O. is a Single Cell Discoveries advisory board member.

STAR★METHODS

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

- **KEY RESOURCES TABLE**
- **EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS**
 - Sample procurement of adult pancreas tissue
 - Pancreatic tissue and isolated islets from LUMC
 - Fetal tissue LUMC
- **METHOD DETAILS**

- Culture of primary human islets
- In vitro human islet glucose stimulated insulin secretion
- Single cell/nuclei preparation
- Single cell/nuclei library preparation and sequencing
- Single cell/nuclei data pre-processing
- Single cell/nuclei adult, islets and fetal pancreas snRNA ambient RNA removal
- Single cell/nuclei adult, islets and fetal pancreas snRNA processing
- Single nuclei doublet detection and removal in snRNA
- Single cell/nuclei integration
- Single cell/nuclei annotation of adult, islets and fetal snRNA data
- snATAC processing
- snATAC annotation of adult whole pancreas data
- snATAC doublet detection and removal
- snATAC peak calling
- snATAC and scRNA integration of adult whole pancreas data
- snATAC and scRNA label transferring between the adult whole pancreas datasets
- snATAC to scRNA label transfer from the adult no diabetes endocrine sample
- Differential gene expression, gene activity and peak accessibility testing
- Converting objects between Seurat and Scanpy
- snATAC trajectory analysis for the adult whole pancreas data
- snATAC and snRNA adult pancreatic network
- Multiome sample processing, merging and annotation
- Pseudotime alignment between snATAC and snRNA data
- Multiome cell type annotation
- Multiome doublet identification
- Conserved markers analysis between the ESPACE and Tosti et al. datasets
- Meta-analysis for consensus endocrine marker identification across sn/scRNA-seq datasets
- MILO analysis
- Xenium in situ sample preparation and sequencing
- Xenium data processing
- Xenium H&E image alignment
- Xenium cell type assignment
- Xenium pCAC identification
- Xenium ionocyte identification
- Xenium multipotency score
- Xenium islet composition analysis
- CODEX Immunohistochemistry staining & slide scanning
- CODEX Indirect immunofluorescence staining & antibody conjugation
- CODEX non-disease adult and fetal highly multiplexed staining
- CODEX Adult - image acquisition & computational image processing
- CODEX Adult - data cleaning, clustering & cell annotation
- CODEX Fetal - image acquisition & computational image processing
- CODEX Fetal - data cleaning, clustering & cell annotation
- Immunofluorescence and immunohistochemistry of adult pancreas tissue for protein level marker validation

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.cmet.2026.07.023>.

Received: October 23, 2025

Revised: May 15, 2026

Accepted: July 24, 2026

Published: August 12, 2026

REFERENCES

1. Storz, P. (2017). Acinar cell plasticity and development of pancreatic ductal adenocarcinoma. *Nat. Rev. Gastroenterol. Hepatol.* **14**, 296–304. <https://doi.org/10.1038/nrgastro.2017.12>.

2. Lodestijn, S.C., van den Bosch, T., Nijman, L.E., Moreno, L.F., Schlingemann, S., Sheraton, V.M., van Neerven, S.M., Koning, J.J., Vieira Braga, F.A., Paauw, N.J., et al. (2021). Continuous clonal labeling reveals uniform progenitor potential in the adult exocrine pancreas. *Cell Stem Cell* 28, 2009–2019.e4. <https://doi.org/10.1016/j.stem.2021.07.004>.
3. Wolny, D., Zhao, S., Everlien, I., Lun, X., Brunken, J., Brüne, D., Ziebell, F., Tabansky, I., Weichert, W., Marciniak-Czochra, A., et al. (2016). Single-Cell Analysis Uncovers Clonal Acinar Cell Heterogeneity in the Adult Pancreas. *Dev. Cell* 39, 289–301. <https://doi.org/10.1016/j.devcel.2016.10.002>.
4. Dor, Y., Brown, J., Martinez, O.I., and Melton, D.A. (2004). Adult pancreatic beta-cells are formed by self-duplication rather than stem-cell differentiation. *Nature* 429, 41–46. <https://doi.org/10.1038/nature02520>.
5. Wilson, M.E., Scheel, D., and German, M.S. (2003). Gene expression cascades in pancreatic development. *Mech. Dev.* 120, 65–80. [https://doi.org/10.1016/s0925-4773\(02\)00333-7](https://doi.org/10.1016/s0925-4773(02)00333-7).
6. Arda, H.E., Benitez, C.M., and Kim, S.K. (2013). Gene Regulatory Networks Governing Pancreas Development. *Dev. Cell* 25, 5–13. <https://doi.org/10.1016/j.devcel.2013.03.016>.
7. Stanger, B.Z., and Hebrok, M. (2013). Control of Cell Identity in Pancreas Development and Regeneration. *Gastroenterology* 144, 1170–1179. <https://doi.org/10.1053/j.gastro.2013.01.074>.
8. Zhou, Q., Law, A.C., Rajagopal, J., Anderson, W.J., Gray, P.A., and Melton, D.A. (2007). A Multipotent Progenitor Domain Guides Pancreatic Organogenesis. *Dev. Cell* 13, 103–114. <https://doi.org/10.1016/j.devcel.2007.06.001>.
9. Baeyens, L., Bonné, S., Bos, T., Rooman, I., Peleman, C., Lahoutte, T., German, M., Heimberg, H., and Bouwens, L. (2009). Notch Signaling as Gatekeeper of Rat Acinar-to- β -Cell Conversion in Vitro. *Gastroenterology* 136, 1750–60.e13. <https://doi.org/10.1053/j.gastro.2009.01.047>.
10. Burdziak, A.C., Alonso-Curbelo, D., Walle, T., and Francisco, M. (2022). Epigenetic plasticity cooperates with emergent cell-cell interactions to drive neoplastic tissue remodeling in the pancreas. <https://doi.org/10.1101/2022.07.26.501417>.
11. 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. <https://doi.org/10.1016/j.cels.2016.09.002>.
12. Migliorini, A., Ge, S., Atkins, M.H., Oakie, A., Sambathkumar, R., Kent, G., Huang, H., Sing, A., Chua, C., Gehring, A.J., et al. (2024). Embryonic macrophages support endocrine commitment during human pancreatic differentiation. *Cell Stem Cell* 37, 1591–1611.e8. <https://doi.org/10.1016/j.stem.2024.09.011>.
13. Tosti, L., Hang, Y., Debnath, O., Tiesmeyer, S., Trefzer, T., Steiger, K., Ten, F.W., Lukassen, S., Ballik, S., Kühl, A.A., et al. (2021). Single-Nucleus and In Situ RNA-Sequencing Reveal Cell Topographies in the Human Pancreas. *Gastroenterology* 160, 1330–1344.e11. <https://doi.org/10.1053/j.gastro.2020.11.010>.
14. Chiou, J., Zeng, C., Cheng, Z., Han, J.Y., Schlichting, M., Miller, M., Mendez, R., Huang, S., Wang, J., Sui, Y., et al. (2021). Single-cell chromatin accessibility identifies pancreatic islet cell type- and state-specific regulatory programs of diabetes risk. *Nat. Genet.* 53, 455–466. <https://doi.org/10.1038/s41588-021-00823-0>.
15. van Gurp, L., Fodoulan, L., Oropeza, D., Furuyama, K., Bru-Tari, E., Vu, A.N., Kaddis, J.S., Rodríguez, I., Thorel, F., and Herrera, P.L. (2022). Generation of human islet cell type-specific identity genesets. *Nat. Commun.* 13, 2020. <https://doi.org/10.1038/s41467-022-29588-8>.
16. Puri, C., Vicinanza, M., Ashkenazi, A., Gratian, M.J., Zhang, Q., Bento, C.F., Renna, M., Menzies, F.M., and Rubinsztein, D.C. (2018). The RAB11A-Positive Compartment Is a Primary Platform for Autophagosome Assembly Mediated by WIPI2 Recognition of PI3P-RAB11A. *Dev. Cell* 45, 114–131.e8. <https://doi.org/10.1016/j.devcel.2018.03.008>.
17. Dhawan, S., Georgia, S., Tschen, S.I., Fan, G., and Bhushan, A. (2011). Pancreatic β Cell Identity Is Maintained by DNA Methylation-Mediated Repression of *Arx*. *Dev. Cell* 20, 419–429. <https://doi.org/10.1016/j.devcel.2011.03.012>.
18. Wang, L., Baek, S., Prasad, G., Wildenthal, J., Guo, K., Sturgill, D., Truongvo, T., Char, E., Pegoraro, G., McKinnon, K., et al. (2026). Predictive prioritization of enhancers associated with pancreatic disease risk. *Cell Genomics* 6, 101040. <https://doi.org/10.1016/j.xgen.2025.101040>.
19. Kaestner, K.H., Powers, A.C., Naji, A.; HPAP Consortium, and Atkinson, M.A. (2019). NIH Initiative to Improve Understanding of the Pancreas, Islet, and Autoimmunity in Type 1 Diabetes: The Human Pancreas Analysis Program (HPAP). *Diabetes* 68, 1394–1402. <https://doi.org/10.2337/db19-0058>.
20. Tosti, L., and Conrad, C. (2020) Nuclei Isolation from Snap Frozen Human Pancreatic Tissue Using a Citric Acid Buffer. <https://doi.org/10.17504/protocols.io.tpxempn>.
21. Guerra, C., Schuhmacher, A.J., Cañamero, M., Grippo, P.J., Verdaguer, L., Pérez-Gallego, L., Dubus, P., Sandgren, E.P., and Barbacid, M. (2007). Chronic pancreatitis is essential for induction of pancreatic ductal adenocarcinoma by K-Ras oncogenes in adult mice. *Cancer Cell* 11, 291–302. <https://doi.org/10.1016/j.ccr.2007.01.012>.
22. Chen, H., Albergante, L., Hsu, J.Y., Lareau, C.A., Lo Bosco, G., Guan, J., Zhou, S., Gorban, A.N., Bauer, D.E., Aryee, M.J., et al. (2019). Single-cell trajectories reconstruction, exploration and mapping of omics data with STREAM. *Nat. Commun.* 10, 1903. <https://doi.org/10.1038/s41467-019-09670-4>.
23. Neuß, T., Wirges, N., Chen, M.-C., Usluer, S., Oellinger, R., Lier, S., Dudek, M., Madl, T., Jastroch, M., Steiger, K., et al. (2022). Acinar-to-ductal metaplasia in the pancreas requires a glycolytic switch and functional mitochondria. Preprint at bioRxiv. <https://doi.org/10.1101/2022.06.27.495427>.
24. Wang, V.M.-Y., Ferreira, R.M.M., Almagro, J., Evan, T., Legrave, N., Zaw Thin, M., Frith, D., Carvalho, J., Barry, D.J., Snijders, A.P., et al. (2019). CD9 identifies pancreatic cancer stem cells and modulates glutamine metabolism to fuel tumour growth. *Nat. Cell Biol.* 21, 1425–1435. <https://doi.org/10.1038/s41556-019-0407-1>.
25. Friedmann Angeli, J.P., Schneider, M., Proneth, B., Tyurina, Y.Y., Tyurin, V.A., Hammond, V.J., Herbach, N., Aichler, M., Walch, A., Eggenhofer, E., et al. (2014). Inactivation of the ferroptosis regulator Gpx4 triggers acute renal failure in mice. *Nat. Cell Biol.* 16, 1180–1191. <https://doi.org/10.1038/ncb3064>.
26. Means, A.L., Meszoely, I.M., Suzuki, K., Miyamoto, Y., Rustgi, A.K., Coffey, R.J., Wright, C.V.E., Stoffers, D.A., and Leach, S.D. (2005). Pancreatic epithelial plasticity mediated by acinar cell transdifferentiation and generation of nestin-positive intermediates. *Development* 132, 3767–3776. <https://doi.org/10.1242/dev.01925>.
27. Jiang, Y., Liao, C., Lai, J., Peng, Y., Chen, Q., and Zheng, X. (2025). KRT7 promotes pancreatic cancer metastasis by remodeling the extracellular matrix niche through FGF2-fibroblast crosstalk. *Sci. Rep.* 15, 6951. <https://doi.org/10.1038/s41598-024-84129-1>.
28. Aney, K.J., Jeong, W.-J., Vallejo, A.F., Burdziak, C., Chen, E., Wang, A., Koak, P., Wise, K., Jensen, K., Pe'er, D., et al. (2024). Novel Approach for Pancreas Transcriptomics Reveals the Cellular Landscape in Homeostasis and Acute Pancreatitis. *Gastroenterology* 166, 1100–1113. <https://doi.org/10.1053/j.gastro.2024.01.043>.
29. Grimont, A., Leach, S.D., and Chandwani, R. (2022). Uncertain Beginnings: Acinar and Ductal Cell Plasticity in the Development of Pancreatic Cancer. *Cell. Mol. Gastroenterol. Hepatol.* 13, 369–382. <https://doi.org/10.1016/j.jcmgh.2021.07.014>.
30. Zhang, H., Corredor, A.L.G., Messina-Pacheco, J., Li, Q., Zogopoulos, G., Kaddour, N., Wang, Y., Shi, B.Y., Gregorieff, A., Liu, J.L., et al. (2021). REG3A/REG3B promotes acinar to ductal metaplasia through binding to EXTL3 and activating the RAS-RAF-MEK-ERK signaling pathway. *Commun. Biol.* 4, 688. <https://doi.org/10.1038/s42003-021-02193-z>.
31. Tritschler, S., Thomas, M., Böttcher, A., Ludwig, B., Schmid, J., Schubert, U., Kemter, E., Wolf, E., Lickert, H., and Theis, F.J. (2022). A

