Multiomics single-cell analysis of human pancreatic islets reveals novel cellular states in health and type 1 diabetes

Maria Fasolino[1,2,3,4,7*], Gregory W. Schwartz[2,3,5,6,8*], Maria L. Golson[1,4,7], Yue J. Wang[1,4,7], Ashleigh Morgan[1,4,7], Chengyang Liu[4,7,9], Jonathan Schug[1], Jinping Liu[1,4,7], Minghui Wu[1,4,7], Daniel Traum[1,4,7], Ayano Kondo[1,4], Catherine L. May[1,4,7], Naomi Goldman[1,2,3,4,7], Wenliang Wang[1,2,3,4,7], the HPAP Consortium[7], Michael Feldman[5,7,8], Jason H. Moore[1,7,8], Alberto S. Japp[7,10], Michael R. Betts[7,10], Robert B. Faryabi[2,3,5,6,8#], Ali Naji[2,4,7,9#], Klaus H. Kaestner[1,3,4,7#], Golnaz Vahedi[1,2,3,4,6,7,8#]

Affiliations

[1]Department of Genetics, [2]Institute for Immunology, [3]Epigenetics Institute, [4]Institute for Diabetes, Obesity and Metabolism, [5]Department of Pathology and Laboratory Medicine, [6]Abramson Family Cancer Research Institute, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA 19104, USA; [7]The Human Pancreas Analysis Program https://hpap.pmacs.upenn.edu/; [8]Institute for Biomedical Informatics; [9]Department of Surgery, [10]Department of Microbiology, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA 19104, USA.

*co-first authors

#Corresponding authors

Abstract

Type 1 diabetes (T1D) is an autoimmune disease of only partially defined etiology in which immune cells destroy insulin-producing beta cells. Using single-cell transcriptomics and an advanced analytical strategy to assess pancreatic islets of T1D, autoantibody-positive, and non-diabetic organ donors, we identified both canonical cell types and rare insulin-expressing cells with a hybrid mixture of endocrine and exocrine gene signatures within all donors. We further found elevated expression of MHC Class II pathway genes in exocrine ductal cells of T1D donors, which we confirmed through CyTOF, *in situ* imaging mass cytometry, and immunofluorescence analysis. Taken together, our multimodal analyses identify novel cell types and processes that may contribute to T1D immunopathogenesis and provide new cellular and molecular insights into human pancreas function.

Introduction

Type 1 diabetes (T1D) is an autoimmune disease which occurs as a consequence of the organ-specific destruction of the insulin-producing beta cells in the islets of Langerhans within the pancreas. This complex disease is characterized by atypical beta-immune interactions including production of beta cell autoantibodies and the immunological attack on beta cells by cytotoxic $CD8^+$ T cells (*1, 2*). The immune destruction of beta cells causes deficits in peripheral tissue glucose uptake, increased hepatic glucose production, and hyperglycemia, the combination of which results in devastating effects on metabolic health and severe complications, including blindness, heart disease, and kidney disease (*3-5*).

T1D autoimmunity has been linked to poorly understood genetic and environmental factors. Genome-wide association studies have implicated multiple loci in T1D, with the major histocompatibility complex (MHC) Class II genes as the dominant susceptibility determinant of this disease (*6*). However, the precise cellular context through which T1D susceptibility genes cause the pathogenic destruction of beta cells remains to be discovered. Addressing this question is particularly challenging since the pancreas is a heterogeneous organ, composed of multiple distinct cell types. The exocrine compartment is crucial for nutrient uptake via releasing digestive enzymes, which are produced by acinar cells, into the gastrointestinal tract via the ductal system. Although the endocrine compartment, i.e. the islets of Langerhans, has been the focus of numerous studies using animal models or individual specimens obtained from autopsy collections of human pancreas, the pathogenic relevance of other pancreatic cell types, in particular the exocrine compartment, remains unclear.

Two nontrivial constraints hamper insights into comprehensive identification of the pathogenic cell types in T1D and a better understanding of the initial molecular perturbations that occur during disease pathogenesis: (1) the inability to safely biopsy the human pancreas of living donors and (2) the significant disease progression and beta cell destruction by the time patients are clinically diagnosed with T1D. Therefore, the majority of T1D studies has been performed on peripheral blood leukocytes, which is not the site of pathogenesis. Recently, the Network for Pancreatic Organ Donors

(nPOD) (*7*) and the Human Pancreas Analysis Program (HPAP) (*8*) collect pancreatic tissues from deceased organ donors diagnosed with T1D. Additionally, given that many T1D patients harbor beta cell autoantibodies (AAbs) in their bloodstream prior to clinical diagnosis, nPOD and HPAP also collect samples from donors with AAbs towards islet proteins but without a medical history of T1D, in hope of elucidating early pathogenic events. Here, we report an unprecedented multimodal analysis of millions of cells using three high-throughput single-cell techniques in pancreatic tissues of human organ donors collected by HPAP, providing novel cellular and molecular insights into T1D pathogenesis.

Results

Single-cell RNA-seq unravels canonical and novel cellular states in the human pancreas

To unmask the molecular perturbations occurring in pancreatic tissues during T1D, we constructed 81,313 single-cell RNA-seq (scRNA-seq) libraries from pancreatic islet cultures of 24 human organ donors representing three categories: individuals with T1D (n = 5), those with AAbs toward pancreatic islet proteins but no clinical diagnosis of T1D ('AAB+'; n = 8), and those with neither AAbs nor a history of T1D ('Control'; n = 11) (Fig. 1A and fig. S1 a and b, and table S1). After stringent quality control, we filtered outlier cells, removed doublets, and employed the cell type classifier 'Garnett' (*9*) (Fig. 1B, fig. S1, c to e, fig. S2, a to c, and fig. S3, a to e), to cluster 69,645 high-quality cells using 'TooManyCells' (*10*), which employs tree visualization to preserve the relationships among cell clusters (Fig. 1C). The resultant classification was confirmed by canonical gene marker expression for each cell type (Fig. 1C and fig. S4, a, c, and d).

Notably, clustering was clearly driven by cell type, as opposed to confounding factors such as autoantibody status, age, BMI, phenotypic group, or other factors (fig. S5, a to k, and fig. S6, a to i). This clustering pattern corroborates the high quality of the resultant data, insinuating the strong biological signal in these scRNA-seq measurements. Additional evidence for the lack of technical noise stems from the observation that cell type clustering was preserved when donors from T1D, AAB+, and Control categories were independently clustered (fig. S7, a to f).

Transdifferentiation of ductal cells to endocrine cells has been reported in models of extreme pancreatic injury (*11-16*). Therefore, we examined the relationship between pancreatic endocrine and ductal cells. First, we subsetted and reanalyzed the endocrine and ductal cells to achieve a more granular clustering (Fig. 1D). Upon reclustering, the major cell types – alpha cells, beta cells, delta cells, epsilon cells, PP cells, ductal cells, and acinar cells – were easily discernible by the cell type classifier and assessment of canonical markers (Fig. 1D and fig. S4, b to d). In instances where there were two transcriptionally distinct canonical cell types (i.e. Beta-1/Beta-2), differential gene expression analysis between populations provided further insights into the underlying molecular differences (table S2 to S5). For example, cells in the Beta-2 cluster expressed higher levels of stress response genes (i.e., *GDF15* and *NPTX2*) when compared to those in the Beta-1 cluster. Notably, comparison of cells in the two Ductal clusters revealed that while cells in the Ductal-1 cluster were enriched for transcription factors (TFs) associated with the endocrine cell fate (i.e., *PDX1* and *NKX6-1*), those in the Ductal-2 cluster only expressed acinar TFs (i.e., *PTF1A* and *GATA4*). These differences in TF gene activation in the two ductal populations were further supported by their position on the dendrogram, and might provide insights into the underlying biology of the acinar-duct-islet axis. For instance, it is plausible that cells in the Ductal-1 cluster might be more prone to transdifferentiation into endocrine cells than those in the Ductal-2 cluster.

Surprisingly, a substantial number of cells (4,001 cells) were not included in these canonical cell type clusters, but formed their own transcriptionally distinct groups in close proximity to one another on the dendrogram. This cluster comprised 5.7% of all profiled cells, with a mixture of cellular classifications and expression of canonical gene markers associated with beta, alpha, ductal, and acinar cells. Therefore, we labeled these cells as 'Hybrid' cells (Fig. 1D and fig. S4b). Notably, the gene expressed most highly and consistently in the Hybrid cell population was *INS* (fig. S4b), and a comprehensive examination of the cells comprising this cluster ruled out the possibility of them being potential doublets (fig. S1d and fig. S2 a to e).

Since immune cell-mediated autoimmune destruction of viable pancreatic cells is a pathognomonic feature of T1D, we further examined the intrapancreatic immune cells profiled by scRNA-seq. First, we subsetted and reclustered the cells comprising the

‘Immune’ cluster from the comprehensive tree (Fig. 1C), and found that this population also contained stellate (*RGS5* high) and Schwann (*PLP1* high) cells along with immune cells (*PTPRC*) (fig. S8a and b). Using the Immunological Genome Project (ImmGen) cell type signatures (*17*), we further found antigen-presenting cells (APCs) and macrophages gene signatures, for example *CD68, SPI1, CD14,* and *CD16,* were most frequently upregulated in immune cell subset (fig. S8b and c), suggesting that these cell types comprise the majority of the identified immune cells.

In addition to successful identification of major endocrine and exocrine cell types and pancreatic immune cells, we also observed that the overall proportions of cell types was in accordance with previous work (*5, 18-22*). Each of the major identified cell types comprised of cells from the three donor groups with varying proportions (Fig. 1, E and F, fig. S8, d to e, and fig. S9 a to b). As expected, we found that there was a lower proportion of beta cells in the T1D cohort compared to the AAB+ or Control groups (fig. S9 a to b). Conversely, both acinar and ductal cells comprised a higher portion in the T1D cohort, reflecting the difficulty of isolating high purity islets from T1D donors. Furthermore, within major cell clusters, there were varying degrees of separation based on donor group, which is to be expected due to likely transcriptomic differences among the three donor states (Fig. 1, E to F, and fig. S8a). Notably, Ductal-1 cells clearly separated into distinct T1D-enriched and Control-enriched groups (Fig. 1F). Taken together, our data indicate that transcriptomic differences amongst cell types and not technical biases drive separation of major cellular clades , and the donor state further segregates each cell type.

AAB+ and T1D endocrine and exocrine cells have common and distinct gene expression programs

Beta cell autoantibodies can be detected in the bloodstream of T1D patients months to years prior to clinical diagnosis. To investigate the molecular events that occur in a possibly prediabetic state, we next compared transcriptomic divergence of AAB+ and T1D cells from Controls (Fig. 2A and tables S6 to S35). For the most part, the degree of overlap between dysregulated genes and pathways in AAB+ and T1D states were cell type-dependent (Fig. 2, B to E, fig. S10, a to b, fig. S11, a to b, fig. S12, a to d, fig. S13, a to d, and fig. S14, a to c). However, some pathways were found to be commonly

dysregulated in multiple cell types across T1D and AAB+; including ‘apoptotic signaling’, various protein folding ontologies, various viral-related ontologies, ‘autophagy’, ‘inflammatory pathways’, and ‘stress response’.

Given the importance of beta cells in T1D pathogenesis, we next examined the transcriptional changes in the two populations of annotated beta cells, Beta-1 and Beta-2. A large number of genes were downregulated in T1D (9,512 genes) and AAB+ (3,666 genes) Beta-1 cells compared to Controls, which frequently overlapped (2,896 genes, 28%; $p < 2.2e-16$) between the two donor groups (Fig. 2B and fig. S10a). Notable pathways that were frequently downregulated in Beta-1 cells were immune/stress response and apoptosis related (Fig. 2B and fig. S10a). Given that beta cells are destroyed by immune cells in T1D, it is possible that these remaining Beta-1 cells captured in our analysis are able to survive and function by downregulating immune and apoptotic signaling. Notably, these results also suggest that cells from AAB+ donors in this beta cell population employ similar protective molecular mechanisms to enhance survival and function. The Beta-2 cell population displayed a small proportion of genes (283 genes; 4%) that were elevated in both T1D and AAB+ cells when compared to Controls (Fig. 2C, and fig. S10b). Additionally, an even smaller number of genes were downregulated in T1D and AAB+ Beta-2 cells when compared to Controls (Fig. 2C and fig. S10b). Nonetheless, a few gene pathways were found to be commonly dysregulated across both donor groups. Two interrelated pathways dysregulated in both T1D and AAB+ Beta-2 cells suggest a dysregulation of protein folding, an essential function for cellular homeostasis. Additionally, the TNF signaling pathways, which have been implicated as an important regulator of autoimmune processes (*23*), were significantly downregulated across the two donor groups (Fig. 2C and fig. S10b; p-value $< 2.2e-16$). Together, our differential expression analyses complement earlier studies on the pathways triggering beta cell dysfunction and death, and can further enable the identification of endogenous cellular responses to restore homeostasis at asymptomatic early stages.

Given the clear segregation of ductal cell populations based on the donor group, we next sought to examine the transcriptional changes in the two populations of ductal cells, Ductal-1 and Ductal-2. A large number of genes were upregulated in T1D (7,175 genes) and AAB+ (4,371 gene) Ductal-1 cells when compared to Controls, a significant

number of which were common between the two donor groups (Fig. 2D and fig. S11a; 2,283 genes; 25%; p-value < 1e-12). Notable pathways that are upregulated in T1D and AAB+ cells are pathways involved in apoptosis, stress, and immune response (Fig. 2D and fig. S11a). In the Ductal-2 cell population, although a large number of upregulated genes were observed in T1D (6,841 genes), there were not nearly as many upregulated genes in AAB+ cells (1,106 genes) when compared to Controls (Fig. 2E and fig. S11b). Furthermore, in the T1D and AAB+ Ductal-2 cell population, there was a modest but significant overlap between upregulated genes (Fig. 2E and fig. S11b; 11%; p-value < 1e-12). Nevertheless, various gene pathways were found to be significantly upregulated across Ductal-1 cells of both disease states, including apoptotic, autophagy, and immune signaling pathways (Fig. 2, D and E, and fig. S11, a and b). Taken together, these findings suggest that although AAB+ donors maintain normoglycemia, significant transcriptional dysregulation is occurring in AAB+ endocrine and exocrine cells that is highly similar to that in T1D.

Next, we directly compared T1D to AAB+ cells (Fig. 2F and table S36 to S50). For both groups of beta cells, genes associated with apoptosis, stress, and immune response pathways were found to be upregulated in AAB+ cells compared to T1D cells (Fig. 2, G and H, and fig. S10, a and b). For AAB+ Beta-1 cells, it is possible that cells with 'Epstein-Barr virus infection' pathway activation are able to survive due to the upregulation of the 'autophagy pathway', since this pathway is known to enhance survival over apoptotic or necrotic pathways (*24*); this line of reasoning may also apply to pathways that quell cellular stressors, such as ER stress (Fig. 2, G and H, and fig. S10, a and b). In addition to several commonly upregulated pathways in Beta-1 and Beta-2 AAB+ cells, apoptotic and adaptive immunity signaling were only upregulated in Beta-2 AAB+ cells. These data suggest that this population is undergoing cell death, indicated by the expression of adaptive immune cell genes and *BCL10*. In Ductal cell populations, there was a larger number of upregulated genes in T1D (Fig. 2, I and J, and fig. S11, a and b). More specifically, apoptotic, stress, and immune responses were activated in T1D ductal cells in comparison to AAB+ ductal cells (Fig. 2, I and J, and fig. S11, a and b). Remarkably, interferon alpha and beta pathways, known to be critical in T1D disease pathogenesis (*25-27*), were significantly elevated in T1D ductal cells compared to either Control or

AAB+ ductal cells (fig. S14d). While ductal cells have not historically been implicated in T1D, these findings position these exocrine cells in disease pathogenesis.

Taken together, our data indicate that both endocrine and exocrine pancreatic cell types show transcriptional dysregulation in T1D. Remarkably, AAB+ cells exhibit significant transcriptional changes similar to those observed in T1D.

Single-cell RNA-seq profiling identifies non-canonical, rare Hybrid cells

We next sought to gain a deeper understanding of the rare, non-canonical cell population, which we termed 'Hybrid' cells (Fig. 1D). This group consisted of 4,001 cells (5.7% of all cells) with a large fraction of the cells expressing high levels of *INS*, with additional activation of the gene signatures associated with alpha, ductal, and acinar cells (Fig. 1D and fig. S4b). We found that approximately 1.3% of cells from each individual donor (median across all donors) were classified as Hybrid cells, with the exception of one ABB+ donor (AAB+ #3), which was overrepresented in the Hybrid cluster (Fig. 1F, fig. S6d, and fig. S9, a and b). However, the Hybrid cluster was present at a similar frequency even when this AAB+ donor was removed, providing evidence of the presence of this cellular state across multiple donors (fig. S15, a and b).

Detailed interrogation and cellular classifications revealed three distinct subpopulations across the Hybrid cluster (Fig. 3, A and B). The 'Ductal-like Hybrids' subpopulation expressed ductal cell signature genes; the 'Endocrine-like Hybrids' subpopulation exhibited alpha-cell gene signature. Finally, the third subpopulation of Hybrid cells, called 'Acinar-like Hybrids' expressed acinar cell signature genes (Fig. 3. A and B). In addition to expressing beta cell marker *INS*, many Hybrid cells across all three subpopulations co-expressed another marker gene indicative of two independent canonical cell types (Fig. 3C, and fig. S15c).