- transcriptional cross species map of pancreatic islet cells. *Mol. Metab.* 66, 101595. <https://doi.org/10.1016/j.molmet.2022.101595>.
32. Farack, L., Golan, M., Egozi, A., Dezorella, N., Bahar Halpern, K., Ben-Moshe, S., Garzilli, I., Tóth, B., Roitman, L., Krizhanovsky, V., et al. (2019). Transcriptional Heterogeneity of Beta Cells in the Intact Pancreas. *Dev. Cell* 48, 115–125.e4. <https://doi.org/10.1016/j.devcel.2018.11.001>.
33. Li, J., Klughammer, J., Farlik, M., Penz, T., Spittler, A., Barbieux, C., Berishvili, E., Bock, C., and Kubicek, S. (2016). Single-cell transcriptomes reveal characteristic features of human pancreatic islet cell types. *EMBO Rep.* 17, 178–187. <https://doi.org/10.15252/embr.201540946>.
34. Elgamal, R.M., Kudtarkar, P., Melton, R.L., Mummey, H.M., Benaglio, P., Okino, M.-L., and Gaulton, K.J. (2023). An Integrated Map of Cell Type-Specific Gene Expression in Pancreatic Islets. *Diabetes* 72, 1719–1728. <https://doi.org/10.2337/db23-0130>.
35. 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. <https://doi.org/10.1016/j.cmet.2016.08.020>.
36. 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. <https://doi.org/10.1016/j.cels.2016.08.011>.
37. Camunas-Soler, J., Dai, X.-Q., Hang, Y., Bautista, A., Lyon, J., Suzuki, K., Kim, S.K., Quake, S.R., and MacDonald, P.E. (2020). Patch-Seq Links Single-Cell Transcriptomes to Human Islet Dysfunction in Diabetes. *Cell Metab.* 31, 1017–1031.e4. <https://doi.org/10.1016/j.cmet.2020.04.005>.
38. Benninger, R.K.P., and Kravets, V. (2022). The physiological role of β -cell heterogeneity in pancreatic islet function. *Nat. Rev. Endocrinol.* 18, 9–22. <https://doi.org/10.1038/s41574-021-00568-0>.
39. Bonnycastle, L.L., Gildea, D.E., Yan, T., Narisu, N., Swift, A.J., Wolfsberg, T.G., Erdos, M.R., and Collins, F.S. (2020). Single-cell transcriptomics from human pancreatic islets: sample preparation matters. *Biol. Methods Protoc.* 5, bpz019. <https://doi.org/10.1093/biomethods/bpz019>.
40. Thiel, G., Müller, I., and Rössler, O.G. (2013). Signal transduction via TRPM3 channels in pancreatic β -cells. *J. Mol. Endocrinol.* 50, R75–R83. <https://doi.org/10.1530/JME-12-0237>.
41. Scoville, D.W., Licht-Kaiser, K., Grimm, S.A., and Jetten, A.M. (2019). GLIS3 binds pancreatic beta cell regulatory regions alongside other islet transcription factors. *J. Endocrinol.* 243, 1–14. <https://doi.org/10.1530/JOE-19-0182>.
42. Kang, R.B., Li, Y., Rosselot, C., Zhang, T., Siddiq, M., Rajbhandari, P., Stewart, A.F., Scott, D.K., Garcia-Ocana, A., and Lu, G. (2023). Single-nucleus RNA sequencing of human pancreatic islets identifies novel gene sets and distinguishes β -cell subpopulations with dynamic transcriptome profiles. *Genome Med.* 15, 30. <https://doi.org/10.1186/s13073-023-01179-2>.
43. Kattner, N., Hang, Y., Krentz, N.A.J., Russell, L.A., Palmer, M., Flaxman, C., Plett, N., Coulthard, R., Al-Selwi, Y., Dyson, N., et al. (2025). Identification of a vimentin-expressing α -cell phenotype in CF and normal pancreas. *J. Endocrinol.* 264, e240190. <https://doi.org/10.1530/JOE-24-0190>.
44. Bouwens, L., and Pipeleers, D.G. (1998). Extra-insular beta cells associated with ductules are frequent in adult human pancreas. *Diabetologia* 41, 629–633. <https://doi.org/10.1007/s001250050960>.
45. Lehrstrand, J., Davies, W.I.L., Hahn, M., Korsgren, O., Alanentalo, T., and Ahlgren, U. (2024). Illuminating the complete β -cell mass of the human pancreas—signifying a new view on the islets of Langerhans. *Nat. Commun.* 15, 3318. <https://doi.org/10.1038/s41467-024-47686-7>.
46. Bonner-Weir, S., Guo, L., Li, W.-C., Ouziel-Yahalom, L., Lysy, P.A., Weir, G.C., and Sharma, A. (2012). Islet Neogenesis: A Possible Pathway for Beta-Cell Replenishment. *Rev. Diabet. Stud.* 9, 407–416. <https://doi.org/10.1900/RDS.2012.9.407>.
47. Liang, R., Liu, T., Zhang, L., Ma, W., Ren, H., and Wang, S. (2026). Pancreatic Duct Cells as a Potential Source for Human Islet Neogenesis: Insights from Imaging Mass Cytometry. Preprint at bioRxiv. <https://doi.org/10.64898/2026.02.07.704527>.
48. Murrall, K., Luckett, T., Lekka, C., Flaxman, C.S., Wyatt, R., Akhbari, P., Kusmartseva, I., Hunter, S.L., Leete, P., Burn, I., et al. (2025). Small things matter: Lack of extraslet β cells in type 1 diabetes. *Sci. Adv.* 11, ead2251. <https://doi.org/10.1126/sciadv.adz2251>.
49. Mardamshina, M., Dorka, N., Thielert, M., Björklund, F., Navarro, F.B., Casals, A.M., Kovacs, F., Manning Fox, J.E.M., Horvath, P., MacDonald, P.E., et al. (2025). Multiplexed Deep Visual Proteomics Unveils Spatial Heterogeneity and Rare Endocrine States in Human Adult Pancreatic Islets. Preprint at bioRxiv. <https://doi.org/10.1101/2025.04.27.650857>.
50. Bernardo, E., De Vas, M.G., Balboa, D., Cuenca-Ardura, M., Bonàs-Guarch, S., Planas-Fèlix, M., Mollandin, F., Torrens-Dinarès, M., Maestro, M.A., García-Hurtado, J., et al. (2025). HNF1A and A1CF coordinate a beta cell transcription-splicing axis that is disrupted in type 2 diabetes. *Cell Metab.* 37, 1870–1889.e10. <https://doi.org/10.1016/j.cmet.2025.07.007>.
51. Weng, C., Gu, A., Zhang, S., Lu, L., Ke, L., Gao, P., Liu, X., Wang, Y., Hu, P., Plummer, D., et al. (2023). Single cell multiomic analysis reveals diabetes-associated β -cell heterogeneity driven by HNF1A. *Nat. Commun.* 14, 5400. <https://doi.org/10.1038/s41467-023-41228-3>.
52. Wang, G., Chiou, J., Zeng, C., Miller, M., Matta, I., Han, J.Y., Kadakia, N., Okino, M.-L., Beebe, E., Mallick, M., et al. (2023). Integrating genetics with single-cell multiomic measurements across disease states identifies mechanisms of beta cell dysfunction in type 2 diabetes. *Nat. Genet.* 55, 984–994. <https://doi.org/10.1038/s41588-023-01397-9>.
53. Acera-Mateos, M., Adiconis, X., Li, J.-K., Marchese, D., Caratù, G., Hon, C.-C., Tiwari, P., Kojima, M., Vieth, B., Murphy, M.A., et al. (2026). Systematic evaluation of single-cell multimodal data integration enhances cell type resolution and discovery of clinically relevant states in complex tissues. *Genome Biol.* 27, 64. <https://doi.org/10.1186/s13059-026-04002-4>.
54. Dann, E., Henderson, N.C., Teichmann, S.A., Morgan, M.D., and Marioni, J.C. (2022). Differential abundance testing on single-cell data using k-nearest neighbor graphs. *Nat. Biotechnol.* 40, 245–253. <https://doi.org/10.1038/s41587-021-01033-z>.
55. Xin, Y., Dominguez Gutierrez, G., Okamoto, H., Kim, J., Lee, A.-H., Adler, C., Ni, M., Yancopoulos, G.D., Murphy, A.J., and Gromada, J. (2018). Pseudotime Ordering of Single Human β -Cells Reveals States of Insulin Production and Unfolded Protein Response. *Diabetes* 67, 1783–1794. <https://doi.org/10.2337/db18-0365>.
56. Viñuela, A., Varshney, A., van de Bunt, M., Prasad, R.B., Asplund, O., Bennett, A., Boehnke, M., Brown, A.A., Erdos, M.R., Fadista, J., et al. (2020). Genetic variant effects on gene expression in human pancreatic islets and their implications for T2D. *Nat. Commun.* 11, 4912. <https://doi.org/10.1038/s41467-020-18581-8>.
57. Boesgaard, T.W., Grarup, N., Jørgensen, T., Borch-Johnsen, K., Meta-Analysis of Glucose and Insulin-Related Trait Consortium (MAGIC), Hansen, T., and Pedersen, O. (2010). Variants at DGKB/TMEM195, ADRA2A, GLIS3 and C2CD4B loci are associated with reduced glucose-stimulated beta cell function in middle-aged Danish people. *Diabetologia* 53, 1647–1655. <https://doi.org/10.1007/s00125-010-1753-5>.
58. Fueger, P.T., Schisler, J.C., Lu, D., Babu, D.A., Mirmira, R.G., Newgard, C.B., and Hohmeier, H.E. (2008). Trefoil Factor 3 Stimulates Human and Rodent Pancreatic Islet β -Cell Replication with Retention of Function. *Mol. Endocrinol.* 22, 1251–1259. <https://doi.org/10.1210/me.2007-0500>.
59. Gradwohl, G., Dierich, A., LeMeur, M., and Guillemot, F. (2000). neurogenin3 is required for the development of the four endocrine cell lineages of the pancreas. *Proc. Natl. Acad. Sci. USA* 97, 1607–1611. <https://doi.org/10.1073/pnas.97.4.1607>.
60. McGrath, P.S., Watson, C.L., Ingram, C., Helmrath, M.A., and Wells, J.M. (2015). The Basic Helix-Loop-Helix Transcription Factor NEUROG3 Is

Required for Development of the Human Endocrine Pancreas. *Diabetes* 64, 2497–2505. <https://doi.org/10.2337/db14-1412>.