To gain further insight into the molecular nature of Hybrid cells, we next compared each Hybrid subpopulation to relevant canonical cell populations to determine their transcriptional divergence (tables S51 to 57) and convergence with respect to gene signatures of common pancreatic cell types (*22, 28*). To this end, we observed the enrichment of the alpha cell signature in Endocrine-like Hybrids, the ductal cell signature in Ductal-like Hybrids, and the acinar cell signature in Acinar-like Hybrids. Conversely,

when each of the Hybrid subpopulations was compared to relevant canonical cell types, the beta cell signature was found to be significantly associated with the Hybrid subpopulations (Fig. 3D).

We next investigated the expression of key endocrine, ductal, and acinar transcription factors (TFs) (*29*) in Hybrid cells (tables S51 to S57). Compared to exocrine cells, Hybrid cells were found to have an enrichment of endocrine TF genes (positive $\log_2$ fold change) (fig. S15d). Conversely, compared to endocrine cells, Hybrid cells were found to have an enrichment of exocrine TF genes (fig. S15d). Furthermore, pseudotime analysis (*30*) of Hybrid cells amongst either exocrine-related or endocrine-related cells demonstrated that the Hybrid cell population is a transcriptionally distinct cellular state with an intermediary mixture of canonical cell types (fig. S16, a to f).

Finally, we assessed standard hallmark gene signatures to determine whether general pathways are enriched in Hybrid cells in comparison to canonical cell types. Across multiple comparisons, Hybrid cells were found to be enriched with the 'epithelial-mesenchymal transition' and 'apical surface' signatures (Fig. 3D and fig. S15e). Notably, the 'epithelial-mesenchymal transition' signature has been implicated in the conversion of ductal cells to endocrine cells (*31, 32*). Taken together, our findings suggest that Hybrid cells are a transcriptionally distinct population of cells across all donors with disparate, yet intermediary, transcriptional signatures of two canonical cell types.

Proteomic validation of Hybrid cells by mass cytometry

Although gene expression measurement by droplet-based scRNA-seq is unbiased, doublets are observed ~1% of the time. We used strict criteria and computationally removed potential doublets from scRNA-seq analysis using two independent methods (see Methods). Nevertheless, to address the possibility of doublet artifact in detection of Hybrid cells, we first employed cytometry by time-of-flight (CyTOF) (*33*). Considering the relatively low frequency of Hybrid cells predicted by scRNA-seq (~5%), we took an integrative approach and combined more than 6 million live single cells from 12 donors, which have also been profiled by scRNA-seq (4 Control, 4 AAB+, and 4 T1D donors). To ascertain if the Hybrid population exists using these deep phenotyping data, we scaled our analytical strategy to millions of cells, and clustered them by

expression levels of 33 proteins (table S58). Despite the inherent variability of antibody staining across donors that limited the integrative analysis of previous studies (*34*), we were able to annotate major cell types based on the expression of canonical proteins (fig. S17, a to e).

Remarkably, our strategy identified a cluster of rare cells representing the Ductal-like Hybrid cells previously identified in scRNA-seq, which are cytokeratin/C-peptide double-positive cells (Fig. 4A). The median frequency of Hybrid cells measured by CyTOF across all donors was around 2.7%, similar to the scRNA-seq results, and cells from Control donors constituted a significantly larger proportion of the Hybrid population (Fig. 4, B and C; p-value < 2.2e-16). Finally, the presence of cytokeratin/C-peptide Hybrid cell population was confirmed by two-parameter analysis for donors predicted by our integrative analysis to have the highest and lowest frequency of Hybrid cells (Fig. 4D and fig. S17o).

Imaging mass cytometry in intact tissues corroborates the presence of Hybrid cells

Having confirmed the existence of Hybrid cells by two orthogonal data modalities, we next sought to study the anatomical-spatial features of Hybrid cells in pancreatic tissues using imaging mass cytometry (IMC) (*5*). Although CyTOF and scRNA-seq rely on the profiling of dissociated cells, IMC removes this confounding effect by analyzing tissues fixed directly from the native human pancreas. We again amended our analytical pipeline to harness the expression levels of 35 proteins quantified by IMC in more than 1 million cells across 147 tissue slides from 19 donors, including 11 individuals not previously assessed by scRNA-seq or CyTOF for an independent validation of our findings (table S59).

IMC analysis not only discerned Ductal-like Hybrid cells amongst ductal and endocrine cells (Fig. 4E and fig. S17, f to j), but also quantified the frequency of Hybrid cells across donor groups (Fig. 4, F and G). Considering that pancreatic tissue sections contain much larger cellular heterogeneity than islet cultures, we found that Hybrid cells accounted for ~0.4% of all cells across donors (Fig. 4, F and G). Despite normalizing the proportion of Hybrid cells by the number of cells per donor, we found that Controls accounted for significant majority of the Hybrid population (Fig. 4 F-G; p-value < 2.2e-16).

We also found that the frequency of immune cells did not correlate with the frequency of Hybrid cells across donor groups (Fig. 4, A, B, C, E, F, and G, and fig. S17, a, d, e, f, i, j, k, and l).

Next, we leveraged the IMC map of spatial architecture of human islets and determined the location of Hybrid cells within the pancreas and interrogated their cellular neighborhood. We found Hybrid cells most frequently located in close proximity to beta and alpha cells, although some Hybrid cells were located near ductal cells (fig. S17, m and n; p-value < 1e-2). Interestingly, the frequency of Hybrid cells was highest in the tail region of the pancreas (Fig. 4H, fig. S17p). Taken together, our multimodal, single-cell profiling approach using scRNA-seq, CyTOF, and IMC enabled the identification and substantiation of non-canonical, rare Hybrid cells. Since the enrichment of Hybrid cells in a particular donor group was not consistent across three technologies, we conclude that Hybrid cells represent a general feature of pancreatic tissues.

MHC Class II expression is enriched in T1D ductal cells

The major genetic susceptibility determinants of T1D have been mapped to the MHC Class II genes (*35, 36*). We therefore sought to determine which cell types or donor states disproportionately express genes in this pathway. Using our scRNA-seq data, we found that genes associated with MHC Class II activity were enriched in Immune, Endothelial, Ductal, and Ductal-like Hybrid clusters (fig. S18, a to d). The lack of enrichment of the immune cell marker *PTPRC* or other genes associated with immune cells across the endocrine and ductal dendrogram supports that this finding is not due to immune cell contamination (fig. S4,a and b, fig. S18g). We next evaluated the expression of *HLA-DPB1*, an MHC class II gene associated with T1D risk, and *KRT19*, a ductal cell marker, across ductal and endocrine cell types. We identified five clusters with high *HLA-DPB1* and high *KRT19* expression, which accounted for 10.9% of all cells (7,588 cells) (Fig. 5, A and B, fig. S18, e and f). Strikingly, cells from T1D donors disproportionately contributed to the population of MHC Class II-expressing ductal cells (Fig. 5C; p-value < 2.2e-16). This observation is not due to sampling issues pertaining to the difficulty of isolating high purity islets from T1D donors. The Ductal-1 cell population consists of very similar numbers of Control and T1D donor ductal cells (4,217 and 4,154 cells, respectively);

however, there is a marked difference in the percentage of Control versus T1D MHC class II-expressing Ductal-1 cells, at 35% and 91%, respectively (Fig. 5C; p-value < 2.2e-16).

Multimodal confirmation of MHC Class II-expressing ductal cells in T1D

Based on our results from single-cell transcriptomic profiling of the human pancreas (Fig. 5), we surmised that MHC Class II molecules on ductal cells might be involved in antigen presentation in T1D. To validate this result, we employed two high-throughput technologies, CyTOF and IMC, in addition to immunofluorescence experiments in native pancreatic tissues. Using CyTOF, we identified a population of ductal cells expressing HLA-DR, an MHC Class II protein (Fig. 6A). Notably, we found that cells from T1D donors constituted the largest percentage of this cluster, in agreement with the findings from scRNA-seq (Fig. 6B; p-value < 2.2e-16). Furthermore, HLA-DR-expressing ductal cells made up a large percentage of total cells across individual T1D donors (Fig. 6C). A two parameter (cytokeratin and HLA-DR) analysis on all single cells analyzed by CyTOF further confirmed the presence of this double-positive population across multiple donors (Fig. 6D and fig. S19, a and b). Notably, these ductal cells did not express CD45, the hallmark of leukocytes (Fig. 6E).

To comprehensively study the existence of MHC Class II-expressing ductal cells in pancreatic tissues independent of islet cultures, we exploited the quantitative analysis of IMC data. By analyzing 147 tissue slides, we confirmed that MHC Class II-expressing ductal cells were predominately present in T1D donors (Fig. 6, F to H, fig. S20). MHC Class II-expressing ductal cells were located primarily in the tail and body regions of the pancreas (fig. S19c). Remarkably, the frequency of immune cells highly correlated with the frequency of MHC Class II expressing ductal cells, given that the T1D donor group had the highest percentage of both immune cells and HLA-DR-expressing ductal cells (fig S19, d and e). Finally, immunofluorescence staining (IF), followed by confocal microscopy verified the existence of MHC Class II-expressing cells in the pancreatic tissues of a Control and a T1D donor. We identified MHC Class II-expressing ductal cells in both donors; however, there was a more pronounced enrichment of MHC Class II-expressing ductal cells in the T1D pancreas (Fig. 6I). Cellular neighborhood analysis in pancreatic tissues further unveiled that HLA-DR-expressing ductal cells were surrounded by immune

cells, in particular $CD4^+$ immune cells and $CD11b^+$ dendritic cells (fig. S19, f and g; p-value < 1e-2). Together, our multi-modal single-cell measurements from transcriptomics to spatial proteomics in ductal cells provide novel insights into pancreas biology and suggest that ductal cells could be involved in the initial pathogenic events of T1D by functioning as antigen-presenting cells.

Discussion

Here, we employed scRNA-seq as a molecular microscope to investigate cellular diversity in pancreatic islets of T1D, AAb+, and non-diabetic human organ donors, comprehensively analyzing over 80,000 cells. Importantly, a web portal of the scRNA-seq data presented here is publicly available for exploration (*37*). We further leveraged high-throughput proteomic assays, including mass cytometry (CyTOF) and multiplexed histological analyses by IMC, to corroborate transcriptomic-based findings.

We found that AAb+ donors exhibit similar transcriptional changes to T1D donors in various endocrine and exocrine cells, despite these donors retaining normoglycemia. This observation suggests that even though AAb+ donors are clinically normal, various cell types of AAb+ donors have a molecular phenotype similar to those of T1D donors. Although it is impossible to discern at present whether these transcriptional changes are contributing to or are byproducts of disease pathogenesis, the mere discovery of molecular phenotypic changes in these cells in AAb+ individuals is novel and advances our understanding of early pancreatic perturbations occurring in T1D.

We identified a rare population of Hybrid cells, with expression signatures of beta, ductal, alpha, and acinar cells. Similar mixed signature cell types have been reported previously using single-cell transcriptomics (*22, 38*). Although the human fetal pancreas has been known to contain multihormonal cells (*39, 40*), as well as cells with dual endocrine-exocrine signatures (*22, 38*), these types of cells have only recently been documented in the adult stage in mice and humans (*22, 38, 41-43*). It is possible that these Hybrid cells in adulthood are multihormonal cells retained from early development or are formed when monohormonal cells differentiate into another cell type during adulthood. Most recently, it has been suggested that multipotent progenitor-like cells in humans defined as $PDX1^+/ALK3^+/CAII^-$ cells can differentiate into all pancreatic lineages

including endocrine cells (*44*). While cells annotated as Hybrid cells in our study did not show high expression levels of *ALK3* (*BMPR1A*) (fig. S18h), it is likely that large scale investigations of exocrine cells will shed more light onto the cellular heterogeneity and regenerative potential of this compartment.

The most striking finding arising from our study is that cells of the exocrine compartment show transcriptional and gene ontological changes in the T1D disease setting. Ductal cells from T1D donors, in a sharp contrast with non-diabetic or AAb+ donors, expressed high levels of MHC Class II and interferon pathways and were surrounded by immune cells. Although our study represents the first report of ductal cells expressing MHC Class II proteins in the T1D context, this finding is in accordance with previous literature documenting an elevation of immune cells in the exocrine pancreas of T1D donors (*5, 27, 45*) and regulation of MHC Class II genes by the interferon signaling pathway (*46*). Recent reports support a role for epithelial cells as facultative, non-professional antigen presenting cells in the gut and lung (*47*), and expression of MHC Class II proteins in non-lymphoid cells in the pancreas (*48*). Notably, two HLA genes with high expression in ductal cells, *HLA-DRB1* and *HLA-DPB1*, are major T1D susceptibility genes (*35*). Together, our study unmasks exocrine ductal cells as potential pathogenic cells in T1D and suggest a new cellular context for examining the role of genetics in T1D.

Acknowledgments

We thank our colleagues for helpful discussions, particularly: Aditi Chandra, Eline Luning Prak, Ben Stanger, Michael Silverman, Gregory Beatty, Ken Zaret, Mitch Lazar, and E. John Wherry. We thank Andrei Georgescu for confocal microscopy and the University of Pennsylvania Diabetes Research Center (DRC) for the use of the Functional Genomics Core (P30-DK19525). This work was supported by NIH grants UC4 DK112217 (to A.N., K.K., M.B., J.M., M.F., and G.V.), R01CA230800 and Susan G. Komen CCR185472448 (to R.B.F.), and R01HL145754, U01DK127768, the Burroughs Wellcome Fund, the Chan Zuckerberg Initiative, the Penn Epigenetics and the Sloan Foundation awards to G.V.

Materials and Methods

Experimental model and subject details

Pancreatic islets were procured by the HPAP consortium under the Human Islet Research Network (https://hirnetwork.org/) with approval from the University of Florida Institutional Review Board (IRB # 201600029) and the United Network for Organ Sharing (UNOS). A legal representative for each donor provided informed consent prior to organ retrieval. For T1D diagnosis, medical charts were reviewed and C-peptide levels were measured in accordance with the American Diabetes Association guidelines (American Diabetes Association 2009). All donors were screened for autoantibodies prior to organ harvest, and AAb positivity was confirmed again post tissue processing and islet isolation.

Organs were recovered and processed as previously described (*7*). Table 1 summarizes donor information. Pancreatic islets were cultured and dissociated into single cells as previously described (*22*). Total dissociated cells were used for single cell capture for each of the donors, except AAB+ donor #1 (HPAP019), which was enriched for beta cells as previously reported prior to capture (*49*).

scRNA-seq islet capture, sequencing, and processing

The Single Cell 3' Reagent Kit v2 or v3 was used for generating scRNA-seq data. 3,000 cells were targeted for recovery per donor. All libraries were validated for quality and size distribution using a BioAnalyzer 2100 (Agilent) and quantified using Kapa (Illumina). For samples prepared using 'The Single Cell 3' Reagent Kit v2', the following chemistry was performed on an Illumina HiSeq4000: Read 1: 26 cycles, i7 Index: 8 cycles, i5 index: 0 cycles, and Read 2: 98 cycles. For samples prepared using 'The Single Cell 3' Reagent Kit v3', the following chemistry was performed on an Illumina HiSeq 4000: Read 1: 28 cycles, i7 Index: 8 cycles, i5 index: 0 cycles, and Read 2: 91 cycles. Cell Ranger (10x Genomics; v3.0.1) was used for bcl2fastq conversion, aligning (using the hg38 reference genome), filtering, counting, cell calling, and aggregating (--normalize=none) (fig. S1, a and b).

scRNA-seq clustering, doublet removal, & cell type classification

Seurat v3.1.5 (*19, 50*) was used for filtering, UMAP generation, and initial clustering. Genes expressed in at least 3 cells were kept, as were cells with at least 200 genes. nFeature, nCount, percent.mt, nFeature vs nCount, and percent.mt vs nCount plots were generated to ascertain the lenient filtering criteria of 200 > nFeature < 8,750, percent.mt < 25, and nCount <125,000. Data was then log normalized, and the top 2,000 variable genes were detected using the "vst" selection method. The data was then linearly transformed ("scaled"), meaning that for each gene, the mean expression across cells is 0 and the variance across cells is 1. Principle component analysis (PCA) was then carried out on the scaled data, using the 2,000 variable genes as input. We employed two approaches to determine the dimensionality of the data, i.e. how many principal components to choose when clustering: (1) a Jackstraw-inspired resampling test that compares the distribution of p-values of each principle component (PC) against a null distribution and (2) an elbow plot that displays the standard deviation explained by each principal component. Based on these two approaches, 17 PCs with a resolution of 1.2 were used to cluster the cells, and non-linear dimensionality reduction (UMAP) was used with 17 PCs to visualize the dataset (fig. S1, c and e).

Two independent methods were used to detect and remove doublets. DoubletFinder v2.0 (*51*) was used to demarcate potential doublets in the data as previously described, with the following details: 17 PCs were used for pK identification (no ground-truth) and the following parameters were used when running doubletFinder_v3: PCs = 17, pN = 0.25, pK = 0.0725, nExp = nExp_poi, reuse.pANN = FALSE, and sct = FALSE (fig. S1d). Scrublet v0.2.1 (*18*) was also used to demarcate potential doublets (fig. S2a). Given that a small percentage of cells were demarcated as doublets by both methods, we removed all cells that were flagged as doublet by both or either approach, leading to the removal of 3,770 cells (fig. S2, b and c).

Following doublet removal, the raw data for the remaining 73,235 cells were filtered using the following criteria, which resulted in 69,645 cells remaining: 200 > nFeature < 8,750, percent.mt < 25, and nCount <100,000 (fig. S2c and fig. S3a). The data were log normalized, the top 2,000 variable genes were detected, the data underwent linear transformation, and PCA was carried out, as described above. Both the Jackstraw-

inspired resampling test and an elbow plot of standard deviation explained by each principal component were used to determine the optimal dimensionality of the data, as described above. Based on these two approaches, 26 PCs with a resolution of 1.2 was used to cluster the cells, and UMAP was used with 26 PCs to visualize the 49 clusters detected (fig.S3, b and c).