61. Smith, S.B., Qu, H.-Q., Taleb, N., Kishimoto, N.Y., Scheel, D.W., Lu, Y., Patch, A.-M., Grabs, R., Wang, J., Lynn, F.C., et al. (2010). Rfx6 directs islet formation and insulin production in mice and humans. *Nature* 463, 775–780. <https://doi.org/10.1038/nature08748>.
62. Petersen, M.B.K., Azad, A., Ingvorsen, C., Hess, K., Hansson, M., Grapin-Botton, A., and Honoré, C. (2017). Single-Cell Gene Expression Analysis of a Human ESC Model of Pancreatic Endocrine Development Reveals Different Paths to β -Cell Differentiation. *Stem Cell Rep.* 9, 1246–1261. <https://doi.org/10.1016/j.stemcr.2017.08.009>.
63. Ramond, C., Beydag-Tasöz, B.S., Azad, A., Van De Bunt, M., Petersen, M.B.K., Beer, N.L., Glaser, N., Berthault, C., Gloyn, A.L., Hansson, M., et al. (2018). Understanding human fetal pancreas development using subpopulation sorting, RNA sequencing and single-cell profiling. *Dev.* 145, dev165480. <https://doi.org/10.1242/dev.165480>.
64. Bechard, M.E., Bankatis, E.D., Hipkens, S.B., Ustione, A., Piston, D.W., Yang, Y.-P., Magnuson, M.A., and Wright, C.V.E. (2016). Precommitment low-level Neurog3 expression defines a long-lived mitotic endocrine-biased progenitor pool that drives production of endocrine-committed cells. *Genes Dev.* 30, 1852–1865. <https://doi.org/10.1101/gad.284729.116>.
65. Ma, Z., Zhang, X., Zhong, W., Yi, H., Chen, X., Zhao, Y., Ma, Y., Song, E., and Xu, T. (2023). Deciphering early human pancreas development at the single-cell level. *Nat. Commun.* 14, 5354. <https://doi.org/10.1038/s41467-023-40893-8>.
66. Olaniru, O.E., Kadlinsky, U., Kannambath, S., Vaikinen, H., Fung, K., Dhami, P., and Persaud, S.J. (2023). Single-cell transcriptomic and spatial landscapes of the developing human pancreas. *Cell Metab.* 35, 184–199.e5. <https://doi.org/10.1016/j.cmet.2022.11.009>.
67. Cogger, K.F., Sinha, A., Sarangi, F., McGaugh, E.C., Saunders, D., Dorrell, C., Mejia-Guerrero, S., Aghazadeh, Y., Rourke, J.L., Sreanot, R.A., et al. (2017). Glycoprotein 2 is a specific cell surface marker of human pancreatic progenitors. *Nat. Commun.* 8, 331. <https://doi.org/10.1038/s41467-017-00561-0>.
68. Desgraz, R., and Herrera, P.L. (2009). Pancreatic neurogenin 3-expressing cells are unipotent islet precursors. *Development* 136, 3567–3574. <https://doi.org/10.1242/dev.039214>.
69. Seymour, P.A., Collin, C.A., Egeskov-Madsen, A. la R., Jørgensen, M.C., Shimojo, H., Imayoshi, I., de Lichtenberg, K.H., Kopan, R., Kageyama, R., and Serup, P. (2020). Jag1 Modulates an Oscillatory Dll1-Notch-Hes1 Signaling Module to Coordinate Growth and Fate of Pancreatic Progenitors. *Dev. Cell* 52, 731–747.e8. <https://doi.org/10.1016/j.devcel.2020.01.015>.
70. Xu, X., Seymour, P.A., Sneppen, K., Trusina, A., Egeskov-Madsen, A. la R., Jørgensen, M.C., Jensen, M.H., and Serup, P. (2023). Jag1-Notch cis-interaction determines cell fate segregation in pancreatic development. *Nat. Commun.* 14, 348. <https://doi.org/10.1038/s41467-023-35963-w>.
71. Qu, X., Afelik, S., Jensen, J.N., Bukys, M.A., Kobberup, S., Schmerr, M., Xiao, F., Nyeng, P., Veronica Albertoni, M., Grapin-Botton, A., et al. (2013). Notch-mediated post-translational control of Ngn3 protein stability regulates pancreatic patterning and cell fate commitment. *Dev. Biol.* 376, 1–12. <https://doi.org/10.1016/j.ydbio.2013.01.021>.
72. Cebola, I., Rodríguez-Seguí, S.A., Cho, C.H.-H., Bessa, J., Rovira, M., Luengo, M., Chhatrivala, M., Berry, A., Ponsa-Cobas, J., Maestro, M.A., et al. (2015). TEAD and YAP regulate the enhancer network of human embryonic pancreatic progenitors. *Nat. Cell Biol.* 17, 615–626. <https://doi.org/10.1038/ncb3160>.
73. Miguel-Escalada, I., Maestro, M.Á., Balboa, D., Elek, A., Bernal, A., Bernardo, E., Grau, V., García-Hurtado, J., Sebé-Pedrós, A., and Ferrer, J. (2022). Pancreas agenesis mutations disrupt a lead enhancer controlling a developmental enhancer cluster. *Dev. Cell* 57, 1922–1936.e9. <https://doi.org/10.1016/j.devcel.2022.07.014>.
74. Gonçalves, C.A., Larsen, M., Jung, S., Stratmann, J., Nakamura, A., Leuschner, M., Hersemann, L., Keshara, R., Perlman, S., Lundvall, L., et al. (2021). A 3D system to model human pancreas development and its reference single-cell transcriptome atlas identify signaling pathways required for progenitor expansion. *Nat. Commun.* 12, 3144. <https://doi.org/10.1038/s41467-021-23295-6>.
75. Carpino, G., Renzi, A., Cardinale, V., Franchitto, A., Onori, P., Overi, D., Rossi, M., Berloco, P.B., Alvaro, D., Reid, L.M., et al. (2016). Progenitor cell niches in the human pancreatic duct system and associated pancreatic duct glands: an anatomical and immunophenotyping study. *J. Anat.* 228, 474–486. <https://doi.org/10.1111/joa.12418>.
76. Martín-Pagola, A., Sisino, G., Allende, G., Domínguez-Bendala, J., Gianani, R., Reijonen, H., Nepom, G.T., Ricordi, C., Ruiz, P., Sageshima, J., et al. (2008). Insulin protein and proliferation in ductal cells in the transplanted pancreas of patients with type 1 diabetes and recurrence of autoimmunity. *Diabetologia* 51, 1803–1813. <https://doi.org/10.1007/s00125-008-1105-x>.
77. Qadir, M.M.F., Álvarez-Cubela, S., Klein, D., van Dijk, J., Muñoz-Anquela, R., Moreno-Hernández, Y.B., Lanzoni, G., Sadiq, S., Navarro-Rubio, B., García, M.T., et al. (2020). Single-cell resolution analysis of the human pancreatic ductal progenitor cell niche. *Proc. Natl. Acad. Sci. USA* 117, 10876–10887. <https://doi.org/10.1073/pnas.1918314117>.
78. Solar, M., Cardalda, C., Houbracken, I., Martín, M., Maestro, M.A., De Medts, N., Xu, X., Grau, V., Heimberg, H., Bouwens, L., et al. (2009). Pancreatic exocrine duct cells give rise to insulin-producing beta cells during embryogenesis but not after birth. *Dev. Cell* 17, 849–860. <https://doi.org/10.1016/j.devcel.2009.11.003>.
79. Rankin, M.M., Wilbur, C.J., Rak, K., Shields, E.J., Granger, A., and Kushner, J.A. (2013). β -Cells are not generated in pancreatic duct ligation-induced injury in adult mice. *Diabetes* 62, 1634–1645. <https://doi.org/10.2337/db12-0848>.
80. Zhao, H., Huang, X., Liu, Z., Pu, W., Lv, Z., He, L., Li, Y., Zhou, Q., Lui, K.O., and Zhou, B. (2021). Pre-existing beta cells but not progenitors contribute to new beta cells in the adult pancreas. *Nat. Metab.* 3, 352–365. <https://doi.org/10.1038/s42255-021-00364-0>.
81. Perez-Frances, M., Abate, M.V., Baronnier, D., Scherer, P.E., Fujitani, Y., Thorel, F., and Herrera, P.L. (2022). Adult pancreatic islet endocrine cells emerge as fetal hormone-expressing cells. *Cell Rep.* 38, 110377. <https://doi.org/10.1016/j.celrep.2022.110377>.
82. Fleming, S.J., Chaffin, M.D., Arduini, A., Akkad, A.-D., Banks, E., Marioni, J.C., Philippakis, A.A., Ellinor, P.T., and Babadi, M. (2023). Unsupervised removal of systematic background noise from droplet-based single-cell experiments using CellBender. *Nat. Methods* 20, 1323–1335. <https://doi.org/10.1038/s41592-023-01943-7>.
83. Cargnelli, C.B., Nielsen, J.V., and Skat Madsen, J.G.S. (2026). Benchmarking computational decontamination of ambient RNA. Preprint at bioRxiv. <https://doi.org/10.64898/2026.01.13.699237>.
84. Janssen, P., Kliesmete, Z., Vieth, B., Adiconis, X., Simmons, S., Marshall, J., McCabe, C., Heyn, H., Levin, J.Z., Enard, W., et al. (2023). The effect of background noise and its removal on the analysis of single-cell expression data. *Genome Biol.* 24, 140. <https://doi.org/10.1186/s13059-023-02978-x>.
85. Slyper, M., Porter, C.B.M., Ashenberg, O., Waldman, J., Drokhlyansky, E., Wakiro, I., Smillie, C., Smith-Rosario, G., Wu, J., Dionne, D., et al. (2020). A single-cell and single-nucleus RNA-Seq toolbox for fresh and frozen human tumors. *Nat. Med.* 26, 792–802. <https://doi.org/10.1038/s41591-020-0844-1>.
86. Janesick, A., Shelansky, R., Gottschö, A.D., Wagner, F., Williams, S.R., Rouault, M., Beliakoff, G., Morrison, C.A., Oliveira, M.F., Sicherman, J.T., et al. (2023). High resolution mapping of the tumor microenvironment using integrated single-cell, spatial and in situ analysis. *Nat. Commun.* 14, 8353. <https://doi.org/10.1038/s41467-023-43458-x>.
87. Park, J., Choi, W., Tiesmeyer, S., Long, B., Borm, L.E., Garren, E., Nguyen, T.N., Tasic, B., Codeluppi, S., Graf, T., et al. (2021). Cell segmentation-free inference of cell types from in situ transcriptomics data. *Nat. Commun.* 12, 3545. <https://doi.org/10.1038/s41467-021-23807-4>.
88. Petukhov, V., Xu, R.J., Soldatov, R.A., Cadinu, P., Khodosevich, K., Moffitt, J.R., and Kharchenko, P.V. (2022). Cell segmentation in

- p>imaging-based spatial transcriptomics.
- Nat. Biotechnol.*
- 40, 345–354.
- <https://doi.org/10.1038/s41587-021-01044-w>
- .
89. Braxton, A.M., Kiemen, A.L., Grahm, M.P., Forjaz, A., Parksong, J., Mahesh Babu, J., Lai, J., Zheng, L., Niknafs, N., Jiang, L., et al. (2024). 3D genomic mapping reveals multifocality of human pancreatic precancers. *Nature* 629, 679–687. <https://doi.org/10.1038/s41586-024-07359-3>.
 90. Reisinger, E., Genthner, L., Kerssemakers, J., Kensche, P., Borufka, S., Jugold, A., Kling, A., Prinz, M., Scholz, I., Zipprich, G., et al. (2017). OTP: An automatized system for managing and processing NGS data. *J. Biotechnol.* 261, 53–62. <https://doi.org/10.1016/j.jbiotec.2017.08.006>.
 91. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. *Cell* 184, 3573–3587.e29. <https://doi.org/10.1016/j.cell.2021.04.048>.
 92. McGinnis, C.S., Murrow, L.M., and Gartner, Z.J. (2019). DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst.* 8, 329–337.e4. <https://doi.org/10.1016/j.cels.2019.03.003>.
 93. Wolock, S.L., Lopez, R., and Klein, A.M. (2019). Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. *Cell Syst.* 8, 281–291.e9. <https://doi.org/10.1016/j.cels.2018.11.005>.
 94. Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P.R., and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* 16, 1289–1296. <https://doi.org/10.1038/s41592-019-0619-0>.
 95. Stuart, T., Srivastava, A., Madad, S., Lareau, C.A., and Satija, R. (2021). Single-cell chromatin state analysis with Signac. *Nat. Methods* 18, 1333–1341. <https://doi.org/10.1038/s41592-021-01282-5>.
 96. Thibodeau, A., Eroglu, A., McGinnis, C.S., Lawlor, N., Nehar-Belaid, D., Kursawe, R., Marches, R., Conrad, D.N., Kuchel, G.A., Gartner, Z.J., et al. (2021). AMULET: a novel read count-based method for effective multiplet detection from single nucleus ATAC-seq data. *Genome Biol.* 22, 252. <https://doi.org/10.1186/s13059-021-02469-x>.
 97. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoutte, J., Johnson, D.S., Bernstein, B.E., Nussbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based Analysis of ChIP-Seq (MACS). *Genome Biol.* 9, R137. <https://doi.org/10.1186/gb-2008-9-9-r137>.
 98. Squire, J.W., Gautier, M., Kathe, C., Anderson, M.A., James, N.D., Hutson, T.H., Hudelle, R., Kaiser, T., Matson, K.J.E., Barraud, Q., et al. (2021). Confronting false discoveries in single-cell differential expression. *Nat. Commun.* 12, 5692. <https://doi.org/10.1038/s41467-021-25960-2>.
 99. Yang, K., Spitzer, H., Sterr, M., Hrovatin, K., de la O, S., Zhang, X., Setyono, E.S.A., Ud-Dean, M., Walzthoeni, T., Flisikowski, K., et al. (2025). A multimodal cross-species comparison of pancreas development. *Nat. Commun.* 16, 9355. <https://doi.org/10.1038/s41467-025-64774-4>.
 100. Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. <https://doi.org/10.1093/bioinformatics/btp616>.
 101. Bravo González-Blas, C., De Winter, S., Hulselmans, G., Hecker, N., Matetovici, I., Christiaens, V., Poovathingal, S., Wouters, J., Aibar, S., and Aerts, S. (2023). SCENIC+: single-cell multiomic inference of enhancers and gene regulatory networks. *Nat. Methods* 20, 1355–1367. <https://doi.org/10.1038/s41592-023-01938-4>.
 102. Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* 19, 15. <https://doi.org/10.1186/s13059-017-1382-0>.
 103. Palla, G., Spitzer, H., Klein, M., Fischer, D., Schaar, A.C., Kuemmerle, L.B., Rybakov, S., Ibarra, I.L., Holmberg, O., Virshup, I., et al. (2022). Squidpy: a scalable framework for spatial omics analysis. *Nat. Methods* 19, 171–178. <https://doi.org/10.1038/s41592-021-01358-2>.
 104. Marconato, L., Palla, G., Yamauchi, K.A., Virshup, I., Heidari, E., Treis, T., Vierdag, W.-M., Toth, M., Stockhaus, S., Shrestha, R.B., et al. (2025). SpatialData: an open and universal data framework for spatial omics. *Nat. Methods* 22, 58–62. <https://doi.org/10.1038/s41592-024-02212-x>.
 105. Harris, C.R., Millman, K.J., van der Walt, S.J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N.J., et al. (2020). Array programming with NumPy. *Nature* 585, 357–362. <https://doi.org/10.1038/s41586-020-2649-2>.
 106. Waskom, M.L. (2021). seaborn: statistical data visualization. *J. Open Source Softw.* 6, 3021. <https://doi.org/10.21105/joss.03021>.
 107. Hunter, J.D. (2007). Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* 9, 90–95. <https://doi.org/10.1109/MCSE.2007.55>.
 108. Ratamero, E.M., Govek, K., Langerak, J.M., Sanchez, F.C., and Mellert, D.J. (2023). Easing OMERO adoption with ezomero. Preprint at bioRxiv. <https://doi.org/10.1101/2023.06.29.546930>.
 109. van der Walt, S. van der, Schönberger, J.L., Nunez-Iglesias, J., Boulogne, F., Warner, J.D., Yager, N., Gouillart, E., and Yu, T.; scikit-image contributors (2014). scikit-image: image processing in Python. *PeerJ* 2, e453. <https://doi.org/10.7717/peerj.453>.
 110. Schmidt, U., Weigert, M., Broaddus, C., and Myers, G. (2018). Cell Detection with Star-convex Polygons. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, A.F. Frangi, J.A. Schnabel, C. Davatzikos, C. Alberola-López, and G. Fichtinger, eds. (Springer International Publishing), pp. 265–273. https://doi.org/10.1007/978-3-030-00934-2_30.
 111. Weigert, M., and Schmidt, U. (2022). Nuclei instance segmentation and classification in histopathology images with StarDist. In *IEEE International Symposium on Biomedical Imaging Challenges (ISBIC)*, pp. 1–4. <https://doi.org/10.1109/ISBIC56247.2022.9854534>.
 112. Weigert, M., Schmidt, U., Haase, R., Sugawara, K., and Myers, G. (2020). Star-convex Polyhedra for 3D Object Detection and Segmentation in Microscopy. In *IEEE Winter Conference on Applications of Computer Vision (WACV) (IEEE)*, pp. 3655–3662. <https://doi.org/10.1109/WACV45572.2020.9093435>.
 113. Blampey, Q., Mulder, K., Gardet, M., Christodoulidis, S., Dutertre, C.-A., André, F., Ginhoux, F., and Courmède, P.-H. (2024). Sopa: a technology-invariant pipeline for analyses of image-based spatial omics. *Nat. Commun.* 15, 4981. <https://doi.org/10.1038/s41467-024-48981-z>.
 114. Shih, H.P., Wang, A., and Sander, M. (2013). Pancreas Organogenesis: From Lineage Determination to Morphogenesis. *Annu. Rev. Cell Dev. Biol.* 29, 81–105. <https://doi.org/10.1146/annurev-cellbio-101512-122405>.
 115. Oczkiewicz, M., Piestrzyska-Kajtoch, A., Piórkowska, K., Rejdach, B., and Różycki, M. (2010). Expression of DLK1 and MEG3 genes in porcine tissues during postnatal development. *Genet. Mol. Biol.* 33, 790–794. <https://doi.org/10.1590/S1415-47572010000400030>.
 116. Zoppi, G., Andreotti, G., Pajno-Ferrara, F., Njai, D.M., and Gaburro, D. (1972). Exocrine Pancreas Function in Premature and Full Term Neonates. *Pediatr. Res.* 6, 880–886. <https://doi.org/10.1203/00006450-197212000-00005>.
 117. Kampf, C., Olsson, I., Ryberg, U., Sjöstedt, E., and Pontén, F. (2012). Production of Tissue Microarrays, Immunohistochemistry Staining and Digitalization Within the Human Protein Atlas. *J. Vis. Exp.* 63, 3620. <https://doi.org/10.3791/3620>.
 118. Casals, A.M., Mitsios, N., Mulder, J., and Lundberg, E. (2023). Indirect Immunofluorescence – Tissue Staining in TMA and Whole Tissue FFPE Sections Protocol v.
 119. Casals, A.M., Sitnik, I., Schürch, C.M., Malmqvist, M., Macdonald, P.E., and Lundberg, E. (2023) Codex® Multiplexed Imaging – tissue staining in TMA and whole tissue FFPE sections protocol v).
 120. Mardamshina, M., Navarro, F.B., Casals, A.M., Avenel, C., Wählby, C., and Lundberg, E. (2025). Streamlining Multiplexed Tissue Image Analysis with PIPeX: An Integrated Automated Pipeline for Image Processing and EXploration for Diverse Tissue Types. Preprint at bioRxiv. 2025.05.04. 652145. <https://doi.org/10.1101/2025.05.04.652145>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-SOX9	Atlas Antibodies	Cat#AMAb90795; RRID: AB_2665670
Anti-ANXA13	The Human Protein Atlas	Cat#HPA019650; RRID: AB_1844856
Goat-anti-rabbit IgG (H+L) Highly Cross-Adsorbed secondary Antibody, Alexa Fluor 647	ThermoFisher	Cat#A-21245; RRID: AB_2535813
Goat-anti-mouse IgG (H+L) Highly Cross-Adsorbed secondary Antibody, Alexa Fluor 555	ThermoFisher	Cat#A-21424; RRID: AB_141780
Anti-Cytokeratin 19 antibody [EP1580Y]	Abcam	Cat#ab52625; RRID: AB_2281020
NKX6-1 antibody	DHSB	Cat#F55A12-c
Donkey anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor™ 488	ThermoFisher	Cat#A-21206; RRID: AB_2535792
Donkey anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor™ 568	ThermoFisher	Cat#A10037; RRID: AB_11180865
Anti-SOX9 Antibody	Sigma-Aldrich	Cat#AB5535
HNF-1 Antibody	Santa Cruz Biotechnology	Cat#Sc-8986; RRID: AB_2116746
Anti-MED12 antibody produced in rabbit	Sigma-Aldrich	Cat#HPA003184; RRID: AB_1079349
IHC-plus Polyclonal Rabbit anti-Human Insulin Antibody	LS Bio	Cat#LS-B2526; RRID: AB_1650507
CODEX primary antibodies (see Table 1)	N/A	N/A
Biological samples		
Adult pancreatic tissue	Leiden University Medical Centre	N/A
Pancreatic TMA cores	Institute of Pathology of Technical University of Munich	N/A
Fetal pancreatic tissue	Leiden University Medical Centre	N/A
Cadaveric donor pancreata	Eurotransplant multiorgan donation programme	N/A
Human Pancreatic Islets	IRCCS San Raffaele Scientific Institute, Leiden University Medical Center, Prodo Laboratories Inc.	N/A
Chemicals, peptides, and recombinant proteins		
HistoChoice	Sigma-Aldrich	Cat#H2779
TrypLE	Life Technologies	Cat#12563029
0.05% Trypsin-EDTA	ThermoFisher	Cat#25300054
Nonidet P 40 Substitute	Sigma-Aldrich	Cat#74385
EDTA Buffer (pH 6)	Akoya Biosciences	Cat#AR900250ML
1x PBS	Gibo	Cat#18912-014
Blocking reagent (pH7.5)	Akoya Biosciences	Cat#FP1020
1M NaOH	Sigma-Aldrich	Cat#S5881
0.3% Triton	Sigma-Aldrich	Cat#T8787
Hoechst nuclear staining	ThermoFisher	Cat#H3570
0.1% poly-L-lysine	Sigma-Aldrich	Cat#P8920
Hydratation Buffer	Akoya Biosciences	Cat#7000017
Staining Buffer	Akoya Biosciences	Cat#7000017

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Storage Buffer	Akoya Biosciences	Cat#7000017
1.6% paraformaldehyde	ThermoFisher	Cat#043368
30% H ₂ O ₂	Sigma-Aldrich	Cat#216763
Methanol	Sigma-Aldrich	Cat#322415
Fixative Solution	Akoya Biosciences	Cat#7000017
AF488 dye	Akoya Biosciences	Cat#7000017
Atto550 dye	Akoya Biosciences	Cat#7000017
Cy5 dye	Akoya Biosciences	Cat#7000017
BOND Epitope Retrieval Solution 1	Leica	Cat#AR9961

Critical commercial assays

Chromium Next GEM Single Cell 3' Kit v3.1	10x Genomics	Cat#PN-1000268
Chromium Single Cell ATAC Library & Gel Bead Kit v1	10x Genomics	Cat#PN-1000175
Xenium Prime 5K Human Pan Tissue & Pathways Panel	10x Genomics	Cat#PN-1000671
Additional custom Xenium probes	10x Genomics	Cat#PN1000766
Cisbio HTRF High Range Insulin Assay Kit	Cisbio	Cat#62IN1PEG

Deposited data

Transcriptomic, epigenomic and multiomic data	EGA	EGAS50000001872
Processed .rds and .h5ad objects	This article	https://singlecell.charite.de
CODEX pipelines	GitHub	https://github.com/CellProfiling/ell_espace
CODEX data	Biolmage Archive	S-BIAD3557
Raw data used to create graphs Figures 4J, 4K, and S17A–S17G	Data S1	This article
Additional Notes and data related to this paper	Mendeley	doi: 10.17632/9vwrk7zt7f.1

Software and algorithms

Cellranger	10x Genomics	v6.1.1
R	CRAN	v4.1.2
Python	Python.org	v3.11
Signac	CRAN	v1.5.0
Seurat	CRAN	v4.4.0
CellBender	GitHub	v0.2.2
DoubletFinder	GitHub	v2.0.3
Scrublet	GitHub	v0.2.3
Harmony	GitHub	v1.2.0
scOMM	GitHub	v0.1.0
EnsDb.Hsapiens.v86 package	Bioconductor	v2.99.0
Azymuth	https://azimuth.hubmapconsortium.org/	v0.5.0
AMULET	GitHub	v1.1
MACS2	CRAN	v2.2.7.1
STREAM	GitHub	v1.1
Sceasy	GitHub	v0.0.7
SeuratDisk	GitHub	v0.0.0.9011
edgeR	Bioconductor	v4.0.16
SCENIC+	GitHub	v1.0
MiloR	Bioconductor	v.1.6.0

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Scanpy	https://scanpy.scverse.org/en/stable/	v1.10.4
Xenium Analyzer	10× Genomics	v3.1.0.4
Squidpy	GitHub	v1.6.2
Spatialdata	GitHub	v0.3.0
Numpy	https://numpy.org/	v1.26.4
Pandas	GitHub	v2.2.3
Seaborn	GitHub	v0.13.2
Matplotlib	GitHub	v3.10.0
Ezomero	GitHub	v3.1.0
HarmonyPy	GitHub	v0.0.10
Scikit-image	GitHub	v0.25.2
StarDist	GitHub	v0.9.1
Geojson	GitHub	v3.2.0
Perseus	https://maxquant.net/perseus/	v2.0.10.0
Sva	Bioconductor	3.5.0
PreprocessCore	Bioconductor	1.7.4
Plotly	GitHub	5.2.0
PIPEX	In-house	N/A
Zeiss Zen	https://www.zeiss.com/microscopy/en/products/software/zeiss-zen.html	v3.13

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Sample procurement of adult pancreas tissue

Pancreatic tissue procurement from adult organ donors was performed at Leiden University Medical Center (Leiden, Netherlands). The tissue acquisition was approved by the local clinical ethic committee and followed their guidelines. The samples were obtained in a standardized manner from four distinct locations (pancreas head, body, tail, processus uncinatus). For this purpose, three slices along the sagittal axis were taken from each organ, divided into one slice through the head of the pancreas, including the processus uncinatus, and one slice each through the central region of the pancreatic body and the pancreatic tail. The first slice from the head region was further divided into a cranial part hereinafter referred to as head and a caudal part representing the processus uncinatus. The grossing and sampling procedure was performed as fast as possible to ensure optimal tissue quality. The tissue from each individual location was further subdivided and, depending on further downstream analysis, either snap frozen with liquid nitrogen or fixated in neutral buffered formalin 10% and shipped to the corresponding consortium partners. Additionally, tissue microarrays (TMA) from the formalin fixated material of the organ donors were constructed. Of 17 organ donors only one was excluded completely from the here described analysis due to the detection of major pathological changes, resulting in 16 organ donors with high quality tissue samples.