Garnett was used for initial cell classification as previously described (*9*). In brief, a cell type marker file (table S60) with 17 different cell types was compiled using various resources (*18-22, 28*), and this marker file was checked for specificity using the "check_markers" function in Garnett by checking the ambiguity score and the relative number of cells for each cell type. A classifier was then trained using the marker file, with "num_unknown" set to 500, and this classifier was then used to classify cells and cell type assignments were extended to nearby cells, "clustering-extended type" (Louvain clustering) (fig. S3d). Upon inspection of cluster purity using canonical gene markers of the major pancreatic cell types, we found that the abundant and transcriptionally distinct cell types form generally distinct and unique clusters: beta cells (*INS* high), alpha cells (*GCG* high), acinar cells (*CPA1* high), ductal cells (*KRT19* high), endothelial cells (*VWF* high) stellate cells (*RSG10* high), and immune cells (*PTPRC,* also known as CD45 or leukocyte common antigen, high) (fig. S3e). In contrast, the rarer and/or less transcriptionally distinct cell types did not clearly segregate, namely delta cells (*SST* high), PP cells (*PPY* high), and epsilon cells (*GHRL* high) (fig. S3e).

To overcome the apparent limitation in grouping the major canonical cell types, we employed the analytical workflow termed 'TooManyCells' (*10*), which implements an efficient divisive hierarchical spectral clustering approach along with tree visualizations to preserve the relationships among cell clusters at varying resolutions. In conjunction with 'TooManyCells' clustering, we invoked the cellular classifier Garnett (*9*), which annotates cell types by training a regression-based classifier from user-provided cell type signatures (table S60; (*34*)). Briefly, for the clustering of all cells, the raw data from the 69,645 cells were normalized by total count and gene normalization by median count (TotalMedNorm) followed by term frequency-inverse document frequency (tf-idf) for clustering. For visualization of the comprehensive clustering, the dendrogram was first pruned using the TooManyCells flags '--min-distance-search "15"' and "--smart-cutoff "15"', followed by

pruning using the flag '--max-step 6'. TooManyCells enabled distinct clustering of all major known cell types associated with pancreatic islets, as confirmed by Garnett cell labels and the inspection of canonical gene markers for each cell type (Fig. 1C and fig. S4a).

The raw data from different cell types were then subsetted from the comprehensive clustering in Figure 1C in order to cluster cells on a cell-type basis. For the clustering of ductal/endocrine cells, data from the ductal/endocrine cell clusters from the comprehensive tree were subsetted and normalized by TotalMedNorm followed by term tf-idf. For visualization of the ductal/endocrine tree, the dendrogram was first pruned using the TooManyCells flags '--min-distance-search "7"' and "--smart-cutoff "7"' followed by pruning using the flag '--max-step 7'. For the clustering of immune cells, data from the immune cell cluster from the comprehensive tree were subsetted and normalized by TotalMedNorm followed by tf-idf. For visualization of the immune tree, the dendrogram was first pruned using the TooManyCells flag '--max-step 4'. When individual genes were painted across any of the dendrograms, 'TotalMedNorm' was employed to normalize gene expression.

Differential Gene Expression, GSEA analysis, and Metascape analysis

Differential genes were found using edgeR through TooManyCells with the –normalization "NoneNorm" to invoke edgeR single cell preprocessing, including normalization and filtering. For Metascape analysis (http://metascape.org/gp/index.html#/main/step1; (*52*)), less than or equal to 3,000 differential genes (FDR< 0.05 and fold change (FC) > 0.1) were subjected to analysis. The top 20 clusters are displayed and a stringent cut-off of 1e-6 was applied to determine significant gene ontology pathways. For gene-set-enrichment-analysis (GSEA) analysis, GSEA Preranked (4.0.1) (*53*) was run on a pre-ranked gene list using either user-provided pancreatic gene expression sets (*22, 28*) or standard hallmark gene signatures provided by the Molecular Signatures Database (MSigDB) (*54*).

Hybrid cell co-expression, DE analysis and heatmaps

With the differentially expressed genes (FDR< 0.05 and fold change (FC) > 0.1) between every two sample groups, we calculated the shared and unique genes in each

cell type, and visualized it with Venn diagrams. The expression levels of the genes in each cell of the three groups were extracted from the median normalized count matrix. Then we aggregate the expression levels in each group by taking the average value of the normalized counts. The mean expression values of the three groups were further normalized by the total expression level of that gene. We visualized the normalized expression level of differential genes with heatmaps.

To examine the co-expression of signature genes of some cell types, we normalized the median normalized matrix with log2(N +1). Then we selected the matrix of selected cell types by marker genes. The distribution of the cells from selected cell types by expression level of two marker genes were shown with geom_density_2d_filled() in ggplot2 package of R.

Transcription Factor analysis

A list of human transcription factors was downloaded (*55*), and the expression of key transcription factors (TFs) of endocrine, ductal, and acinar cell fates (*29*) were assessed in Hybrid cells compared to ductal, acinar, beta, or alpha cells (tables S51 to S57).

Monocle Analysis

Ductal, Alpha Major, Beta Major, Hybrid, Alpha/Beta, and Beta Minor cells were analyzed using the standard Monocle3 pipeline (*30, 56, 57*). Only genes within the pancreatic gene expression set (*22, 28*) were used for this analysis.

CyTOF data collection, input files, and preprocessing

Flow CyTOF was performed as described previously (*22*). Briefly, after isolating the dissociated cells, barcoding was conducted for donors following the manufacturer's protocol (Fluidigm, 101-0804 B1). Following barcoding, metal-conjugated antibody labeling was carried out in 'FoxP3 permeabilization buffer' (eBioscience, 00-8333) with 1% FBS (Hyclone, Cat# 7207) for 12 hours at 4°C at a concentration of up to 3 million cells per 300 µl of antibody cocktail, followed by twice washing with FoxP3 permeabilization buffer. Cells were then incubated with the DNA intercalator Iridium

(Fluidigm, 201192A) at a dilution of 1:4,000 in 2% paraformaldehyde (Electron Microscopy Sciences, 15714) in DPBS (Corning, 21-031-CV) at RT for 1 hr. Mass cytometry data were acquired on CyTOF (Fluidigm). Flow CyTOF data analyses of endocrine cell composition was performed using the Cytobank implement (https://www.cytobank.org/).

Normalized FCS files were pre-processed prior to TooManyCells analysis and visualization using FlowJo Version 10.6.1 by gating all events on singlets according to event length and DNA content and then on live cells based on cisplatin exclusion. The Singlet/Live gated population was exported to a CSV file for TooManyCells analysis. Two dimensional plots were visualized for combinations of individual channels.

TooManyCells clustering for CyTOF

TooManyCells was used to generate cell clades of CyTOF data. Cells with less than a total of 1e-16 signal were removed, leaving 6,945,575 cells. Upon inspection of protein levels across a tree with all cells (fig S17 a to c), endocrine and exocrine compartments were further subsetted leading to a refined analysis of 4,521,988 cells (Fig. 4A and 6A). Quantile normalization of the raw counts was used in the clustering step. The resulting tree was pruned by collapsing nodes with less than (7 MAD X median # cells in nodes) cells within them into their parent nodes.

Imaging mass cytometry (IMC) analysis and Cell Segmentation

IMC was performed as described previously (*5*). Cell segmentation of all images was performed with the Vis software package (Visiopharm). All image channels were pre-processed with a 3x3 pixel median filter, then cells were segmented by applying a polynomial local linear parameter-based blob filter to the Iridium-193 DNA channel of each image to select objects representing individual nuclei. Identified nuclear objects were restricted to those greater than 10μm^2, then dilated up to 7 pixels to approximate cell boundaries. Per-cell object mean pixel intensities were then exported for further analysis.

TooManyCells clustering for IMC

TooManyCells was used to generate cell clades of IMC data. Cells with less than a total of $1e^{-16}$ signal were removed. Upon inspection of protein levels across a tree with all 1,170,001 cells (fig S17 f to h), endocrine and exocrine compartments were further subsetted, leading to the refined analysis of 130,428 cells for the investigation of the hybrid cell frequency (Fig. 4E). The full tree with 1,170,001 cells was used for the assessment of HLA-DR-expressing ductal cells (Fig 6F). Quantile normalization of the raw counts was used in the clustering step. The resulting tree was pruned by collapsing nodes with less than (5 MAD X median # cells in nodes) cells within them into their parent nodes. Subsetting of the tree was done with "--root-cut 3" to focus on node 3 in relevant analyses, with additional pruning of (3 MAD X median # cells in nodes).

Cell-neighborhood analysis for IMC

Three labels were given to cells in the IMC neighborhood analysis: base, neighbor, and distant. Base cells originated from the chosen node, here node 16 in the node 3-focused IMC tree, or node 10 in the complete pruned tree which includes the former node 16. Given the x- and y-coordinates from IMC per cell, each cell's Euclidean distance to a base cell was calculated. If that distance was less than or equal to the chosen value, here 20 for the complete pruned tree, the cell was assigned the neighbor label. Otherwise the cell was designated as distant.

IF and confocal microscopy

Tissues were fixed in 10% buffered formalin overnight, washed several times in PBS, then dehydrated through series of ethanol, embedded in paraffin and sectioned to 4-8um. Following deparaffinization through xylene and sequential rehydration, slides were subjected to heat antigen retrieval in a pressure cooker with Bulls Eye Decloaking buffer (Biocare). Slides were incubated in primary antibody overnight and secondary antibody conjugated to peroxidase and then developed using Tyramide Signal Amplification (TSA, Akoya Biosciences). Slides were counter stained with DAPI, mounted and imaged on Zeiss LSM800. Primary antibodies used for staining were Mouse anti-CK19 (Santa Cruz sc-6278) and Rabbit anti-HLA-DR (Abcam ab92511).

Statistical analysis of box plots with Control, AAB+, and T1D donor states

The D'Agostino & Pearson omnibus normality test was used to assess whether the data from each group was normally distributed. If any group failed the D'Agostino & Pearson omnibus normality test, the Kruskal-Wallis test was applied. If none of the groups failed the D'Agostino & Pearson omnibus normality test, the one-way ANOVA test was applied.

Statistical analysis of cellular neighborhoods

Differential marker expression significance for neighbors in the IMC analysis was determined using permutation tests. For each marker, the distribution of that marker value for each of the designated n neighbors was compared against 100 distributions derived from n random cells across the entire IMC tree. The resulting p-value was calculated by the ratio of the number of permutations that had a lower median marker value than the observed marker value to the total number of permutations. If this value was < 0.5, the value was subtracted from 1 to switch directionality (number of permutations with a higher median value). To account for the two-tailed test, this value was multiplied by 2 for the final p-value calculation.

Figure 1

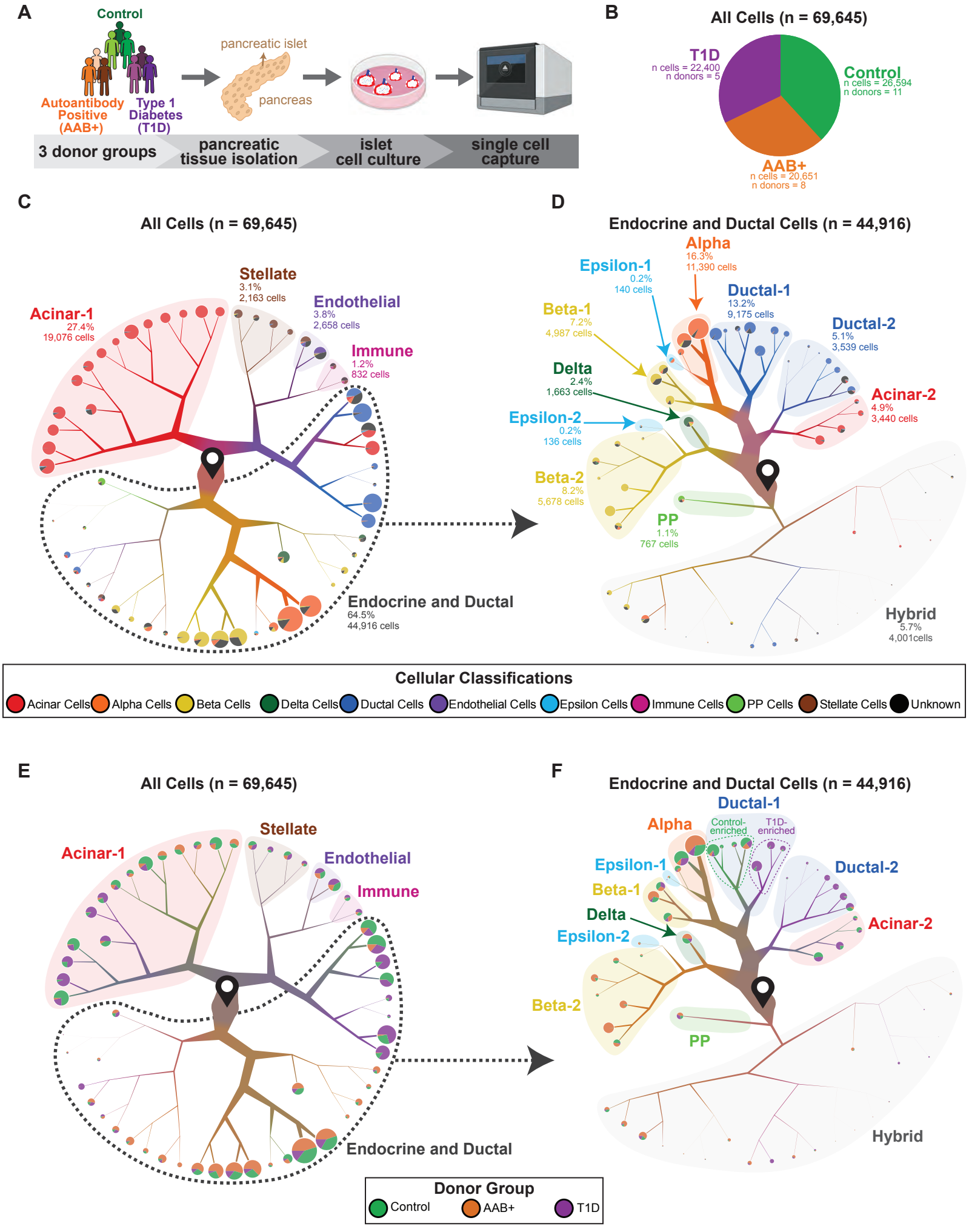

A
Control
Autoantibody Positive (AAB+)
Type 1 Diabetes (T1D)
pancreatic islet
pancreas
3 donor groups
pancreatic tissue isolation
islet cell culture
single cell capture
B
All Cells (n = 69,645)
T1D
n cells = 22,400
n donors = 5
Control
n cells = 26,594
n donors = 11
AAB+
n cells = 20,651
n donors = 8
C
All Cells (n = 69,645)
Stellate
3.1%
2,163 cells
Endothelial
3.8%
2,658 cells
Acinar-1
27.4%
19,076 cells
Immune
1.2%
832 cells
Endocrine and Ductal
64.5%
44,916 cells
D
Endocrine and Ductal Cells (n = 44,916)
Alpha
16.3%
11,390 cells
Epsilon-1
0.2%
140 cells
Ductal-1
13.2%
9,175 cells
Beta-1
7.2%
4,987 cells
Ductal-2
5.1%
3,539 cells
Delta
2.4%
1,663 cells
Acinar-2
4.9%
3,440 cells
Epsilon-2
0.2%
136 cells
Beta-2
8.2%
5,678 cells
PP
1.1%
767 cells
Hybrid
5.7%
4,001cells
Cellular Classifications
Acinar Cells
Alpha Cells
Beta Cells
Delta Cells
Ductal Cells
Endothelial Cells
Epsilon Cells
Immune Cells
PP Cells
Stellate Cells
Unknown
E
All Cells (n = 69,645)
Stellate
Acinar-1
Endothelial
Immune
Endocrine and Ductal
F
Endocrine and Ductal Cells (n = 44,916)
Ductal-1
Alpha
Control-enriched
T1D-enriched
Epsilon-1
Ductal-2
Beta-1
Delta
Acinar-2
Epsilon-2
Beta-2
PP
Hybrid
Donor Group
Control
AAB+
T1D

Figure Legends

Figure 1: Discernment of human pancreatic cell types using single-cell RNA-seq

(A) The transcriptome of single cells from pancreatic islets of 3 donor types (health Control donors, autoantibody positive (AAB+) donors, and donors with Type 1 diabetes (T1D)) was ascertained using the 10x Genomics platform.

(B) Pie chart displaying the proportion of cells comprised by each donor group.

(C) TooManyCell dendrogram visualization and clustering of all cells. Cells begin at the start pin symbol, and there are then partitioned based on transcriptional similarities and differences. The color within the branches indicates the proportion of the cells that are classified by the Garnett cellular classification tool. Each bifurcation denotes significant transcriptional differences between the two cell groups. Pie charts at the end of the branches display the breakdown of Garnett cellular classification of cells within that terminal cluster. Highlighting or dotted lines surrounding particular clusters of cells with labels define cell types based on Garnett cellular classifications and canonical gene expression (fig. S4a). Branch thickness and piechart size is proportional to cell number. Branch length is not indicative of any factor, but is merely a means by which to display cells within a defined space. Beta cells (*INS* high), alpha cells (*GCG* high), delta cells (*SST* high), PP cells (*PPY* high), epsilon cells (*GHRL* high), acinar cells (*CPA1* high), ductal cells (*KRT19* high), endothelial cells (*VWF* high), stellate cells (*RSG10* high), and immune cells (*PTPRC,* also known as CD45 or leukocyte common antigen, high).

(D) Dendrogram visualization and clustering of ductal and endocrine cells. Highlighting or dotted lines surrounding particular clusters of cells with labels define cell types based on Garnett cellular classifications and canonical gene expression (fig. S4b).