Furthermore, from archived material from the Institute of Pathology of Technical University of Munich (Munich, Germany) we generated TMA blocks containing pancreas tissue with microscopic normal architecture on morphological level. These tissue samples derived from transplant cases or surgical specimens in which inconspicuous pancreas tissue could be identified distant from the pathological lesion. TMA generation and distribution was approved by the local clinical ethic committee (2022-207-S-NP).

Pancreatic tissue and isolated islets from LUMC

Human cadaveric donor pancreata were procured through the Eurotransplant multiorgan donation programme. Pancreatic tissue was used if the pancreas could not be used for clinical pancreas or islet transplantation, according to Dutch national laws, and if research consent was present.

Isolated human cells were cultured in CMRL 1066 (Corning, 99-663-CV, 5 mmol/l glucose), supplemented with 10% FCS, 2 mmol/l l-glutamine, 100 U/ml penicillin, 100 µg/ml streptomycin, 50 µg/ml gentamycin, 20 µg/ml ciprofloxacin, 10 mmol/l HEPES and 1.2 mg/ml nicotinamide.

Fetal tissue LUMC

Fetal pancreatic tissue from elective abortions was used after written (mother) informed consent was provided. Collection and use of human fetal tissue for research was approved by the medical ethics committee of the LUMC. A complete list of samples is presented in [Table S1](#)

METHOD DETAILS

Culture of primary human islets

Human pancreatic islets were obtained from organ donors through three sources: the Diabetes Research Institute at the IRCCS San Raffaele Scientific Institute (Milan, Italy), Leiden University Medical Center (Leiden, Netherlands) and Prodo Laboratories Inc. (Aliso Viejo, CA, USA). Islet isolation procedures were conducted in accordance with local ethical guidelines and approvals, specifically by the Human Research Ethics Boards at each respective site. The use of anonymized donor tissue was approved by the Clinical Research Ethics Committee of Hospital Parc de Salut Mar. In all cases, informed consent for research use of pancreatic tissue was provided by the donors' families. Islets were shipped to Centre for Genomic Regulation (Barcelona, Spain) in CRML-1066 transport medium and, upon arrival, cultured at 37°C in a humidified incubator with 5% CO₂. The islets were maintained in glucose-free RPMI 1640 medium supplemented with 10% fetal calf serum, 100 U/mL penicillin, 100 U/mL streptomycin, and 4.5 mM, 11 mM or 25 mM glucose, for a period of 72 to 96 hours prior to downstream processing.

In vitro human islet glucose stimulated insulin secretion

Dynamic glucose-stimulated insulin secretion was assessed using a custom-designed perfusion system. The setup included an Ismatec 8-channel peristaltic pump (Model ISM931A, Ismatec), equipped with 2-stop, 0.38 mm Tygon tubing (Fisher Scientific), which directed the flow through a Biorep PERI-CHAMBER perfusion chamber. The outflow was routed via 1.016 mm Tygon tubing (Fisher Scientific) to a Biorep PERI-NOZZLE to ensure uniform drop collection. A total of 50 human islets were suspended in Krebs-Ringer Buffer maintained at 37°C. Perfusion was carried out at a constant flow rate of 250 µL/min, using buffer containing varying glucose concentrations (low glucose 3 mM, high glucose 15 mM) and additional secretagogues (10 nM Exendin 4, 30 mM KCl). Effluent samples were manually collected every 5 minutes into a 96 deep-well plate. At the conclusion of the perfusion, the islets were collected for total insulin content analysis. The islets were first sonicated in water for 20 seconds, then transferred to acidic ethanol (1.5% HCl in 100% ethanol). Insulin concentrations in the 5-minute effluent fractions and total insulin content were quantified using the Cisbio HTRF High Range Insulin Assay Kit (Cisbio 62IN1PEG), according to the manufacturer's protocol. Measurements were performed in 96-well plates and read using a Tecan SPARK 10M multimode plate reader.

Single cell/nuclei preparation

A rice-grain-sized piece of frozen tissue was cut with a scalpel from the frozen samples. For the fetal samples the whole piece was used. To isolate the nuclei we used the citric acid nuclei isolation protocol published by Tosti and Conrad.²⁰

In short, the frozen tissue piece was placed into 1 mL of pre-cooled S25 buffer inside a glass douncer tissue grinder (DWK Life Sciences, CAT. 357358). The tissue was crushed using a loose pestle and incubated for 5 min on ice. After repeating that step the tissue was homogenized with five more strokes of the loose pestle followed by five more strokes with the tight pestle. The suspension was transferred through a 35 µm strainer into a 1.5 mL Eppendorf tube and the strainer was washed with an additional 250 µL S25 buffer.

The nuclei suspension was centrifuged at 500 g for 5 min at 4 °C. Afterwards the supernatant was carefully removed and resuspended in 500 µL S25 buffer. After checking the nuclei number and quality the suspension was centrifuged as before or, if a lot of debris was visible, a density centrifugation was done instead. For this 500 µL of S88 buffer is carefully added to the bottom of the Eppendorf tube and centrifuged at 1000 g for 10 min at 4 °C. Following the removal of the supernatant, the nuclei were resuspended in 300 µL S25 buffer and after checking the number of nuclei, the suspension was centrifuged at 500 g for 5 min at 4 °C.

The supernatant was removed as much as possible and the nuclei were resuspended in the appropriate amount of resuspension buffer for the desired concentration. The nuclei suspension was diluted to load 10,000 nuclei per sample onto the 10x Genomics Chromium controller. After Gel Bead-in-Emulsion (GEM) generation and the RT-PCR, the GEMs were stored at 4 °C until all adult samples from that batch (8-16 samples) were at this point, with a maximum waiting time of one hour.

For cultured human islets, approximately 2000 islets were harvested into 15 mL Falcon tubes, pelleted by centrifugation (200 rcf, 5 min), and washed with ice-cold phosphate-buffered saline (PBS). After removing the PBS, islets were enzymatically dissociated using a 1:1 mixture of TrypLE (Life Technologies, #12563029) and 0.05% Trypsin-EDTA (ThermoFisher, #25300054). The dissociation was carried out for 10 minutes at 37°C in a shaking water bath, with gentle pipette mixing every 3–5 minutes to facilitate breakdown into single cells. To stop the enzymatic reaction, an ice-cold solution of 5% fetal bovine serum (FBS) in PBS was added. The resulting cell suspension was filtered through a 33 µm strainer to remove clumps and extracellular DNA. Cells were then centrifuged at 250 rcf for 3 minutes, resuspended in ice-cold 0.04% bovine serum albumin (BSA) in PBS, and stained with Trypan Blue for viability assessment. Single-cell dissociation efficiency was confirmed microscopically, and cells were manually counted using a hemocytometer. Following dissociation and quality control, the single-cell suspension was aliquoted into separate tubes for scRNA-seq and scATAC-seq. For scATAC-seq, nuclei were isolated following 10x Genomics Nuclei Isolation for single-cell ATAC sequencing protocol (CG000169). Briefly, single-cell suspension was resuspended in lysis buffer (10 mM Tris-HCl (pH 7.4), 10 mM NaCl, 3 mM MgCl₂, 0.1%, Nonidet P 40 Substitute (Sigma 74385), 0.1% Tween-20 and 0.01% Digitonin, 1% BSA), and incubated on ice for 5 min. Lysis was stopped by adding 1 mL of chilled wash buffer (10 mM Tris-HCl (pH 7.4), 10 mM NaCl, 3 mM MgCl₂, 0.1% Tween-20, 1% BSA). Nuclei were centrifuged 500 g for 5 min at 4°C and resuspended in Diluted Nuclei Buffer.

Single cell/nuclei library preparation and sequencing

Fetal: From each fetal sample 10,000 nuclei were loaded onto two lanes of a 10x Genomics Chip. The libraries were generated with the Chromium Next GEM Single Cell 3' Kit v3.1 (PN-1000268) following the corresponding protocol (CG000315).

Adult: Library generation was performed with the Chromium Next GEM Single Cell 3' Kit v3.1 (PN-1000268) following the corresponding protocol (CG000315) was adapted to be automated with the Biomek i7 (Beckman Coulter, Life Sciences). After library generation, few libraries underwent an additional manual clean-up before they were sent out for sequencing. For single-cell ATAC, libraries were generated using the Chromium Single Cell ATAC Library & Gel Bead Kit v1 (10x Genomics, PN-1000175) according to the manufacturer's instructions.

Cultured islets: From each cultured islet sample, 7,000 nuclei were loaded onto a 10x Genomics Chromium controller channel, following manufacturers' instructions. Library generation was performed with the Chromium Next GEM Single Cell 3' Kit v3.1 (PN-1000268) following the corresponding protocol (CG000315) for scRNA-seq, and with Chromium Next GEM Single Cell ATAC Reagent Kits v1.1 (protocol CG000204) for scATAC-seq.

All libraries were sequenced on an Illumina NovaSeq600 (S4 flow cell, paired-end)

Single cell/nuclei data pre-processing

The fastq-files of the sequenced samples were loaded into the One Touch Pipeline (OTP), which is an automation platform managing Next-Generation Sequencing (NGS) data and calling bioinformatic pipelines for processing these data.⁹⁰ The used pipeline consisted of cellranger (v6.1.1) and the alignment to the GRCh38 genome.

Single cell/nuclei adult, islets and fetal pancreas snRNA ambient RNA removal

To reduce ambient RNA contamination, which is particularly relevant in pancreatic tissue due to high RNA content and fragility of acinar cells, we applied CellBender (v0.2.2)⁸² to all samples. For each sample, the parameters *-expected-cells* and *-total-droplets-included* were optimized based on the CellRanger web summary and knee plots, respectively. This ensured the inclusion of cell-containing droplets and a sufficient number of empty droplets to accurately model the background ambient profile. The model was trained for 150 epochs with a false positive rate (FPR) of 0.01.

Post-correction quality control included visual inspection of highly expressed marker genes (e.g. *PRSS1*, *INS*) to confirm the restriction of transcripts to their expected cell populations.

All downstream analyses were performed using the CellBender-corrected count matrix. To ensure transparency and allow for alternative re-analysis, both the corrected and the raw, uncorrected count matrices have been made available as described in the data availability section.

Single cell/nuclei adult, islets and fetal pancreas snRNA processing

Initial data analysis was performed using R (v4.1.2) and Seurat (v4.4.0).⁹¹ Cells with fewer than 1,000 UMIs and fewer than 200 genes (for adult samples) and 300 genes (for fetal samples) were excluded as low-quality cells. Upper thresholds for number of genes per cell were determined individually based on their gene expression distributions, with an average cutoff around 6,000 genes per cell. Data were then normalized, log1p transformed, the top 2,000 variable genes were identified, scaled, and principal component analysis (PCA) was performed following standard Seurat workflows.

Single nuclei doublet detection and removal in snRNA

Potential doublets were identified in each sample separately using two tools: DoubletFinder (v2.0.3)⁹² and Scrublet (v0.2.3),⁹³ both run with standard parameters as described in their respective vignettes. Only cells that were identified by both tools as doublets were removed from the dataset.

Single cell/nuclei integration

Fetal and adult samples were processed separately following the same integration workflow. For each dataset, individual samples were merged using Seurat's *merge()* function. The merged objects were normalized, log1p transformed and the top 2,000 variable genes were identified. Data were then scaled and principal component analysis (PCA) was performed. Batch correction was applied using Harmony (v1.2.0),⁹⁴ using sample identity as the integration variable. Dimensionality reduction was performed with UMAP based on the Harmony-corrected components. Cell-cell neighborhood graphs were constructed, and clustering was performed across a range of resolutions (0.1 to 1.5) to identify transcriptionally distinct populations.

Single cell/nuclei annotation of adult, islets and fetal snRNA data

Cell type annotation was performed in three iterative levels of increasing resolution. Initially, a reference-based approach was used to guide broad classification. Specifically, label transfer using scOMM⁵³ was applied using the pancreas dataset¹³ to generate preliminary cell type predictions. These initial labels were refined by identifying differentially expressed genes (DEGs) per Leiden cluster and manually assigning broad cell type categories (annotation level 1: epithelial, mesenchymal, immune) based on canonical markers.

Each broad category was then subset into a separate object and reprocessed independently, including normalization, identification of variable genes, scaling, dimensionality reduction, clustering, and DEG analysis. This enabled finer classification into specific

cell types (annotation level 2). Subsequently, each identified cell type was again subsetted, reprocessed, and analyzed to resolve cell states or subtypes using DEGs and known markers from the literature (annotation level 3).

Throughout this iterative process, additional doublets not identified during initial filtering were detected and removed. After completion of all three annotation levels, the fully annotated subsets were merged back into a single object for downstream analysis.

snATAC processing

R (v4.1.2) and Signac (v1.5.0)⁹⁵ were used for the initial snATAC data analysis. Aggregated samples were annotated using the hg38 annotation extracted from EnsDb.Hsapiens.v86 package (v2.99.0) annotation with the UCSC annotation style. Initial normalization and dimensionality reduction were performed according to the standard Signac pipeline. After an initial exploration, low quality cells were discarded with more relaxed parameters than the original pipeline (TSS enrichment >2 and % in reads >25). Then, the dimensional reduction was recalculated and Harmony (v1.2.0) was run over dimension 2 to 40 to correct sampling batch effect and clusters were calculated using the SLM algorithm available in the Seurat package (v4.4.0)⁹¹ at default resolution value. The gene activity matrix was also calculated using the *GeneActivity* function available on Signac, with a 2000 bp extension of the upstream region.

snATAC annotation of adult whole pancreas data

Based on the gene activity data, manual and automatic annotation was employed to assign biological identity to the clusters. Manual cell types were assigned according to expression of the markers described in Tosti et al.¹³ Azimuth (v0.5.0) and the pancreatic AzimuthDB were used to contrast the performed manual annotation.

To further refine the annotation, unsupervised label transferring was performed through ScoMM⁵³ with the multiome dataset as a reference. A single layer of 300 hidden nodes was trained and markers were limited to shared peaks only.

snATAC doublet detection and removal

After the initial clustering, AMULET (v1.1)⁹⁶ was used to detect potential doublet cells and unwanted cells were removed. Furthermore, Acinar cluster 3 was removed since it presented a high proportion of doublets and accumulation of high-count cells, hinting at a technical cluster. Similarly, Acinar cluster 4 was also removed since only two cells remained.

snATAC peak calling

MACS2 (v2.2.7.1)⁹⁷ was used to call the peaks per each single identified acinar and ductal sub-state as well as the rest of annotated cell types individually. Afterwards, the different peak callings were performed, the results were merged into a single combined object. After the new set of peaks was calculated, the snATAC was rebuilt based on these peaks and the initial analysis was repeated till the cell types were newly annotated and the gene activity matrix recalculated. Nevertheless, an acinar cluster had to be filtered out in the quality control step since it presented a low FRiP.

snATAC and scRNA integration of adult whole pancreas data

To integrate the scRNA-seq and snATAC-seq data from whole adult pancreas samples the bridge integration available in the Seurat documentation was conducted. A multiome dataset served as the bridge between both adult datasets. The first 30 dimensions were used for the integration (avoiding the first dimension in the case of snATAC data since it is heavily correlated to sequencing depth). Donors were considered as a batch variable for the integration.

snATAC and scRNA label transferring between the adult whole pancreas datasets

ScoMM⁵³ was the software used to transfer labels between both single-assay objects. To better bridge the gap between both modalities, an intermediate step was added by employing the multiomic data. For the RNA to ATAC label transfer, the RNA to multiome model was limited to the common marker genes identified in the snRNA data. The model was trained with the following parameters: a single hidden node layer (1000 nodes), 30 epochs, a batch size of 32, relu activation, 0.001 learning rate, dropouts enabled and patience value of 20. Afterwards, the multiome to ATAC model was trained, common accessible peaks identified in the multiome following the RNA label transfer. The parameters for training this model were the following: two hidden node layers (1500 and 500 nodes), 30 epochs, a batch size of 32, a dropout rate of 0.7, relu activation, batch normalization and a learning rate of 0.001. For the snRNA to multiome transfer, a 70-30 split between training and validation was used while a 50-50 ratio was employed in the multiome to snATAC transfer.

snATAC to scRNA label transfer from the adult no diabetes endocrine sample

A direct label transfer between the snATAC gene activity matrix and the scRNA data was performed for the adult nodiabetes (ND) endocrine samples using ScoMM.⁵³ In this case, three separate models were first trained and later stacked to bridge the gap between accessibility and transcriptomic data. Models were trained with default parameters except the dropout rate, set to 0.3, and the prediction threshold, lowered to 0.2. The selected models had a single layer of hidden nodes with 1300, 1400 and 2000 nodes. To stack these models into a single one, first a model fitting was required, where a validation split of 0.2 and a patience of 5 over 50 epochs was considered. Then, a stacked model with a single layer (1900 hidden nodes) was trained. A dropout percentage of 0.3 was considered and l2 was set to 0.01.

Differential gene expression, gene activity and peak accessibility testing

Seurat's FindMarkers() function was generally used to explore differential gene expression or accessibility, using Wilcoxon rank-sum tests with default filtering (logfc.threshold = 0.0, min.pct = 0.1, only.pos = FALSE). To account for inter-donor variability, we performed additional pseudobulk differential expression analyses using donor as the unit of replication, following the approaches described in Squair et al. (2021)⁹⁸ and Yang et al.⁹⁹ We calculated pseudobulk expression by aggregating the counts per donor and condition within each relevant comparison. Gene expression was modeled using a generalized linear model with donor as covariate using edgeR (v4.0.16).¹⁰⁰ We employed a likelihood ratio test corrected with Benjamini-Hochberg method for multiple testing to identify significantly different expressed genes (FDR<0.05) (See Supplementary Tables).

Converting objects between Seurat and Scanpy

Most analyses were carried inside the Seurat and Signac environments. Nevertheless, some analyses were carried out with Python exclusive packages, such as STREAM. The object conversion between environments was done employing two main packages: sceasy (v0.0.7) to convert from Seurat to H5AD and SeuratDisk (v0.0.0.9015) to convert from H5AD to H5Seurat.

snATAC trajectory analysis for the adult whole pancreas data

For the trajectory analysis two groups of cells were subset from the main whole pancreas adult snATAC object: Acinar 4 subcluster and Ductal cell simultaneously expressing *CPA1*, *PRSS1* and *CELA3A*, three known acinar TFs. The Seurat object containing these cells was then rescaled and the gene matrix was recalculated. The best dimensional reduction was obtained considering the first 15 UMAP components calculated over 20 PCA dimensions corrected with harmony. Afterwards, the clustering was performed with the Leiden algorithm and a 0.5 resolution. The object was then transformed to H5AD format and trajectories were inferred with the STREAM package (v1.1).²² Based on the dimensional reduction calculated with Seurat, 10 initial clusters were set for the elastic principal graph, which was optimized with the following parameters: *epg_alpha*=0.01, *epg_mu*=0.05, *epg_lambda*=0.01. Afterwards, branching was recalculated by adding 65 nodes to the elastic principal graph. Finally, the elastic principal graph branches were collapsed based on the number of edges (*epg_collapse_mode*='EdgesNumber', *epg_collapse_par*=2).

snATAC and snRNA adult pancreatic network

To build the snRNA and snATAC network a subsample of 60,000 cells was generated based on the snATAC data. Additionally, in order to balance the proportion of cells, the number of acinar cells selected was equated to the number of ductal cells. After subsetting the snATAC data, both snATAC and snRNA objects were transformed to python objects. Once data was transformed, SCENIC+ (v. 1.0)¹⁰¹ was employed to generate the adult pancreatic network and downstream analysis. Pipeline was run according to the available tutorial for unpaired data modality. Five cells were considered for each pseudo-cell built by the software and the downstream tutorials were used as a guide for further analysis.

Multiome sample processing, merging and annotation

An individual Seurat object was built for each of the two samples. Similar to the adult whole pancreas scATAC data, the hg38 annotation from the EnsDb.Hsapiens.v86 package was used for annotation, with the UCSC annotation style. However, only the peaks previously identified in the whole adult pancreas were kept to maintain consistency between datasets. Quality controls were performed according to the standard Signac pipeline and cells with less than 50,000 ATAC counts, less than 25,000 RNA counts and a TSS enrichment higher than two were kept. Once both objects were evaluated, they were merged into a single object, and after additional QC evaluations, cells with lower than 300 ATAC counts and 1000 RNA counts were discarded. Data was then normalized using SCTransform for the scRNA information and a PCA and UMAP (40 dimensions) was built. For the scRNA data, a UMAP reduction was built with dimension 2 to 40 of LSI reduction. Finally, a multimodal nearest neighbor matrix was used to generate a multimodal UMAP visualization and cells were clustered at 0.8 resolution and the gene activity matrix was calculated.

Pseudotime alignment between snATAC and snRNA data

Pseudotime values extracted from the trajectory analysis with STREAM²² were added to the corresponding snRNA/snATAC object. From the snRNA object, cells in the S0 to S1, S1 to S2 and S2 to S3 were selected. From the snATAC data, cells prior to the branching event in S3 were selected regardless of the branch they belonged to. After the cells were selected, pseudotime values were scaled between 0 and 1. Expression matrices were limited to the unique top 50 branch markers per each of the branches described before, sorted by log2FC and q-value. Finally, global alignment was performed by running cellAlign v.0.1.0 according to the default parameters.

Multiome cell type annotation

Manual and automatic cell annotation were employed. For the manual annotation, the set of markers reported in the Tosti et al.¹³ paper allowed us to define the different cell types. Automatic cell annotation through scOMM was performed to revise the manual annotation. The dataset published together with the Tosti et al.¹³ paper served as a reference for label transferring. Similarly, our own adult whole pancreas scATAC dataset was employed as a reference for the label transferring of the scATAC modality.

Multiome doublet identification

Three different software were used for doublet detection: AMULET for the scATAC data and Scrublet (v0.2.3)⁹³ and DoubletFinder (v2.0.3)⁹² for the sn/scRNA. Each software was run per sample and the different predictions evaluated. The predictions from AMULET and Scrublet were kept for filtering out double positive cells for further analysis.

Conserved markers analysis between the ESPACE and Tosti et al. datasets

Both datasets were subsetted to five shared cell types: Acinar-i, Acinar-s, Acinar-trans, Ductal, and Ductal-MucPos (referred to as MUC5B+ Ductal in Tosti et al.). Conserved markers for each cell type were identified using the FindConservedMarkers function from the Seurat package. Canonical markers for each population were subsequently selected based on their degree of cross-dataset conservation and visualised as a dot plot displaying normalised expression levels across both datasets.

Meta-analysis for consensus endocrine marker identification across sn/scRNA-seq datasets

Consensus markers were calculated for endocrine cells based on the methodology described by Van Gurp et al.¹⁵ First, pairwise Differential Expression (DE) analysis was performed for all the endocrine cells in the different datasets. Afterwards, for each individual comparison, the results across the datasets were integrated, which were then further integrated by target cell type in order to procure the final list of DE genes per cell type across datasets.

MILO analysis

We applied the MiloR R package (v1.6.0)⁵⁴ to perform differential abundance (DA) testing on endocrine cell population from cultured human islets dataset. A k-nearest neighbor graph was computed on harmonized PCA embeddings, and Milo neighborhoods were constructed using buildNhoodGraph(). We tested for differential abundance between donors with type 2 diabetes (T2D) and no diabetes (ND) using testNhoods() and donor-level metadata. Spatial FDR correction was applied, and neighborhoods with FDR < 10% were retained.

To group spatially adjacent neighborhoods with similar DA signals, we used groupNhoods() with max.lfc.delta = 3. To assign DA groups back to single cells, we used the Milo@nhoods matrix to determine which cells belonged to which neighborhoods. For each cell, we identified all associated neighborhoods, retrieved their corresponding NhoodGroups, and recorded the groups in a binary matrix.