(E) Group donor type projected across the dendrogram visualization and clustering of all cells from Figure 1C. Pie charts at the end of the branches display the breakdown donor type within that terminal cluster.

(F) Group donor type projected across the dendrogram visualization and clustering of endocrine and ductal cells from Figure 1D. Pie charts at the end of the branches display the breakdown of donor type within that terminal cluster.

Figure 2

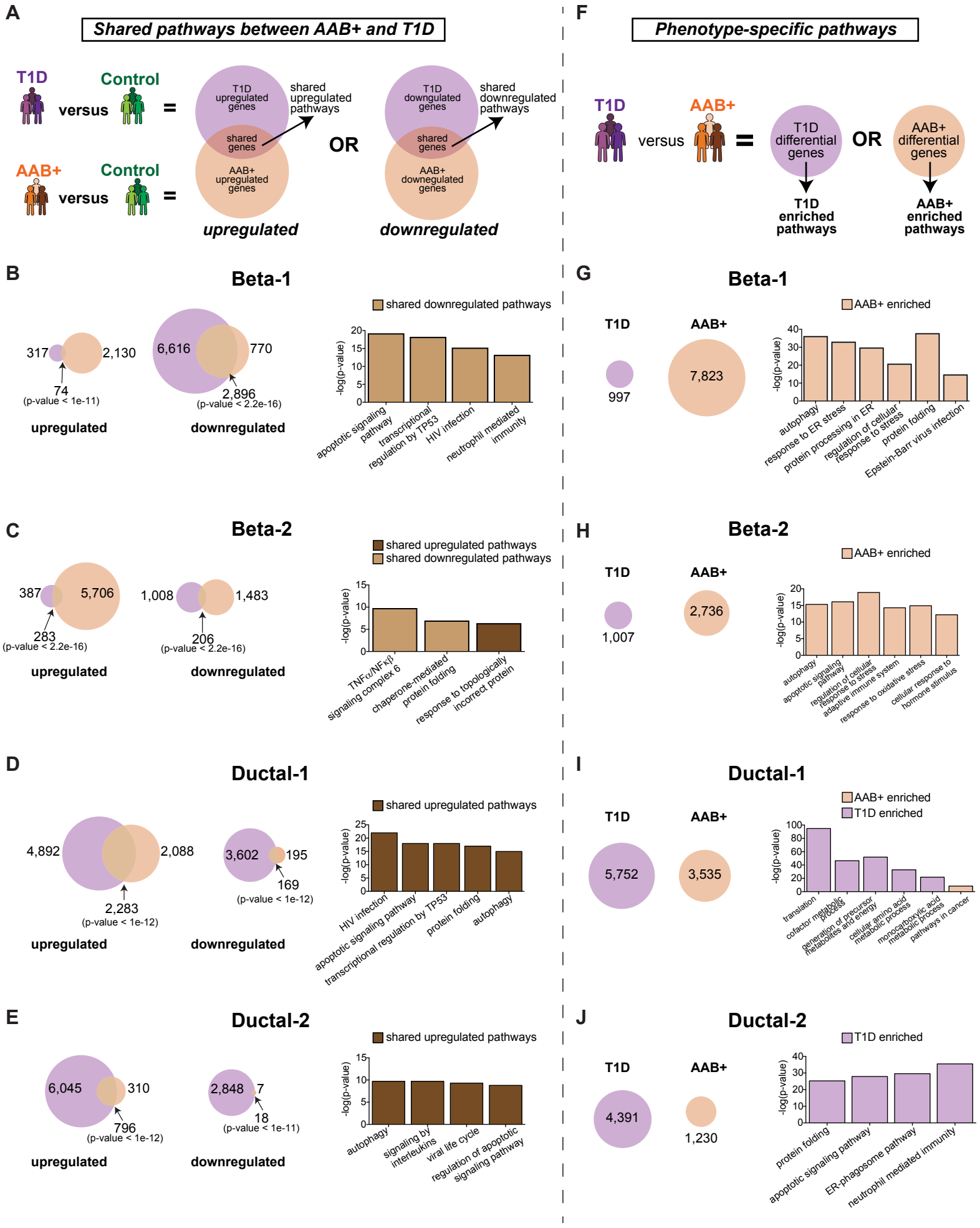

A
Shared pathways between AAB+ and T1D
T1D
versus
Control
=
AAB+
versus
Control
=
T1D upregulated genes
shared genes
AAB+ upregulated genes
shared upregulated pathways
OR
T1D downgulated genes
shared genes
AAB+ downegulated genes
shared downregulated pathways
upregulated
downregulated
F
Phenotype-specific pathways
T1D
versus
AAB+
=
T1D differential genes
OR
AAB+ differential genes
T1D enriched pathways
AAB+ enriched pathways
B
Beta-1
317
2,130
74
(p-value < 1e-11)
upregulated
6,616
770
2,896
(p-value < 2.2e-16)
downregulated
shared downregulated pathways
-log(p-value)
apoptotic signaling pathway
transcriptional regulation by TP53
HIV infection
neutrophil mediated immunity
G
Beta-1
T1D
AAB+
997
7,823
AAB+ enriched
-log(p-value)
autophagy
response to ER stress
protein processing in ER
regulation of cellular response to stress
protein folding
Epstein-Barr virus infection
C
Beta-2
387
5,706
283
(p-value < 2.2e-16)
upregulated
1,008
1,483
206
(p-value < 2.2e-16)
downregulated
shared upregulated pathways
shared downregulated pathways
-log(p-value)
TNFα/NFκβ signaling complex 6
chaperone-mediated protein folding
response to topologically incorrect protein
H
Beta-2
T1D
AAB+
1,007
2,736
AAB+ enriched
-log(p-value)
autophagy
apoptotic signaling pathway
regulation of cellular response to stress
adaptive immune system
response to oxidative stress
cellular response to hormone stimulus
D
Ductal-1
4,892
2,088
2,283
(p-value < 1e-12)
upregulated
3,602
195
169
(p-value < 1e-12)
downregulated
shared upregulated pathways
-log(p-value)
HIV infection
apoptotic signaling pathway
transcriptional regulation by TP53
protein folding
autophagy
I
Ductal-1
T1D
AAB+
5,752
3,535
AAB+ enriched
T1D enriched
-log(p-value)
translation
cofactor metabolic process
generation of precursor metabolites and energy
cellular amino acid metabolic process
monocarboxylic acid metabolic process
pathways in cancer
E
Ductal-2
6,045
310
796
(p-value < 1e-12)
upregulated
2,848
7
18
(p-value < 1e-11)
downregulated
shared upregulated pathways
-log(p-value)
autophagy
signaling by interleukins
viral life cycle
regulation of apoptotic signaling pathway
J
Ductal-2
T1D
AAB+
4,391
1,230
T1D enriched
-log(p-value)
protein folding
apoptotic signaling pathway
ER-phagosome pathway
neutrophil mediated immunity

Figure 2: AAB+ and T1D donors have both common and distinct transcriptomic changes in endocrine and exocrine cell types

(A) For each cell type, two pairwise differential comparisons were carried out: (1) T1D versus Control (referred to as 'T1D upregulated' (T1D/Control) or 'T1D downregulated' (Control/T1D)) and (2) AAB+ versus Control (referred to as 'AAB+ upregulated' (AAB+/Control) or 'AAB+ downregulated' (Control/AAB+)). T1D upregulated genes were then compared to AAB+ upregulated genes to find commonly upregulated genes, and subsequently commonly upregulated gene ontologies and pathways, across these two donor groups; this exact same approach was carried out for downregulated genes as well.

(B-E) (Left) For each cell type, Venn diagrams indicate the numbers of upregulated and downregulated genes, as well as overlapping genes, across the two donor states. (Right) Bar graph displaying notable gene ontologies that are shared across disease states for upregulated and downregulated genes. P-values presented are the results of hypergeometric CDF tests (one-tailed test for overrepresentation).

(F) Transcriptional differences between cells from T1D and AAB+ donors were determined by directly comparing T1D to AAB+ cells to generate lists of differentially expressed genes that are enriched in T1D cells or AAB+ cells, and enriched gene ontology pathways were discovered from these differential gene lists.

(G-J) (Left) For each cell type, circles indicate the numbers of genes that are 'T1D enriched' or 'AAB enriched'. (Right) Bar graph displaying notable gene ontologies that are enriched for each donor state.

Figure 3

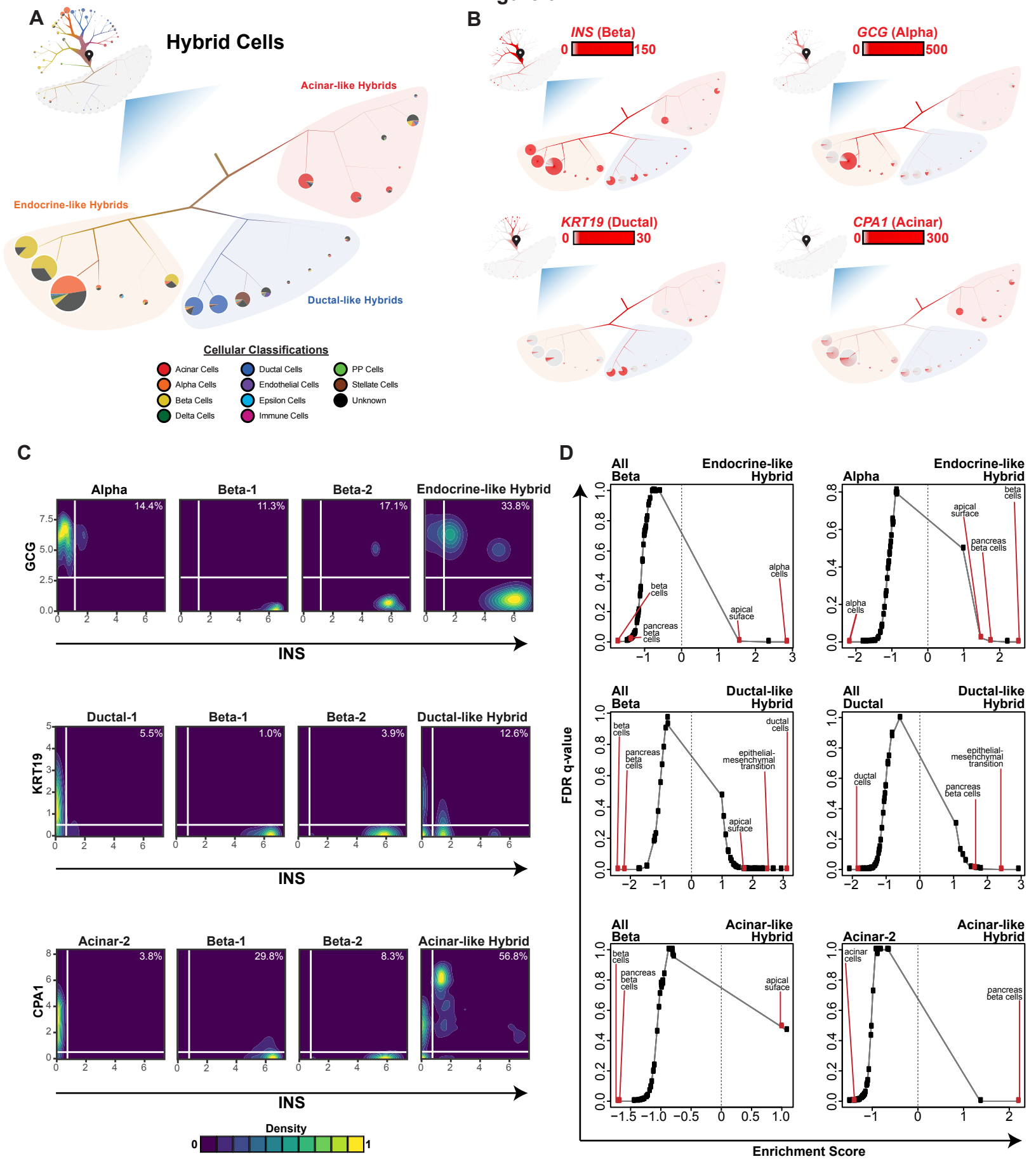

A
Hybrid Cells
Acinar-like Hybrids
Endocrine-like Hybrids
Ductal-like Hybrids
Cellular Classifications
Acinar Cells
Alpha Cells
Beta Cells
Delta Cells
Ductal Cells
Endothelial Cells
Epsilon Cells
Immune Cells
PP Cells
Stellate Cells
Unknown
B
INS (Beta)
0 150
GCG (Alpha)
0 500
KRT19 (Ductal)
0 30
CPA1 (Acinar)
0 300
C
Alpha
Beta-1
Beta-2
Endocrine-like Hybrid
14.4%
11.3%
17.1%
33.8%
GCG
INS
Ductal-1
Beta-1
Beta-2
Ductal-like Hybrid
5.5%
1.0%
3.9%
12.6%
KRT19
INS
Acinar-2
Beta-1
Beta-2
Acinar-like Hybrid
3.8%
29.8%
8.3%
56.8%
CPA1
INS
Density
0
1
D
All Beta
Endocrine-like Hybrid
beta cells
pancreas beta cells
alpha cells
apical suface
Alpha
Endocrine-like Hybrid
alpha cells
apical surface
pancreas beta cells
beta cells
All Beta
Ductal-like Hybrid
beta cells
pancreas beta cells
epithelial-mesenchymal transition
apical suface
ductal cells
All Ductal
Ductal-like Hybrid
ductal cells
epithelial-mesenchymal transition
pancreas beta cells
All Beta
Acinar-like Hybrid
beta cells
pancreas beta cells
apical suface
Acinar-2
Acinar-like Hybrid
acinar cells
pancreas beta cells
FDR q-value
Enrichment Score

Figure 3: Single cell RNA-seq profiling identifies non-canonical, rare Hybrid cells

(A) (Inset) Dendrogram visualization and clustering of ductal and endocrine cells from Figure 1D. Highlighted in grey are Hybrid cells. (Main) Magnified view of the Hybrid cell population from the ductal and endocrine dendrogram. The color within the branches indicates the proportion of the cells that are classified by the Garnett cellular classifier. Cells begin at the start pin symbol, and from there are partitioned based on transcriptional similarities and differences. Pie charts at the end of the branches display the breakdown of Garnett cellular classification of cells within that terminal cluster. Branch thickness and pie chart size is proportional to cell number. Branch length is not indicative of any factor, but is merely a means by which to display cells within a defined space.

(B) Dendrogram visualization of the expression of each canonical gene marker for each major cell type across the dendrograms from A. Scale bars represent normalized transcript numbers.

(C) Co-expression analysis of *INS* and the relevant gene indicative of each subpopulation of Hybrid cells (*KRT19* for Ductal-like Hybrids, *GCG* for Endocrine-like Hybrids, and *CPA1* for Acinar-like Hybrids) at the single-cell level across each Hybrid subpopulation and related canonical cell types. For each gene, values represent normalized transcript numbers. Percentages in the upper right-hand corner indicate the percentage of co-expressing cells.

(D) GSEA analysis plots of FDR q-value vs Normalized Enrichment Score. Each Hybrid cell subpopulation was compared to relevant canonical cell populations to determine differentially enriched gene sets. Labels in red demarcated signatures of interest.