To assign cells to disease-associated groups, we examined all neighborhoods each cell belonged to and determined whether they were part of T2D- or ND-associated NhoodGroups (ND: groups 1, 4, 6, 11, 12, 13, 16; T2D: groups 2, 5, 8, 10, 15). Cells with neighborhoods assigned to both sets were labeled as "mixed" in a new Milo_T2D_ND_clean column. Only cells consistently associated with either T2D or ND were used for differential gene expression analysis. We then used Seurat's FindMarkers() function to compare gene expression between T2D- and ND-associated alpha, beta and delta cells (based on Milo_T2D_ND_clean), using Wilcoxon rank-sum tests with default filtering (logfc.threshold = 0.0, min.pct = 0.1, only.pos = FALSE).

Xenium in situ sample preparation and sequencing

For the experiment the pre-designed Xenium Prime 5K Human Pan Tissue & Pathways Panel (PN-1000671) as well as additional 100 custom probes (PN1000766, Table S14) were used. Tissue sectioning and sample preprocessing were performed according to the manufacturer's protocol. Briefly, 5-μm FFPE sections were mounted on Xenium slides and dried to optimize adherence. Subsequently, the samples went through a 3-day protocol (CG000760, Rev C) including multiple steps such as priming hybridization, RNase treatment, polishing, probe hybridization, ligation, amplification and applying cell segmentation staining reagents. The pre-processed samples were analyzed using the Xenium Analyzer with the instrument software version 3.1.0.0 and analysis version 3.1.0.4. Subsequently, Hematoxylin & Eosin (H&E) staining of the tissue sections was performed according to the manufacturer's protocol for post-Xenium stainings.

Xenium data processing

The analysis was performed in Python (v3.12.8) using scanpy (v1.10.4),¹⁰² squidpy (v1.6.2),¹⁰³ and spatialdata (v0.3.0),¹⁰⁴ with numpy (v1.26.4)¹⁰⁵ and pandas (v2.2.3) for data handling. Visualizations were generated using seaborn (v0.13.2)¹⁰⁶ and matplotlib (v3.10.0).¹⁰⁷ To facilitate labelling of individual cores, the images from the Xenium experiment were uploaded to Omero, a web server based image repository, where a region of interest (ROI) around each core on each TMA was annotated and uniquely labeled. After loading the Xenium data using spatialdata the ROI coordinates were imported using ezomero (v3.1.0),¹⁰⁸ to assign each cell to their corresponding TMA core.

For QC all TMAs were merged after ensuring unique cell IDs. Cell identification was based on the Xenium segmentation output and cells with less than 10 detected transcripts as well as transcripts expressed in less than 5 cells were removed. The samples were normalized to the median of total counts followed by log1p transformation and dimensional reduction. Fifty principal components were used for integration with HarmonyPy (v0.0.10), followed up by nearest neighbor construction and UMAP visualization. In the end the object with all TMAs consisted of 1,585,738 cells.

Xenium H&E image alignment

To achieve accurate alignment between H&E-stained images of the corresponding DAPI-stained sections, we implemented a multi-step registration pipeline using scikit-image (v0.25.2).¹⁰⁹ First, downsampled images for both modalities were used for feature detection with Scale-Invariant Feature Transform (SIFT) algorithm (maximum ratio of 0.8) followed by robust model estimation using the RANSAC algorithm (residual threshold of 2 pixels and 1000 maximum trials) to estimate an affine transformation between the images. This transformation was then adapted and applied to the full-resolution H&E image. To further refine the alignment, ROIs were manually selected across the images, and nuclei within these ROIs were segmented using StarDist (v0.9.1).^{110–112} SIFT was then used to extract additional common features from the segmented nuclei and a final affine transformation was estimated using RANSAC. The aligned H&E images were saved in the OME-TIFF format.

Xenium cell type assignment

Similar to the single-cell data analysis, this was an iterative process. First, Leiden clusters were computed, followed by the identification of DEGs. These DEGs were used to identify marker genes reported in Tosti et al. and our single-cell dataset, allowing us to broadly annotate the cells into Acinar, Ductal, Endocrine, Mesenchymal, and Immune categories (annotation_I1). Next, cells sharing the same annotation were subset. Each subset was normalized, log1p-transformed, subjected to dimensionality reduction (PCA), integrated using HarmonyPy (v0.0.10), followed by nearest-neighbors computation, Leiden clustering, and DEG analysis. Using the DEGs a finer annotation of each subset was performed (annotation_I2). The endocrine compartment received an even finer annotation based on their DEGs, which is reflected in 'annotation_I2.5'.

The new labels were transferred into the object consisting of all TMAs and using Sopa (v2.0.0)¹¹³ exported into a format compatible for visualization with 10x Genomics Xenium Explorer 3.

Xenium pCAC identification

Using the H&E stained images a pathologist manually annotated in three TMAs a total of ~200 cells that based on location and phenotype were centroacinar cells. Annotations of the centroacinar cells were imported (geojson v3.2.0) and their coordinates were transformed to match the spatial transcriptomics data (see Xenium H&E image alignment.) Cells overlapping with the annotations were labelled as 'centroacinar cells'. The final transformed annotations were integrated into the AnnData object and the 258 cells were visualized using Xenium Explorer.

Due to imperfect cell segmentation, we manually selected intact cells in Xenium Explorer and annotated them using their cell IDs in the AnnData object as 'CAC' in the 'annotation_I2_cac' meta.data column. This reduced our 258 detected cells to 198 intact cells. This label was used to compute DEGs, from which we retained genes with a positive log fold change and an adjusted p-value < 0.05. These genes were used to compute a robust module score using Scanpy's score_genes() function, which was run 100 times with different random seeds, and cells within the top 5% of scores across all iterations were selected. The same approach was applied to DEGs from cluster 5 from the scRNA 'centro-acinar' object (see Figure 2D), but before the robust module score was calculated the genes were filtered to retain only genes detected in the Xenium experiment. Finally, the intersection of the top 5% scoring cells from both robust module scores was used to classify plastic centroacinar cells (pCAC).

Xenium ionocyte identification

For identification of the ionocyte population in the Xenium data we also used the robust module score approach utilizing Scanpy's score_genes() function. As input for the score the genes identified for the ionocyte population (also see Figure S4D) were used, namely: HEPACAM2, FOXI1, KIT, TFCEP2L1, PLCG2, CLNK, DMRT2, ATP6V1B1, ATP6V1B1, ATP6V1C2, KCNMA1, KRT19 and CFTR. Only the top 0.1% cells were classified as potential ionocytes and visualized in the Xenium dataset (also see Figure S4E).

Xenium multipotency score

The multipotency score is calculated using the AddModuleScore() function from Seurat. As input multipotency genes that were expressed in the snRNA cluster 5 (see Figure 2D) were used, namely: PDX1, NKX6-1, SOX9, HNF1B, FOXA2, RHOV, MNX1, INSM1, ST18, PAX6, JAG2, DLL1, HEY1, HES1, NTRK2, SLC4A4, TM4SF1, KRT19, LYZ, SPP1, H19, FXD2, MKI67, TOP2A, PCNA, DLK1, SHISA2, SOX4, TFF1, KLF5, CLU and AGR2.^{114–116}

Xenium islet composition analysis

The xenium AnnData object was subset to only include endocrine cells. To identify spatially distinct islet structures, we computed spatial neighborhoods using squidpy's spatial_neighbors() with radius=20 μ m, grouping cells by TMA using the library_key parameter. A spatial connectivity graph was constructed and used to identify spatial clusters via connected component analysis (scipy.sparse.csrgraph.connected_components) so that each cell was assigned to a spatial cluster.

For each spatial cluster, summary statistics were computed including total cell area, nucleus area, x/y coordinates, and cell count. Clusters with a total islet area below 1000 μ m² were defined as "small islets," while the remaining were considered "large islets." Differential gene expression between small and large islet-associated cells was performed using Scanpy's rank_genes_groups() with the Wilcoxon rank-sum test, comparing the small islet group against the large islet group.

CODEX Immunohistochemistry staining & slide scanning

Sections (4 μm thick) of a TMA block containing 102 cores from adult pancreas FFPE archived material underwent deparaffinization in xylene, followed by hydration in graded alcohols and blocking for endogenous peroxidase in 0.3% hydrogen peroxide diluted in 95% ethanol. The next step was Heat-Induced Epitope Retrieval (HIER) treatment, using pH 6 EDTA buffer (Lab Vision, Fremont, CA) in a decloaking chamber (Biocare Medical, Walnut Creek, CA) for 4 min at 125 $^{\circ}\text{C}$, after which the samples were allowed to cool to 90 $^{\circ}\text{C}$. The automated immunohistochemistry was performed as previously described¹¹⁷ using anti-SOX9 antibody (Atlas Antibodies, Cat. No. AMAb90795, dilution 1:1500) and anti-ANXA13 antibody (The Human Protein Atlas, Cat. No. HPA019650, dilution 1:1000). The stained slides were scanned using a Digital Slide Scanner Axioscan 7 (ZEISS), equipped with a Hitachi 203 camera at 20x with resolution of 0.22 μm /pixel. The images were manually annotated by two different experts.

CODEX Indirect immunofluorescence staining & antibody conjugation

Prior to in-house conjugation of antibodies with a unique oligonucleotide sequence for multiplexing experiments, the specificity of the antibodies was assessed by performing single indirect immunofluorescence staining¹¹⁸ on single pancreatic tissue sections from a FFPE block. The block was sectioned to a thickness of 5 μm , mounted on positively charged microscope slides (VWR, Eppredia, Cat. No. 76406-502), and baked at 55 $^{\circ}\text{C}$ for 25 min. Each tissue section was deparaffinized by immersing it in HistoChoice (Sigma, Cat. No. H2779), for 5 min twice, and then rehydrated through a series of ethanol solutions (100% ethanol twice, followed by 90%, 70%, 50%, 30%, ddH₂O twice; each step lasted 5 min). The next step was Heat-Induced Epitope Retrieval (HIER) treatment, using pH 6 EDTA buffer (Akoya Biosciences, Cat. No. AR900250ML) in a pressure cooker (Bio SB TintoRetriever, Cat. No. BSB-7087) at a temperature of 114–121 $^{\circ}\text{C}$, high pressure, for 20 min. Every tissue sample was equilibrated at room temperature, followed by two 2-minute incubations in ddH₂O. To minimize tissue autofluorescence, each tissue section underwent a photobleaching protocol where the slide was placed in a 50 ml falcon containing 25 ml 1x PBS (Gibco, Cat. No. 18912-014), 4.5 ml 30% H₂O₂ (Sigma, Cat. No. 216763) and 0.8 ml 1M NaOH (Sigma, Cat. No. S5881). The section in solution was exposed to two LED lamps (GLIME, Light Therapy Lamp UV Free 32000), positioned on either side and incubated for 45 minutes. A new photobleaching solution was prepared and replaced for an additional 45-minute incubation under the LED lamps, followed by four washes with 1x PBS washes, each step for 2 min. Each primary antibody tested was diluted in antibody diluent containing 0.3 % Triton (Sigma Aldrich, T8787) 1x PBS pH 7.4. The diluted antibody was then applied to the tissue section to be incubated overnight at 4 $^{\circ}\text{C}$. The next day, the material was washed with TBS-T three times, with each wash lasting 15 minutes. To prevent nonspecific binding, the tissue was incubated for 30 min in TNB buffer containing 0.1 M Tris-HCl/0.15 M NaCl and 0.5% blocking reagent pH 7.5 (Akoya, SKU FP1020). Depending on the species reactivity, the secondary antibodies used were goat-anti-rabbit IgG (H+L) Highly Cross-Adsorbed secondary Antibody, Alexa Fluor™ 647 (ThermoFisher, Cat. No. A-21245) or goat-anti-mouse IgG (H+L) Highly Cross-Adsorbed secondary Antibody, Alexa Fluor™ 555 (ThermoFisher, Cat. No. A-21424). The secondary antibody was diluted 1:800 in TNB containing nuclear staining Hoechst (ThermoFisher, Cat. No. H3570) and added to the tissue sections for a 90-minute incubation at room temperature. Finally, the stained material was mounted with Fluoromount-G (Thermo Fisher Scientific, 00-4958-02) on a coverslip for image acquisition.

Each single tissue section was image acquired using an inverted DMI8 microscope (Leica Microsystems, Mannheim, Germany), equipped with a digital CMOS camera (ORCA-Flash 4.0 V3, Hamamatsu), employing a 20x/0.75 NA dry objective with a resolution of 0.325 μm /pixel.

The antibody was selected for oligonucleotide conjugation based on the following criteria: 1) The staining pattern was as expected (using The Human Protein Atlas database and antibody datasheets as references); 2) The signal-to-noise ratio was higher than 3. The custom antibody conjugations were carried out in accordance with the CODEX protocol from Akoya, User Manual Rev. C.