Figure 4

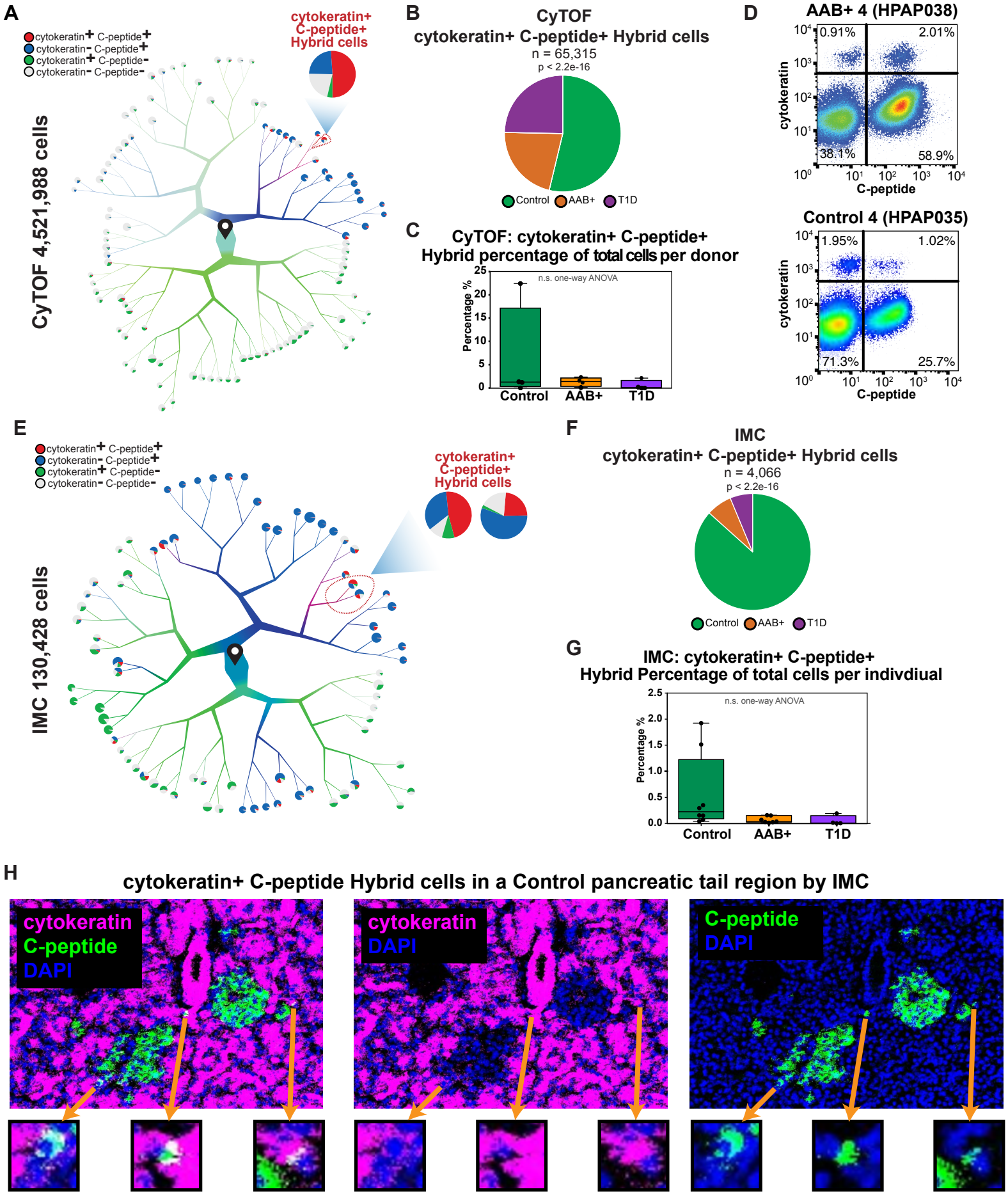

A
cytokeratin+ C-peptide+
cytokeratin- C-peptide+
cytokeratin+ C-peptide-
cytokeratin- C-peptide-
cytokeratin+ C-peptide+ Hybrid cells
CyTOF 4,521,988 cells
B
CyTOF
cytokeratin+ C-peptide+ Hybrid cells
n = 65,315
p < 2.2e-16
Control
AAB+
T1D
C
CyTOF: cytokeratin+ C-peptide+
Hybrid percentage of total cells per donor
n.s. one-way ANOVA
Percentage %
Control
AAB+
T1D
D
AAB+ 4 (HPAP038)
0.91%
2.01%
38.1%
58.9%
cytokeratin
C-peptide
Control 4 (HPAP035)
1.95%
1.02%
71.3%
25.7%
cytokeratin
C-peptide
E
cytokeratin+ C-peptide+
cytokeratin- C-peptide+
cytokeratin+ C-peptide-
cytokeratin- C-peptide-
cytokeratin+ C-peptide+ Hybrid cells
IMC 130,428 cells
F
IMC
cytokeratin+ C-peptide+ Hybrid cells
n = 4,066
p < 2.2e-16
Control
AAB+
T1D
G
IMC: cytokeratin+ C-peptide+
Hybrid Percentage of total cells per indivdiual
n.s. one-way ANOVA
Percentage %
Control
AAB+
T1D
H
cytokeratin+ C-peptide Hybrid cells in a Control pancreatic tail region by IMC
cytokeratin
C-peptide
DAPI
cytokeratin
DAPI
C-peptide
DAPI

Figure 4: Flow cytometry by time-of-flight (CyTOF) and imaging mass cytometry (IMC) validates the presence of Hybrid cells

(A) Dendrogram visualization of co-expression of cytokeratin and C-peptide proteins in individual cells analyzed with flow cytometry by time-of-flight (CyTOF). Cells begin at the start pin symbol, and from there are partitioned based on similarities and differences in protein markers.

(B) Pie chart displaying cytokeratin+ C-peptide+ Hybrid cells and the relative proportions of each donor group from the CyTOF data. P-values are the result of the Chi-squared test.

(C) Box plots displaying Hybrid cell percentage of total cells per individual across donor groups from the CyTOF data.

(D) Two parameter CyTOF analysis of cytokeratin and C-peptide expression in single cells from AAB+ donor #4 (HPAP038) and Control donor #4 (HPAP035).

(E) Dendrogram visualization of co-expression of cytokeratin and C-peptide proteins in individual cells analyzed by imaging mass cytometry (IMC). Cells begin at the start pin symbol, and from there are partitioned based on similarities and differences in protein markers.

(F) Pie chart displaying cytokeratin+ C-peptide+ Hybrid cells and the relative proportions of each donor group from the IMC data. P-values are the result of the Chi-squared test.

(G) Box plots displaying Hybrid cell percentage of total cells per individual across donor groups from the IMC data.

(H) Representative region of interest (ROI) from the tail pancreatic region of a Control donor displaying three cytokeratin+C-peptide+ Hybrid cells.

Figure 5

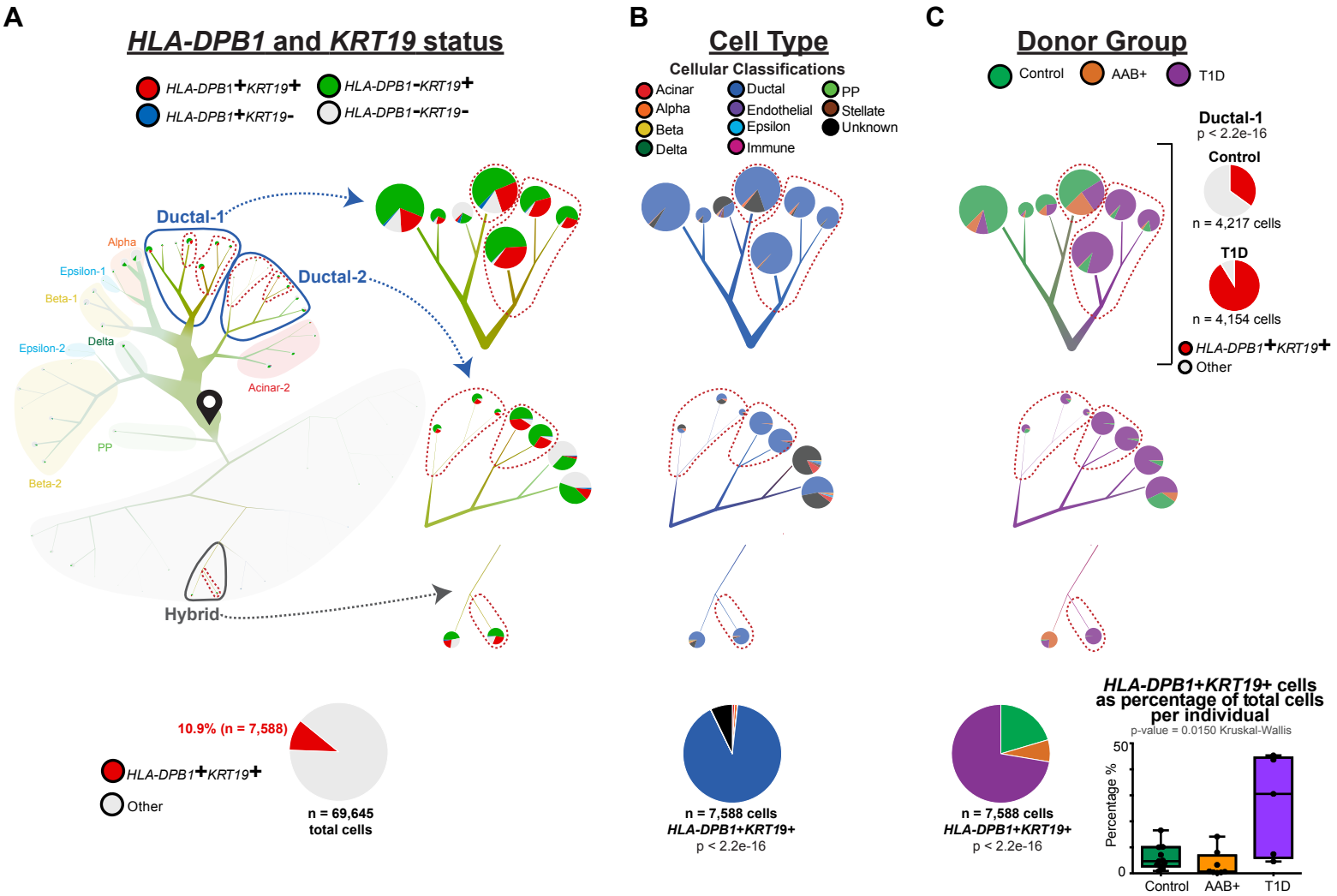

A
HLA-DPB1 and KRT19 status
HLA-DPB1+KRT19+
HLA-DPB1-KRT19+
HLA-DPB1+KRT19-
HLA-DPB1-KRT19-
Ductal-1
Ductal-2
Alpha
Epsilon-1
Beta-1
Epsilon-2
Delta
Acinar-2
PP
Beta-2
Hybrid
10.9% (n = 7,588)
HLA-DPB1+KRT19+
Other
n = 69,645
total cells
B
Cell Type
Cellular Classifications
Acinar
Alpha
Beta
Delta
Ductal
Endothelial
Epsilon
Immune
PP
Stellate
Unknown
n = 7,588 cells
HLA-DPB1+KRT19+
p < 2.2e-16
C
Donor Group
Control
AAB+
T1D
Ductal-1
p < 2.2e-16
Control
n = 4,217 cells
T1D
n = 4,154 cells
HLA-DPB1+KRT19+
Other
n = 7,588 cells
HLA-DPB1+KRT19+
p < 2.2e-16
HLA-DPB1+KRT19+ cells
as percentage of total cells
per individual
p-value = 0.0150 Kruskal-Wallis
Percentage %
50
0
Control
AAB+
T1D

Figure 5: Single cell RNA-seq profiling enables the identification of MHC Class II expressing ductal cells enriched in T1D

(A) (Left Top) Dendrogram visualization of co-expression of *HLA-DPB1* and *KRT19* gene transcripts in individual cells by scRNA-seq across the ductal and endocrine dendrogram from Figure 1D. (Left Bottom) Pie chart demonstrating *HLA-DPB1+KRT19+* cells as percentage of total cells. (Right) Magnified view of the clusters of cells with high percentage (25% or greater) of *HLA-DPB1+KRT19+* cells with *HLA-DPB1* and *KRT19* status displayed across these clusters (outlined in red dashed lines) and neighboring clusters of cells. Cells begin at the start pin symbol, and from there are partitioned based on similarities and differences in gene expression.

(B) (Top) Dendrogram visualization of cellular classification status across the magnified clusters of cells with high percentage (25% or greater) of *HLA-DPB1+KRT19+* cells (outlined in red dash lines) and neighboring clusters of cells. (Bottom) Pie chart displaying the relative proportion of cellular classification status of *HLA-DPB1+KRT19+* cells. P-value presented is the result of the Chi-squared test.

(C) (Top Left) Dendrogram visualization of donor group across the magnified clusters of cells with high percentage (25% or greater) of *HLA-DPB1+KRT19+* cells (outlined in red) and neighboring clusters of cells. (Top Right) Pie chart displaying the relative proportion of *HLA-DPB1+KRT19+* cells in Control (top) or T1D (bottom) Ductal-1 cells. P-value presented is the result of the Fisher exact test. (Bottom Left) Pie chart displaying the relative proportion of donor group of *HLA-DPB1+KRT19+* cells. P-value presented is the result of the Chi-squared test. (Bottom Right) Box plots displaying the *HLA-DPB1+KRT19+* cell percentage of total cells per individual across donor groups.

Figure 6

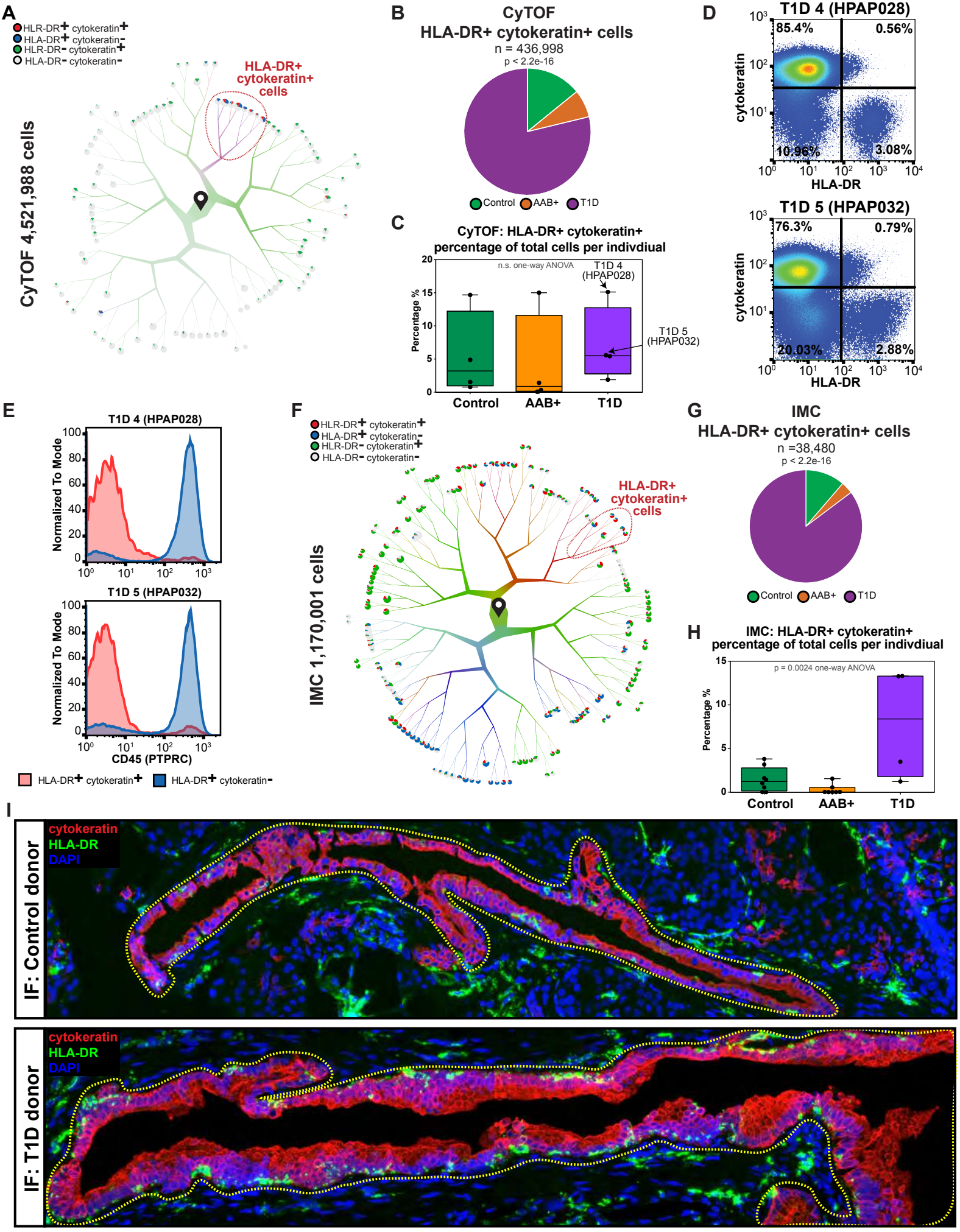

A
HLR-DR+ cytokeratin+
HLA-DR+ cytokeratin-
HLR-DR- cytokeratin+
HLA-DR- cytokeratin-
HLA-DR+ cytokeratin+ cells
CyTOF 4,521,988 cells
B
CyTOF
HLA-DR+ cytokeratin+ cells
n = 436,998
p < 2.2e-16
Control
AAB+
T1D
C
CyTOF: HLA-DR+ cytokeratin+ percentage of total cells per indivdiual
n.s. one-way ANOVA
T1D 4 (HPAP028)
T1D 5 (HPAP032)
Percentage %
Control
AAB+
T1D
D
T1D 4 (HPAP028)
85.4%
0.56%
10.96%
3.08%
T1D 5 (HPAP032)
76.3%
0.79%
20.03%
2.88%
cytokeratin
HLA-DR
E
T1D 4 (HPAP028)
T1D 5 (HPAP032)
Normalized To Mode
CD45 (PTPRC)
HLA-DR+ cytokeratin+
HLA-DR+ cytokeratin-
F
HLR-DR+ cytokeratin+
HLA-DR+ cytokeratin-
HLR-DR- cytokeratin+
HLA-DR- cytokeratin-
HLA-DR+ cytokeratin+ cells
IMC 1,170,001 cells
G
IMC
HLA-DR+ cytokeratin+ cells
n =38,480
p < 2.2e-16
Control
AAB+
T1D
H
IMC: HLA-DR+ cytokeratin+ percentage of total cells per indivdiual
p = 0.0024 one-way ANOVA
Percentage %
Control
AAB+
T1D
I
IF: Control donor
cytokeratin
HLA-DR
DAPI
IF: T1D donor
cytokeratin
HLA-DR
DAPI

Figure 6: Three single-cell resolution protein-based approaches corroborate the MHC Class II-expressing ductal cells in T1D

(A) Dendrogram visualization of co-expression of HLA-DR and cytokeratin proteins in single cells analyzed with flow cytometry by time-of-flight (CyTOF). Cells begin at the start pin symbol, and from there are partitioned based on similarities and differences in protein levels.

(B) Pie chart displaying HLA-DR+ cytokeratin+ cells and the relative proportions of each donor group from the CyTOF data. P-value presented is the result of the Chi-squared test.

(C) Box plots displaying HLA-DR+ cytokeratin+ cell percentage of total cells per individual across donor groups from the CyTOF data.

(D) Two parameter CyTOF analysis of HLA-DR and cytokeratin protein expression in single cells from T1D donor #4 (HPAP028) and T1D donor #5 (HPAP032).

(E) CD45 (PTPRC) expression levels in HLA-DR+ cytokeratin+ and HLA-DR+ cytokeratin- single cells.

(F) Dendrogram visualization of co-expression of HLA-DR and cytokeratin proteins in single cells analyzed by imaging mass cytometry (IMC). Cells begin at the start pin symbol, and from there are partitioned based on similarities and differences in protein levels.

(G) Pie chart displaying HLA-DR+ cytokeratin+ cells and the relative proportions of each donor group from the IMC data. P-value presented is the result of the Chi-squared test.