CODEX non-disease adult and fetal highly multiplexed staining

The highly multiplexed panels include antibodies pre-conjugated by Akoya Biosciences and in-house conjugated antibodies (Table S1: adult_antibody_panel_CODEX, fetal_antibody_panel_CODEX). The multiplex antibody staining protocol used for the adult and fetal pancreas tissues was the same as that described by¹¹⁹, in accordance with the CODEX protocol from Akoya, User Manual Rev. C, with the exception of the glass format due to the instrument version: the former version is referred to as CODEX, while the newer version is referred to as PhenoCycler Fusion; for simplicity, we will refer to it as CODEX. Tissue MicroArray (TMA) adult sections were placed on glass coverslips 22x22mm (Marienfeld, Cat. No. 102052) coated with 0.1% poly-L-lysine (Sigma, Cat. No. P8920): 184 tissue cores were included, out of 198 arrayed across 9 coverslips (Table S1: TMA_adult_CODEX); the 14 excluded cores comprise 2 lost during sectioning and 12 not image-acquired. Whereas the FFPE fetal materials were sectioned and mounted on positively charged microscope slides 75x25 mm (VWR, Eppredia, Cat. No. 76406-502): 11 tissues divided into 2 different slides (Table S1: TMA_fetal_CODEX). Each FFPE tissue section underwent sample preparation as detailed in the previous section. After the material was treated with the photobleaching protocol, each glass was then transferred to a new 6-well plate (for coverslip format)/Coplin Jar (for microscope slide format) containing Hydration Buffer (Akoya Biosciences, Cat. No. 7000017) and incubating for 2 min, repeating this step twice. Following this, the glass was placed in a reservoir with a Staining Buffer (Akoya Biosciences, Cat. No. 7000017) and equilibrated for 30 min. During this time, the antibody cocktail solution was prepared by diluting all the oligonucleotide-barcoded antibodies in Staining Buffer containing N, J, G, and S blockers (Akoya Biosciences, Cat. No. 7000017). Once prepared, the mixed antibodies were applied to the tissues for an overnight incubation at 4 $^{\circ}\text{C}$. The following day, the antibody cocktail solution was removed from the tissue by washing with Staining Buffer, followed by three different consecutive post-fixation steps with washes between them: 1) Incubation in 1.6% paraformaldehyde (ThermoFisher, Cat No. 043368) diluted in Storage Buffer

(Akoya Biosciences, Cat No. 7000017) for 10 min incubation at room temperature; 2) Incubation in ice-cold methanol (Sigma, Cat No. 322415) for 5 min at 4 °C; 3) Incubation in Fixative Solution (Akoya Biosciences, Cat No. 7000017) diluted at 1:50 in 1x PBS. Each glass was then stored in a Storage Buffer until the image acquisition step. Per each experiment, a 96-well reporter plate was prepared, containing a fluorescent dye conjugated to an oligonucleotide sequence complementary to one specific antibody barcode. Unique reporters were added in groups of up to three spectrally different dyes, including AF488, Atto550, and Cy5 (Akoya Biosciences), along with nuclear staining in each single well.

CODEX Adult - image acquisition & computational image processing

Image acquisition was performed on the CODEX instrument (currently the PhenoCycler-Fusion system, Akoya Biosciences, Marlborough, MA), connected to an inverted DMI8 microscope (Leica Microsystems, Mannheim, Germany). Images were acquired with a digital CMOS camera (ORCA-Flash 4.0 V3, Hamamatsu), employing a 20x/0.75 NA dry objective with a resolution of 0.325 $\mu\text{m}/\text{pixel}$, Z-steps = 5 of $z = 1.5 \mu\text{m}$. The light source used was a SOLA-SM-II. The CODEX Instrument Manager (version 1.30.0.12) was utilized to control CODEX and integrate the microscope control software. The generated 16-bit.tiff images were processed using CODEX® Processor application (version 1.7.0.6), which performed image alignment across cycles, tile stitching, background subtraction, deconvolution, extended depth of field, and shading correction.

Every single tissue core from each TMA image/section (TMA $n = 9$) was cropped, and the images were cell segmented separately using our in-house developed toolkit PIPEX,¹²⁰ as described in the previous section. Single cells were cell segmented based on Hoechst nuclear staining (nuclei_diameter=20px, nuclei_expansion=20px); this segmentation was refined using the CDH1 (E-cadherin) membrane marker (membrane_diameter=25px). Direct calculations of segmented cell data (mean pixel intensities, markers binarization) were generated from each tissue core and saved in a Comma-Separated Value (CSV) file for downstream analysis, where additional metadata was included.

CODEX Adult - data cleaning, clustering & cell annotation

Data cleaning, clustering, cell annotation, and analysis were primarily conducted using R version 4.2.2 (RStudio) and Python version 3.12 (Jupyter Notebook and PyCharm), along with Perseus software (version 2.0.10.0). We combined datasets from nine tissue microarray (TMA) slides, resulting in a comprehensive dataset of 1,850,689 segmented cells. This dataset underwent filtration to remove technical artifacts, yielding a refined total of 1,756,482 cells. The filtration process involved excluding cells based on size, specifically removing those in the first and 100th percentiles, followed by filtering out the top 1% based on Hoechst, CTNNB1, and CDH1 intensity.

Next, we applied the log1p transformation to the data and conducted batch correction using the ComBat method from the "sva" R package. We then performed quantile normalization using the "preprocessCore" R package to ensure data harmonization.

For cell type annotation, we utilized Leiden clustering, which categorized the cells into distinct types as follows: acinar ($n=846,840$), ductal ($n=303,139$), endothelial ($n=142,586$), endocrine ($n=65,548$), MHC I+ ($n=77,806$), proliferative ($n=24,266$), Schwann ($n=5,198$), stromal ($n=256,633$), and other ($n=38,419$). To identify subpopulations of endocrine cells, we isolated the endocrine dataset and employed automated Otsu thresholding for annotation using the *scikit-image* Python library (version 0.22.0). This analysis revealed distinct subpopulations of endocrine cells, including alpha ($n=16,883$), beta ($n=12,856$), delta ($n=495$), alphabeta ($n=3,783$), betadelta ($n=1,373$), polyhormonal ($n=2,014$), and "other" ($n=28,144$). The "other" category included cells neighboring GCG+ cells, which were excluded from analysis due to technical staining artifacts that led to signal bleed-through.

For cell clustering and unsupervised analysis, we utilized the "scanpy," "squidpy," and "numpy" packages in Python,^{102,103,105} with visualizations, including a Sankey plot, generated using the Plotly library. For spatial analysis, we calculated Euclidean distances between annotated cell populations using Perseus software, and we visualized the results using the "matplotlib" and "seaborn" libraries in Python.^{106,107} Additionally, the expression of endocrine markers was illustrated in a dot plot format using the "matplotlib.pyplot" library.

CODEX Fetal - image acquisition & computational image processing

Image acquisition was performed on the Phenolmager Fusion 2.2.0 instrument (Akoya Biosciences, Marlborough, MA) where the three fluorescent oligo reporters were applied to the tissue in iterative imaging cycles. The images were captured using a widefield microscope with an Olympus UCPlanFL 20x/0.70 NA dry objective with a resolution of 0.51 $\mu\text{m}/\text{pixel}$. Whole-slide images were automatically aligned using DAPI as a reference, and the background was subtracted to be assembled into the final 8-bit QTIFFs image files.

Every tissue from each QTIFFs image file ($n = 2$) was cropped and converted into a separate.tiff for each marker channel. These images were cell segmented using our in-house developed toolkit PIPEX.¹²⁰ The subsequent strategy was followed to be able to obtain single-cell quantification: 1) Nuclei mask: Stardist ML¹¹² model was applied on the DAPI images (nuclei_diameter=10px), selection nuclear expansion (nuclei_expansion=20px) and additional settings to be able to obtain single cell information in these tissues with tightly packed cells (nuclei_definition=0.1, nuclei_closeness=0.3; 2) Membrane mask: a watershed algorithm was run using CDH1 (E-cadherin) images (membrane_diameter=15, membrane_compactness=0.7) to obtain the membrane boundary shape information. The nuclei mask expansions were cut to fit the membrane boundaries calculated in the membrane mask. Direct calculations of segmented cell data (mean pixel intensities, markers binarization) were generated from each tissue and saved in a Comma-Separated Value (CSV) file for downstream analysis, where additional metadata was included.

CODEX Fetal - data cleaning, clustering & cell annotation

The data cleaning step was performed using Python (v. 3.10) in the PyCharm IDE (v. 2024.2.5). Per each single tissue, a sanity filtering step was performed using cell size to discard segmented cells that were too small or too large; and DAPI intensities. Visual inspection of each image per marker, including the cell segmentation mask, facilitated the removal of staining artifacts. It also allowed for exclusion of regions that were presumably not part of the pancreas tissue based on expected canonical protein expressions. The cleaned data were imported into the PIPEX software for Min-Max normalization after merging all CSV files. Batch correction was subsequently performed using ComBat, with slide number used as the batch variable. The combined datasets from the 11 tissue sections included 224,713 cells. For further analysis in PIPEX, k-means ($k = 10$) clustering using ACTA2, CD31, CD45, CHGA, C-Pep, E-cadherin, GCG, HLA-ABC, HLA-DR, Pan-Cytokeratin, RBP4, SST, SYP, TTR, VIM as markers, was performed separately for each week and cell type annotations were assigned based on ranked protein expression and visual image inspection following annotation_11 categories. Following the annotations, clusters corresponding to the same cell type were combined. A more in-depth analysis of the endocrine population (18,741 cells) was performed by re-clustering for each individual donor. Eight protein markers were used: CHGA, C-Pep, GCG, Pan-Cytokeratin, RBP4, SST, SYP, and TTR. K-means unsupervised clustering was performed with $K=5$ for all cases, except for weeks 13.3, 14.1, and 14.5, where $K=10$ was used and clusters with the same protein expression profile were merged. Supplementary data contain whole tissue images with the corresponding cell annotations. The Min-Max normalized data were used to create violin plots (PanCytokertatin and CHGA), and the binarized data from Ki67 expression were used to generate a horizontal bar plot categorized by different ages.

Immunofluorescence and immunohistochemistry of adult pancreas tissue for protein level marker validation

Immunofluorescence stainings: Pancreatic tissue samples were fixed in 4% formaldehyde solution and embedded paraffin. Samples were sectioned into 4 μ m slides using the Leica microtome (RM2255, Leica, Germany). Slides were deparaffinized and rehydrated. Antigen retrieval was performed using 10mm Citrate buffer pH 6.0 followed by 15 minutes of permeabilization with 1% Triton X-100. Blocking was done with 5% normal donkey serum followed by overnight incubation at 4 °C of KRT19 1:200 (Abcam, ab52625) and NKX 6.1 1:100 (DSBH, F55A12-c) primary antibodies. After washing secondary antibodies donkey a rabbit AF488 (Thermo Fisher) and donkey a mouse AF568 (Thermo Fisher) 1:500 were incubated for an hour at room temperature. Subsequently counterstained with Hoechst 33258 and mounted with Prolong Gold (Thermo Fisher). Immunofluorescence was detected with a Zeiss LSM 980 Airyscan 2, islets were found with a 40x lens and a ~40 tile-scan of the surrounding was created as an overview image of the pancreas. These were analysed for double positive cells with Zeiss Zen 3.13 and counted with ImageJ.

Immunohistochemistry stainings: Formalin fixed paraffin embedded tissue sections were processed using a Bond RXm autostainer (Leica, Wetzlar, Germany). After deparaffinization and rehydration, heat-induced epitope retrieval was performed using BOND Epitope Retrieval Solution 1 (Leica, AR9961) corresponding to citrate buffer (pH 6.0), for 30 min (Sox9/SPP1) or 40 min (HNF1A/Insulin), followed by peroxidase blocking. Sequential double immunohistochemistry with the following primary antibodies was performed using polymer-based detection: Sox9 (1:2000, polyclonal rabbit, Sigma-Aldrich, AB5535) or HNF1A (1:50, polyclonal rabbit, Santa Cruz Biotechnology, sc-8986) were visualized with diaminobenzidine (DAB), followed by SPP1 (1:75, polyclonal rabbit, Sigma-Aldrich, HPA003184) or Insulin (1:100, polyclonal rabbit, LS Bio, LS-B2526) detected with a red chromogen. Sections were counterstained with hematoxylin, dehydrated, and mounted. HNF1A/Insulin immunohistochemistry was performed on TMA cores from the cohort of adult pancreatic samples. In each sample ($n = 67$), 50 cells showing strong cytoplasmatic staining for Insulin were evaluated for their nuclear staining reaction for HNF1A, with the exception of one sample in which only 45 Insulin positive cells were captured on the TMA cores. We categorized Insulin positive cells that showed a positive nuclear reaction for HNF1A as HNF1A-high, whereas Insulin positive cells without any nuclear staining reaction for HNF1A, were classified as HNF1A-low.