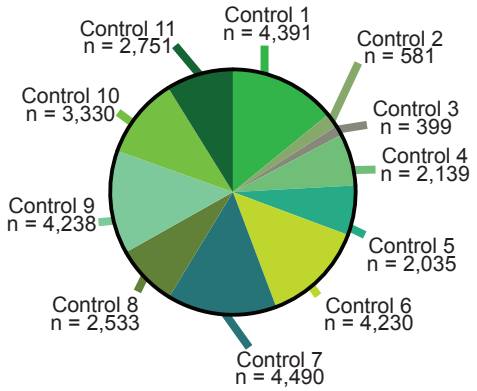
(H) Box plots displaying HLA-DR+ cytokeratin+ cell percentage of total cells per individual across donor groups from the IMC data.

(I) Representative region of interest (ROI) from the pancreas of a T1D donor (top) or Control donor (bottom) displaying HLA-DR+ cytokeratin+ cells via immunohistochemistry (IHC) followed by confocal microscopy.

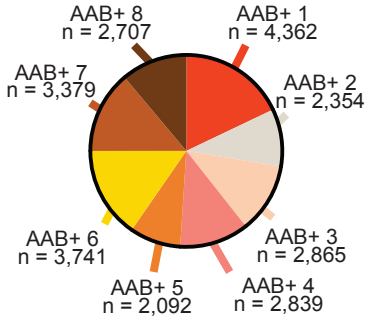
Figure S1

A

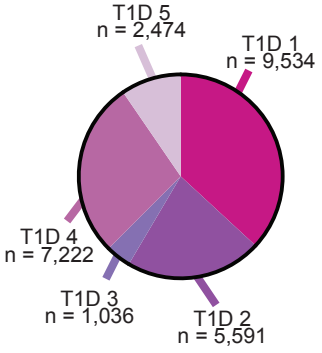
Before Seurat Clustering and Doublet Removal (n = 81,313 cells)



Control Donors
cell n = 31,117
donor n = 11

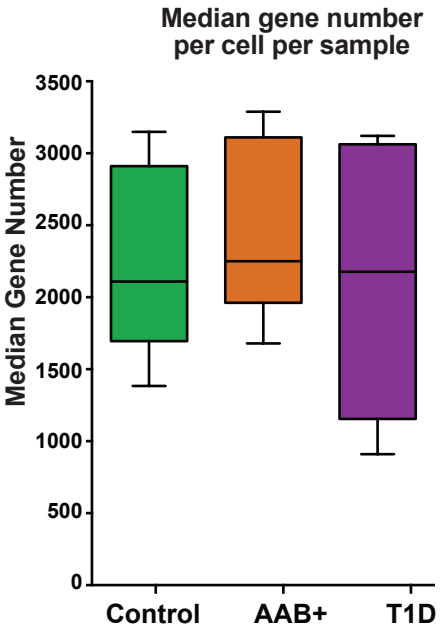


AAB+ Donors
cell n = 24,339
donor n = 8

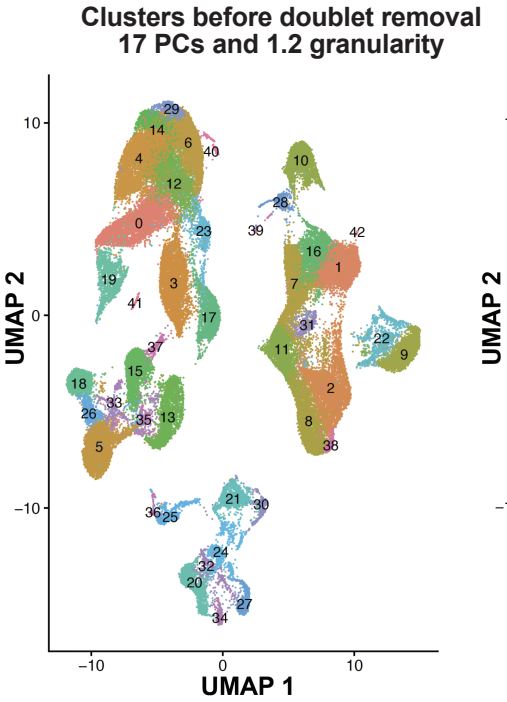


T1D Donors
cell n = 25,857
donor n = 5

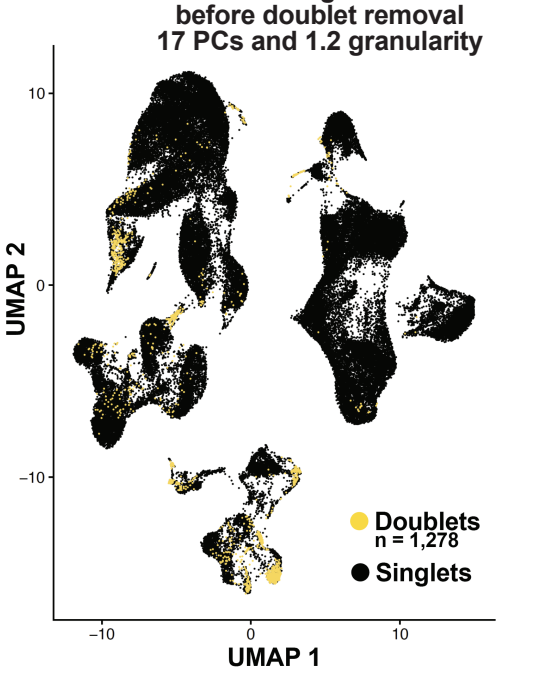
B



C



D



E

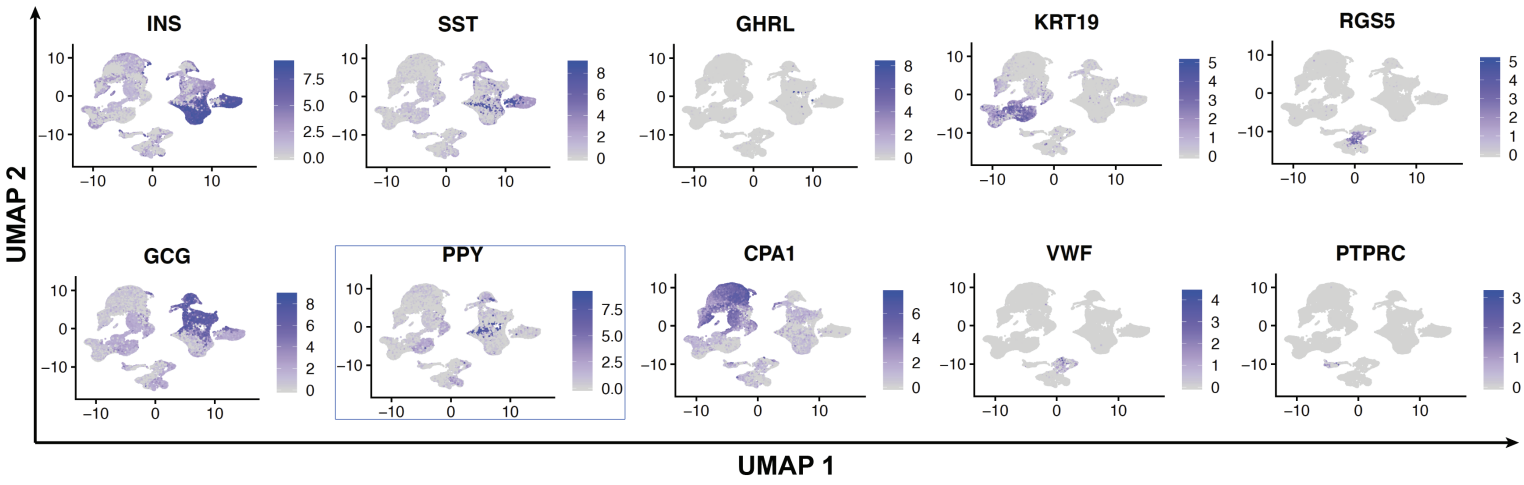
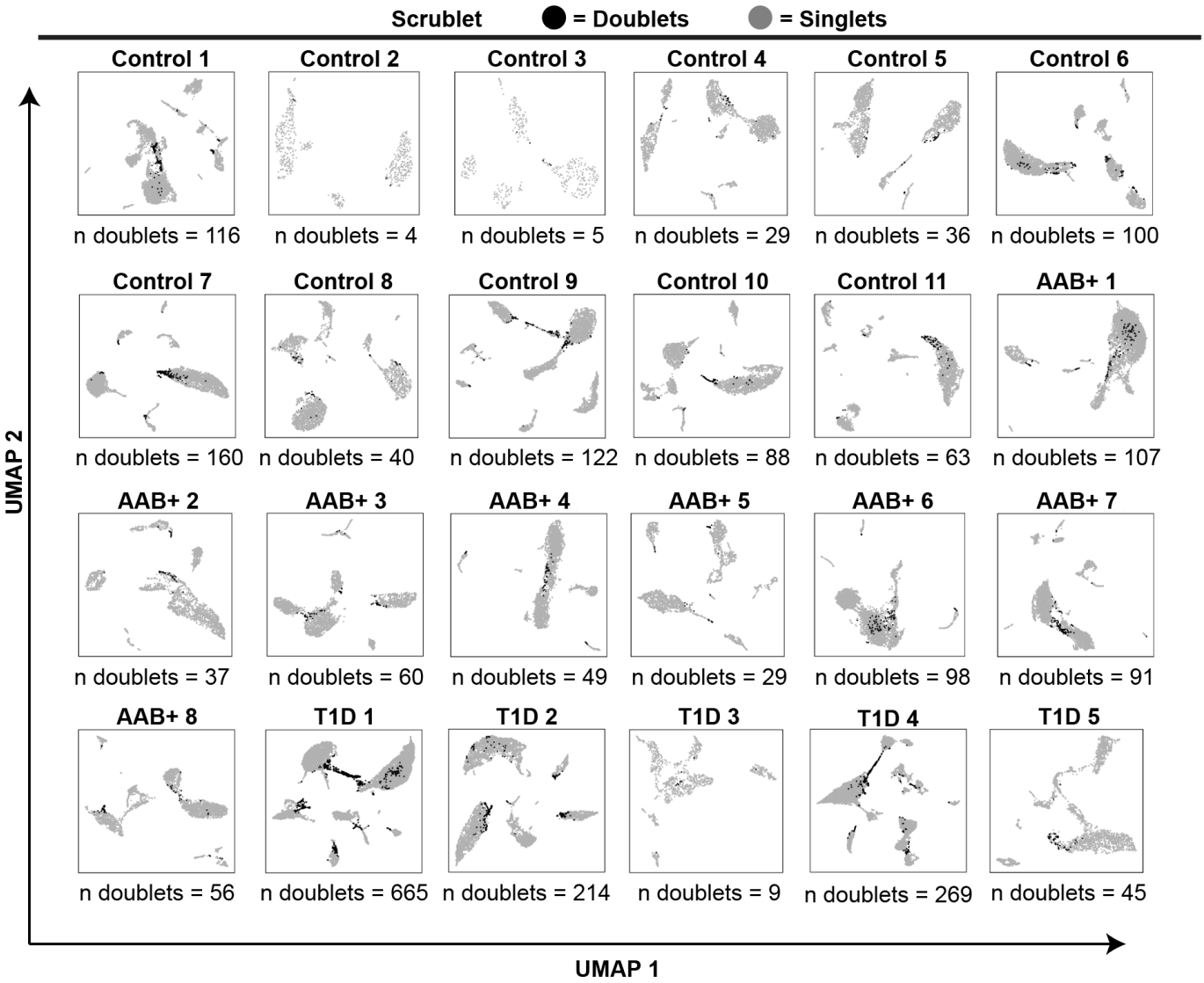
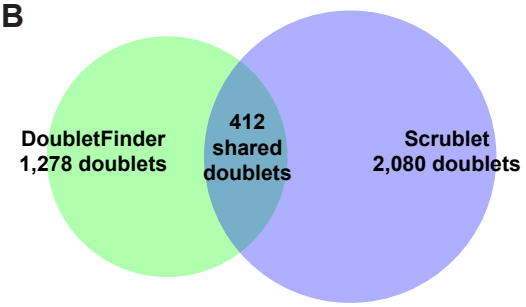


Figure S2

A



B



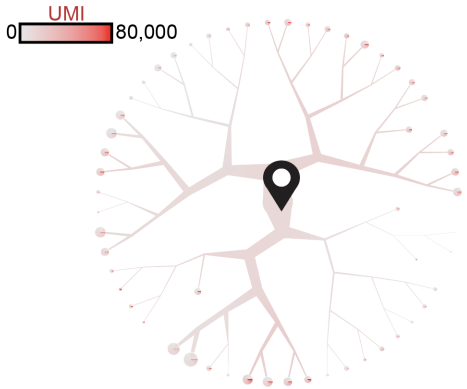
n = 3,770 total doublets removed

C

	Cell Number Removed	Total Cell Number
Starting Number of Cells		81,313
Filtering Before Doublet Removal	4,308	77,005
Doublet Removal	3,770	73,235
Filtering After Doublet Removal	3,590	69,645

D

UMI across all cells tree (Fig. 1C)



E

UMI across endocrine and ductal cell tree (Fig. 1D)

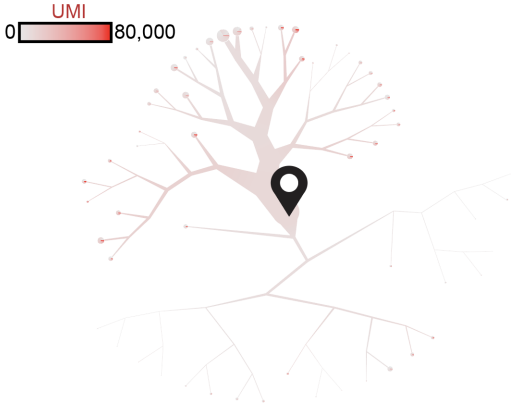
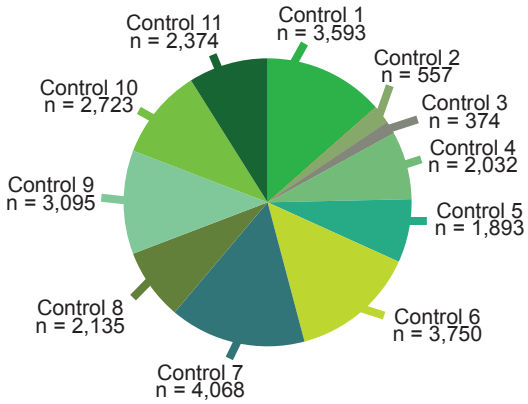


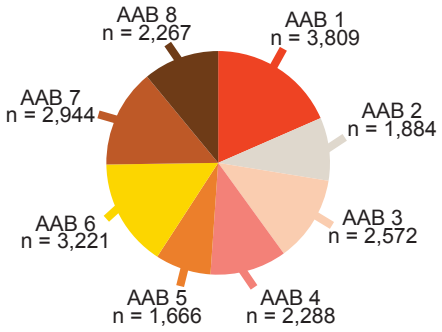
Figure S3

A

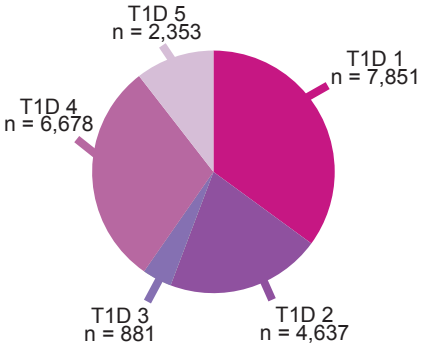
After Filtering and Doublet Removal (n = 69,645)



n = 26,594 cells



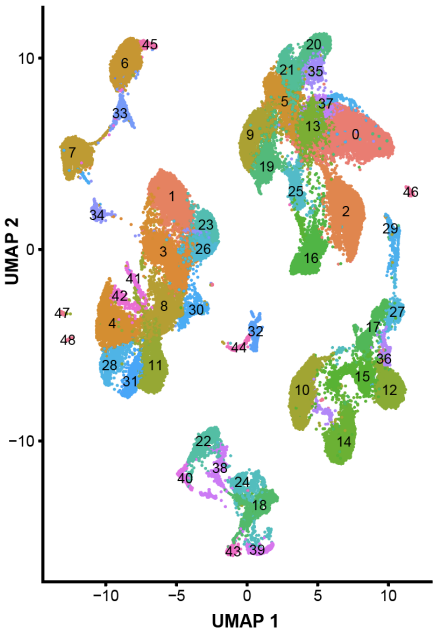
n = 20,651 cells



n = 22,400 cells

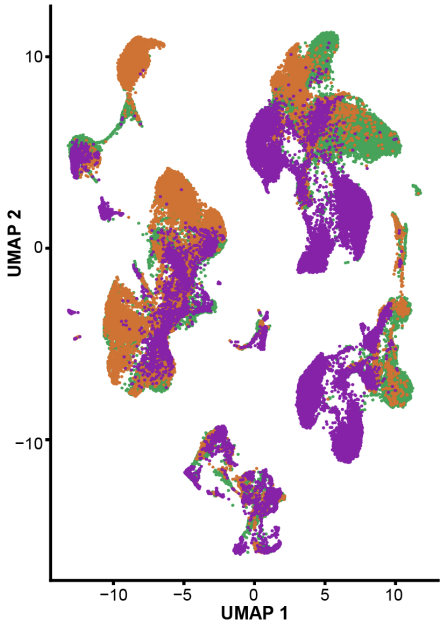
B

49 clusters after doublet removal



C

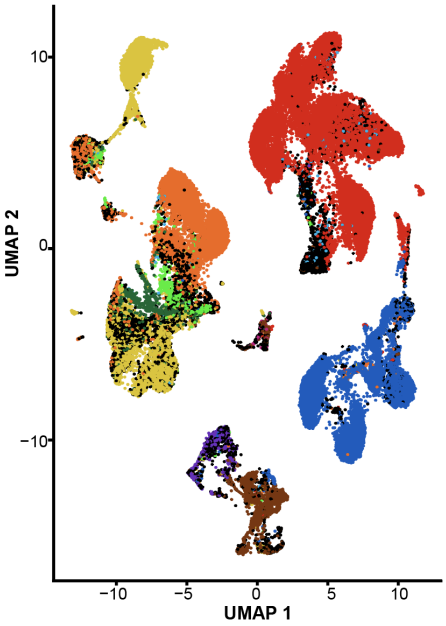
49 clusters after doublet removal with donor group labels



- Control
- AAB+
- T1D

D

49 clusters after doublet removal with cellular classification labels



- Acinar Cells
- Alpha Cells
- Beta Cells
- Delta Cells
- Ductal Cells
- Endothelial Cells
- Epsilon Cells
- Immune Cells
- PP Cells
- Stellate Cells
- Unknown

E

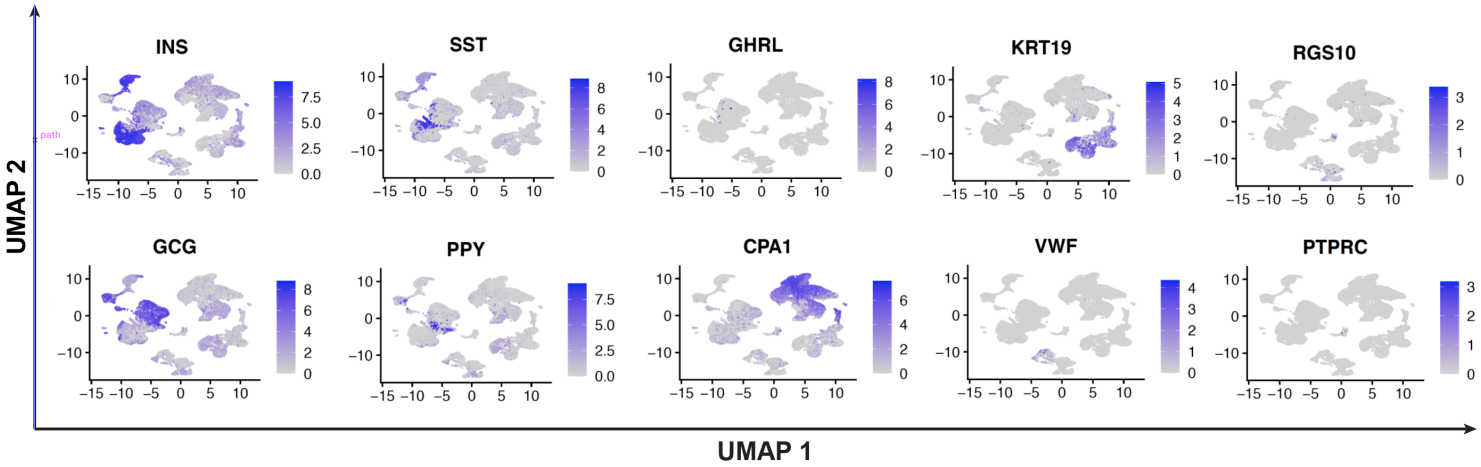


Figure S4

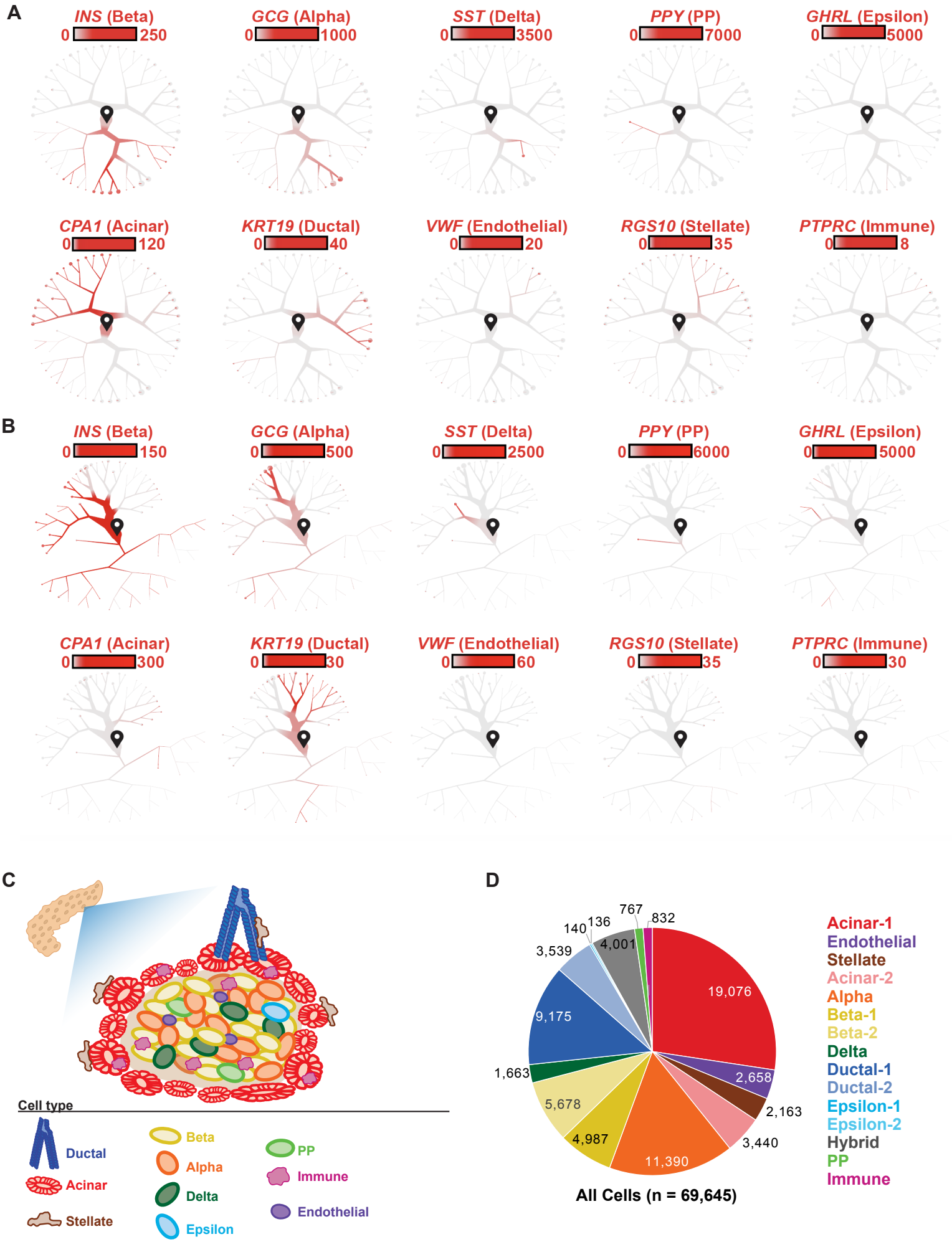

A
INS (Beta) 0 250
GCG (Alpha) 0 1000
SST (Delta) 0 3500
PPY (PP) 0 7000
GHRL (Epsilon) 0 5000
CPA1 (Acinar) 0 120
KRT19 (Ductal) 0 40
VWF (Endothelial) 0 20
RGS10 (Stellate) 0 35
PTPRC (Immune) 0 8
B
INS (Beta) 0 150
GCG (Alpha) 0 500
SST (Delta) 0 2500
PPY (PP) 0 6000
GHRL (Epsilon) 0 5000
CPA1 (Acinar) 0 300
KRT19 (Ductal) 0 30
VWF (Endothelial) 0 60
RGS10 (Stellate) 0 35
PTPRC (Immune) 0 30
C
Cell type
Ductal
Acinar
Stellate
Beta
Alpha
Delta
Epsilon
PP
Immune
Endothelial
D
19,076
2,658
2,163
3,440
11,390
4,987
5,678
1,663
9,175
3,539
140
136
4,001
767
832
All Cells (n = 69,645)
Acinar-1
Endothelial
Stellate
Acinar-2
Alpha
Beta-1
Beta-2
Delta
Ductal-1
Ductal-2
Epsilon-1
Epsilon-2
Hybrid
PP
Immune

Figure S5

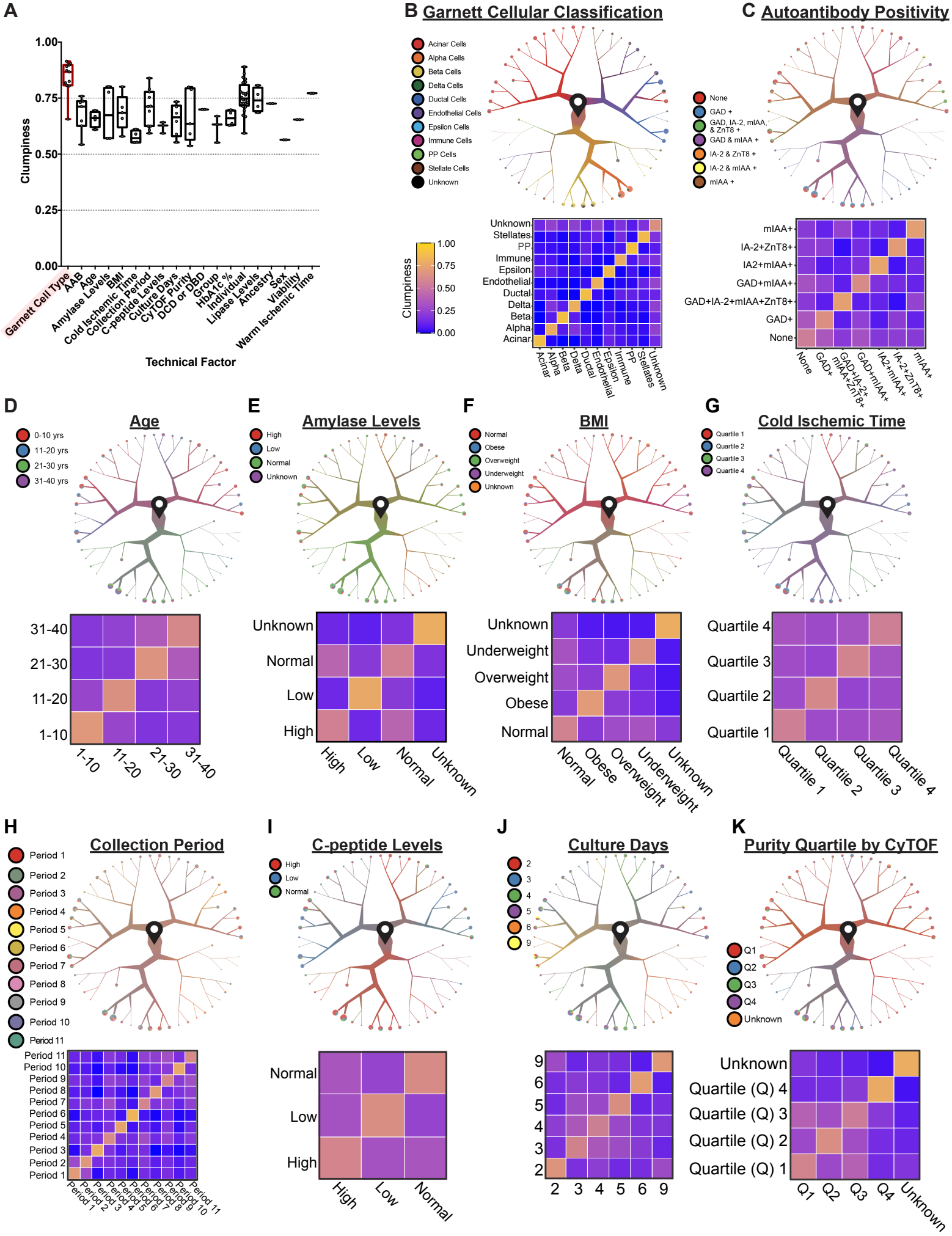

A
Clumpiness
1.00
0.75
0.50
0.25
0.00
Garnett Cell Type
AAB
Age
Amylase Levels
BMI
Cold Ischemic Time
Collection Period
C-peptide Levels
Culture Days
CyTOF Purity
DCD or DBD
Group
HbA1c %
Individual
Lipase Levels
Ancestry
Sex
Viability
Warm Ischemic Time
Technical Factor
B Garnett Cellular Classification
Acinar Cells
Alpha Cells
Beta Cells
Delta Cells
Ductal Cells
Endothelial Cells
Epsilon Cells
Immune Cells
PP Cells
Stellate Cells
Unknown
Clumpiness
1.00
0.75
0.50
0.25
0.00
Unknown
Stellates
PP
Immune
Epsilon
Endothelial
Ductal
Delta
Beta
Alpha
Acinar
C Autoantibody Positivity
None
GAD +
GAD, IA-2, mIAA, & ZnT8 +
GAD & mIAA +
IA-2 & ZnT8 +
IA-2 & mIAA +
mIAA +
mIAA+
IA-2+ZnT8+
IA2+mIAA+
GAD+mIAA+
GAD+IA-2+mIAA+ZnT8+
GAD+
None
D Age
0-10 yrs
11-20 yrs
21-30 yrs
31-40 yrs
31-40
21-30
11-20
1-10
E Amylase Levels
High
Low
Normal
Unknown
F BMI
Normal
Obese
Overweight
Underweight
Unknown
G Cold Ischemic Time
Quartile 1
Quartile 2
Quartile 3
Quartile 4
H Collection Period
Period 1
Period 2
Period 3
Period 4
Period 5
Period 6
Period 7
Period 8
Period 9
Period 10
Period 11
I C-peptide Levels
High
Low
Normal
J Culture Days
2
3
4
5
6
9
K Purity Quartile by CyTOF
Q1
Q2
Q3
Q4
Unknown
Quartile (Q) 4
Quartile (Q) 3
Quartile (Q) 2
Quartile (Q) 1

Figure S6

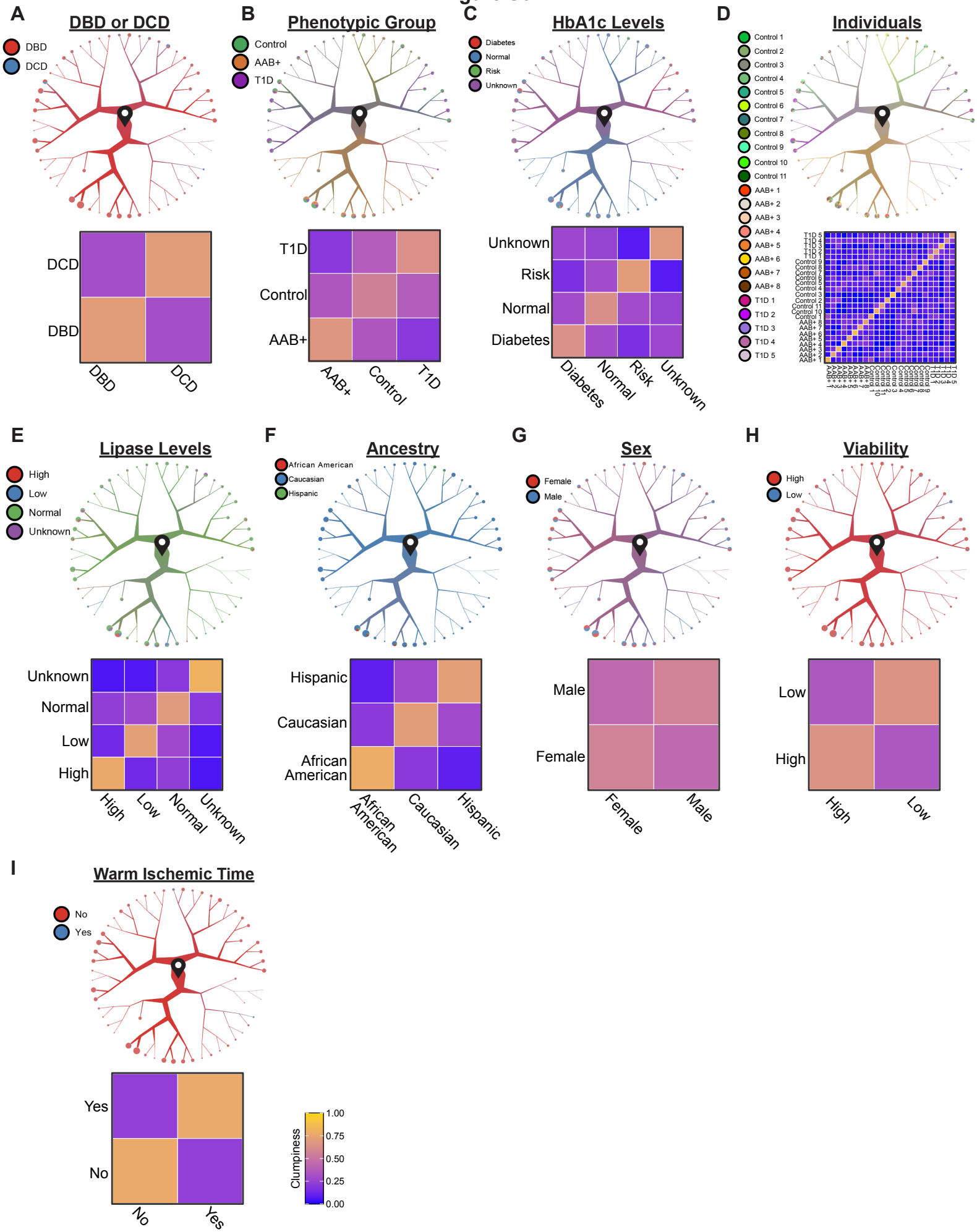

A
DBD or DCD
DBD
DCD
DCD
DBD
DBD
DCD
B
Phenotypic Group
Control
AAB+
T1D
T1D
Control
AAB+
AAB+
Control
T1D
C
HbA1c Levels
Diabetes
Normal
Risk
Unknown
Unknown
Risk
Normal
Diabetes
Diabetes
Normal
Risk
Unknown
D
Individuals
Control 1
Control 2
Control 3
Control 4
Control 5
Control 6
Control 7
Control 8
Control 9
Control 10
Control 11
AAB+ 1
AAB+ 2
AAB+ 3
AAB+ 4
AAB+ 5
AAB+ 6
AAB+ 7
AAB+ 8
T1D 1
T1D 2
T1D 3
T1D 4
T1D 5
E
Lipase Levels
High
Low
Normal
Unknown
Unknown
Normal
Low
High
High
Low
Normal
Unknown
F
Ancestry
African American
Caucasian
Hispanic
Hispanic
Caucasian
African American
African American
Caucasian
Hispanic
G
Sex
Female
Male
Male
Female
Female
Male
H
Viability
High
Low
Low
High
High
Low
I
Warm Ischemic Time
No
Yes
Yes
No
No
Yes
Clumpiness
1.00
0.75
0.50
0.25
0.00

Figure S7

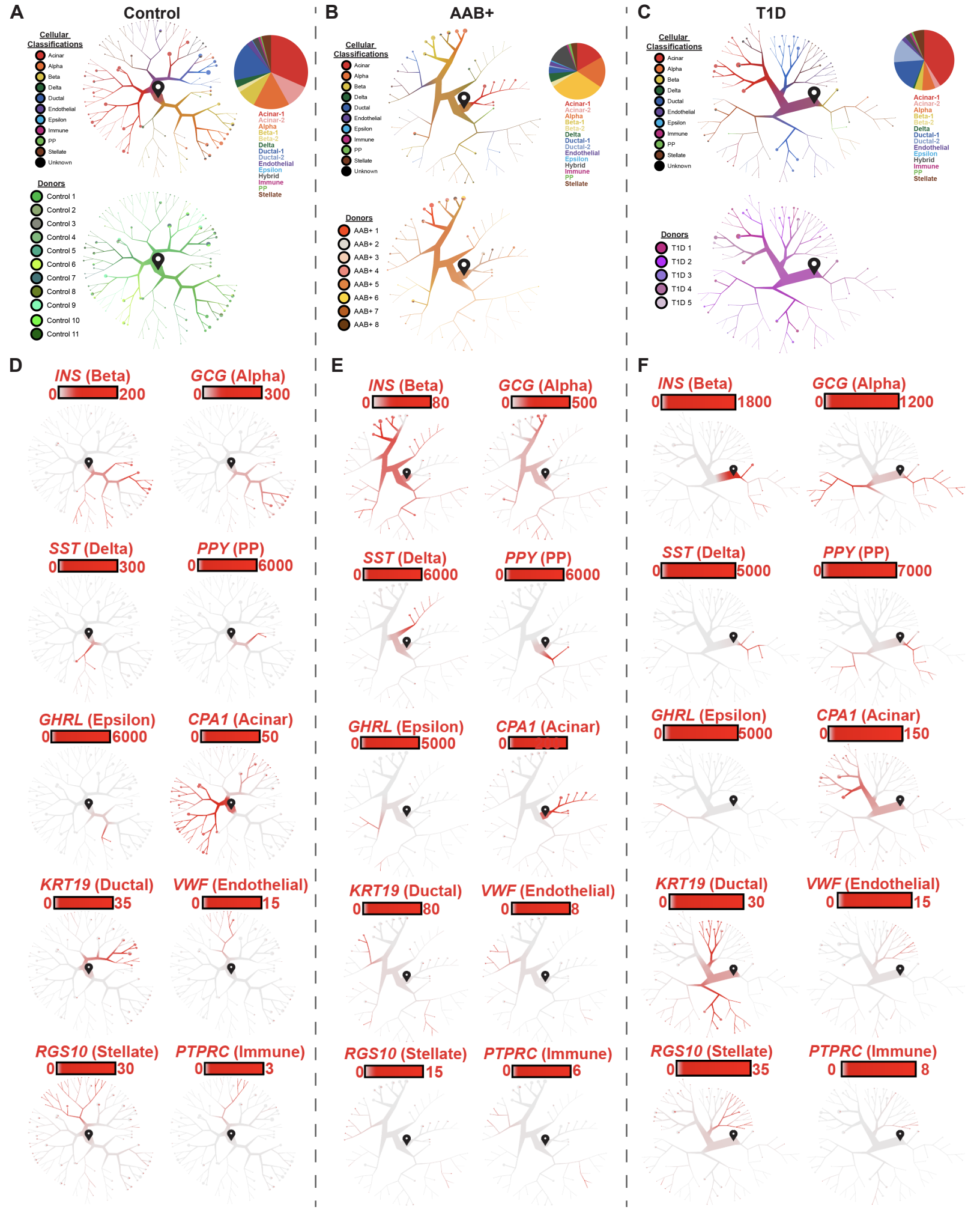
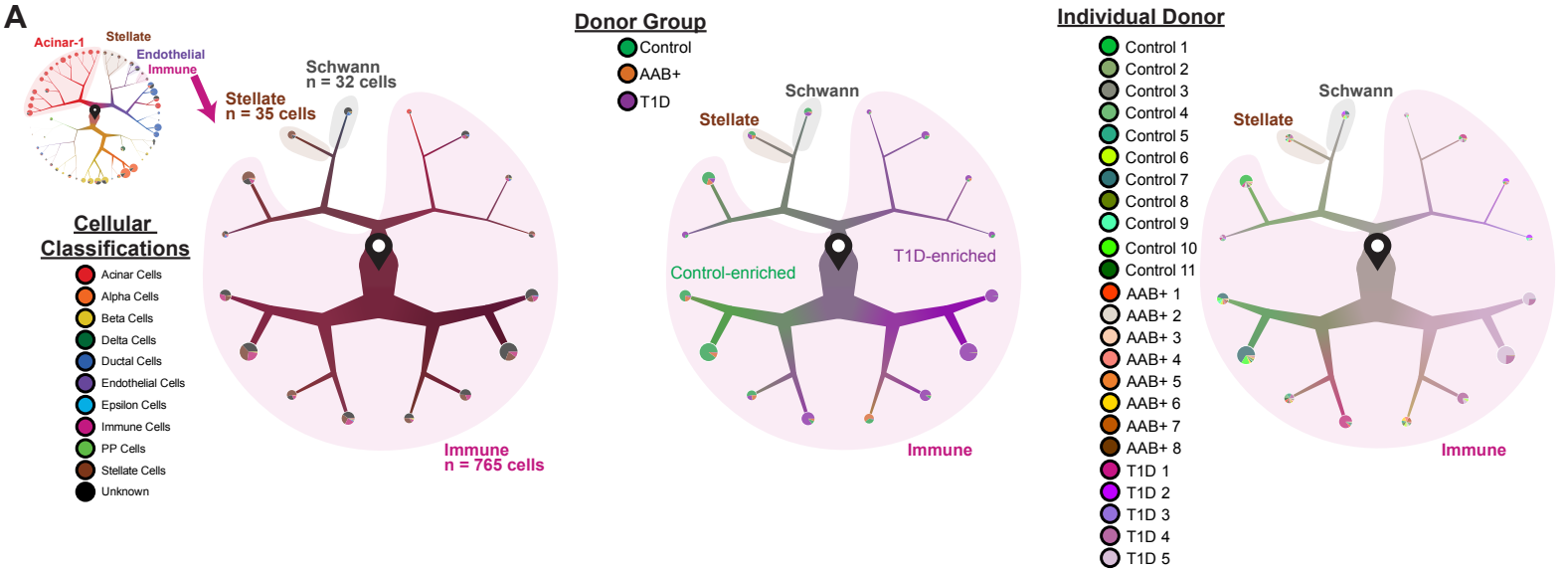
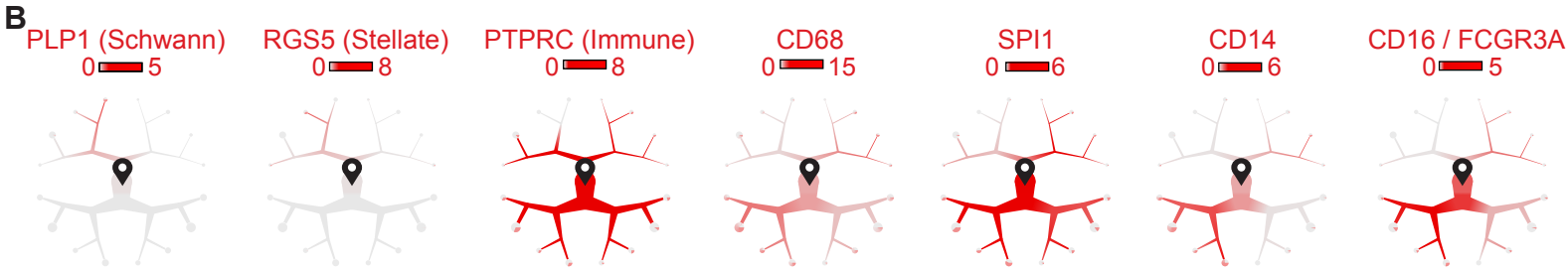
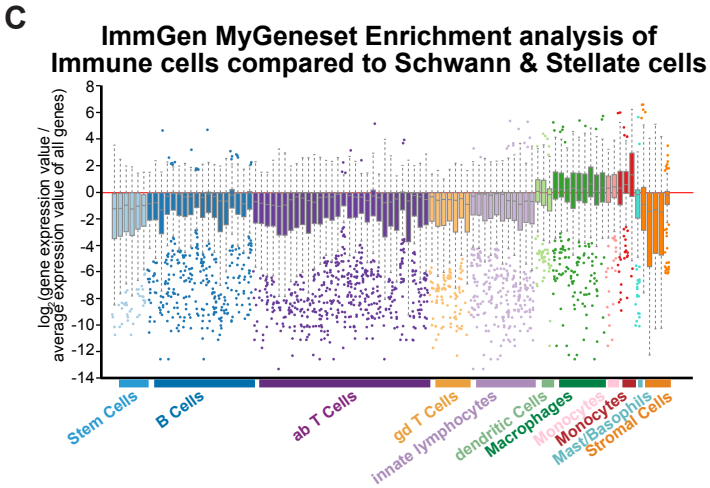
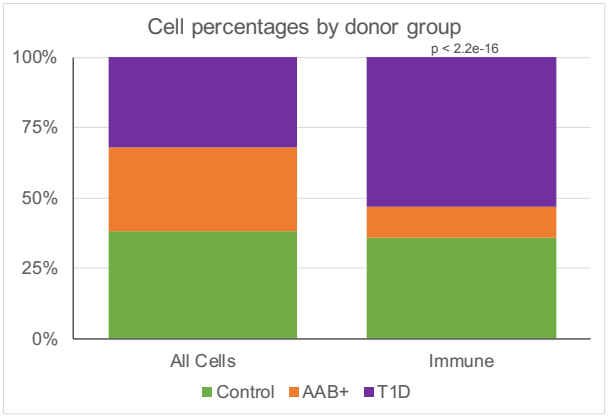

A
Control
Cellular Classifications
Acinar
Alpha
Beta
Delta
Ductal
Endothelial
Epsilon
Immune
PP
Stellate
Unknown
Donors
Control 1
Control 2
Control 3
Control 4
Control 5
Control 6
Control 7
Control 8
Control 9
Control 10
Control 11
Acinar-1
Acinar-2
Alpha
Beta-1
Beta-2
Delta
Ductal-1
Ductal-2
Endothelial
Epsilon
Hybrid
Immune
PP
Stellate
B
AAB+
Donors
AAB+ 1
AAB+ 2
AAB+ 3
AAB+ 4
AAB+ 5
AAB+ 6
AAB+ 7
AAB+ 8
C
T1D
Donors
T1D 1
T1D 2
T1D 3
T1D 4
T1D 5
D
INS (Beta) 0 200
GCG (Alpha) 0 300
SST (Delta) 0 300
PPY (PP) 0 6000
GHRL (Epsilon) 0 6000
CPA1 (Acinar) 0 50
KRT19 (Ductal) 0 35
VWF (Endothelial) 0 15
RGS10 (Stellate) 0 30
PTPRC (Immune) 0 3
E
INS (Beta) 0 80
GCG (Alpha) 0 500
SST (Delta) 0 6000
PPY (PP) 0 6000
GHRL (Epsilon) 0 5000
CPA1 (Acinar) 0
KRT19 (Ductal) 0 80
VWF (Endothelial) 0 8
RGS10 (Stellate) 0 15
PTPRC (Immune) 0 6
F
INS (Beta) 0 1800
GCG (Alpha) 0 1200
SST (Delta) 0 5000
PPY (PP) 0 7000
GHRL (Epsilon) 0 5000
CPA1 (Acinar) 0 150
KRT19 (Ductal) 0 30
VWF (Endothelial) 0 15
RGS10 (Stellate) 0 35
PTPRC (Immune) 0 8

Figure S8


A
Acinar-1
Stellate
Endothelial
Immune
Schwann
n = 32 cells
Stellate
n = 35 cells
Immune
n = 765 cells
Cellular Classifications
Acinar Cells
Alpha Cells
Beta Cells
Delta Cells
Ductal Cells
Endothelial Cells
Epsilon Cells
Immune Cells
PP Cells
Stellate Cells
Unknown
Donor Group
Control
AAB+
T1D
Schwann
Stellate
Control-enriched
T1D-enriched
Immune
Individual Donor
Control 1
Control 2
Control 3
Control 4
Control 5
Control 6
Control 7
Control 8
Control 9
Control 10
Control 11
AAB+ 1
AAB+ 2
AAB+ 3
AAB+ 4
AAB+ 5
AAB+ 6
AAB+ 7
AAB+ 8
T1D 1
T1D 2
T1D 3
T1D 4
T1D 5
Schwann
Stellate
Immune


B
PLP1 (Schwann)
0 5
RGS5 (Stellate)
0 8
PTPRC (Immune)
0 8
CD68
0 15
SPI1
0 6
CD14
0 6
CD16 / FCGR3A
0 5


C
ImmGen MyGeneset Enrichment analysis of
Immune cells compared to Schwann & Stellate cells
log2 (gene expression value / average expression value of all genes)
8
6
4
2
0
-2
-4
-6
-8
-10
-12
-14
Stem Cells
B Cells
ab T Cells
gd T Cells
innate lymphocytes
dendritic cells
Macrophages
Monocytes
Monocytes
Mast/Basophils
Stromal Cells


D
Cell percentages by donor group
p < 2.2e-16
100%
75%
50%
25%
0%
All Cells
Immune
Control
AAB+
T1D

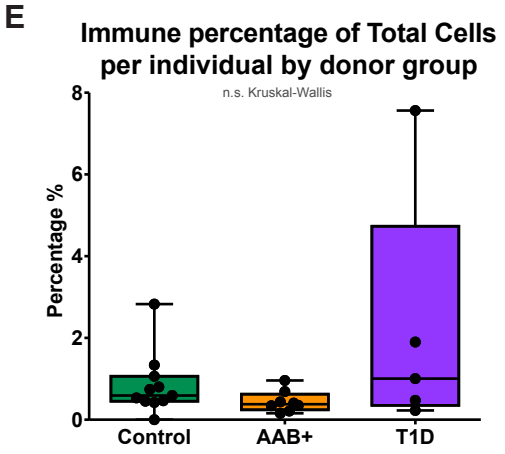
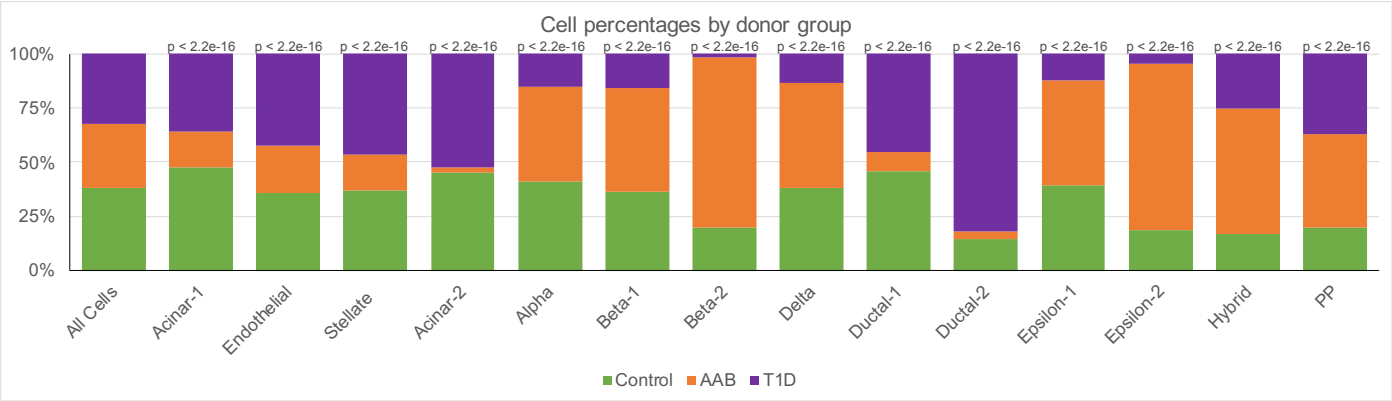

E
Immune percentage of Total Cells
per individual by donor group
n.s. Kruskal-Wallis
Percentage %
8
6
4
2
0
Control
AAB+
T1D

Figure S9

A


Cell percentages by donor group
p < 2.2e-16
100%
75%
50%
25%
0%
All Cells
Acinar-1
Endothelial
Stellate
Acinar-2
Alpha
Beta-1
Beta-2
Delta
Ductal-1
Ductal-2
Epsilon-1
Epsilon-2
Hybrid
PP
Control
AAB
T1D

B

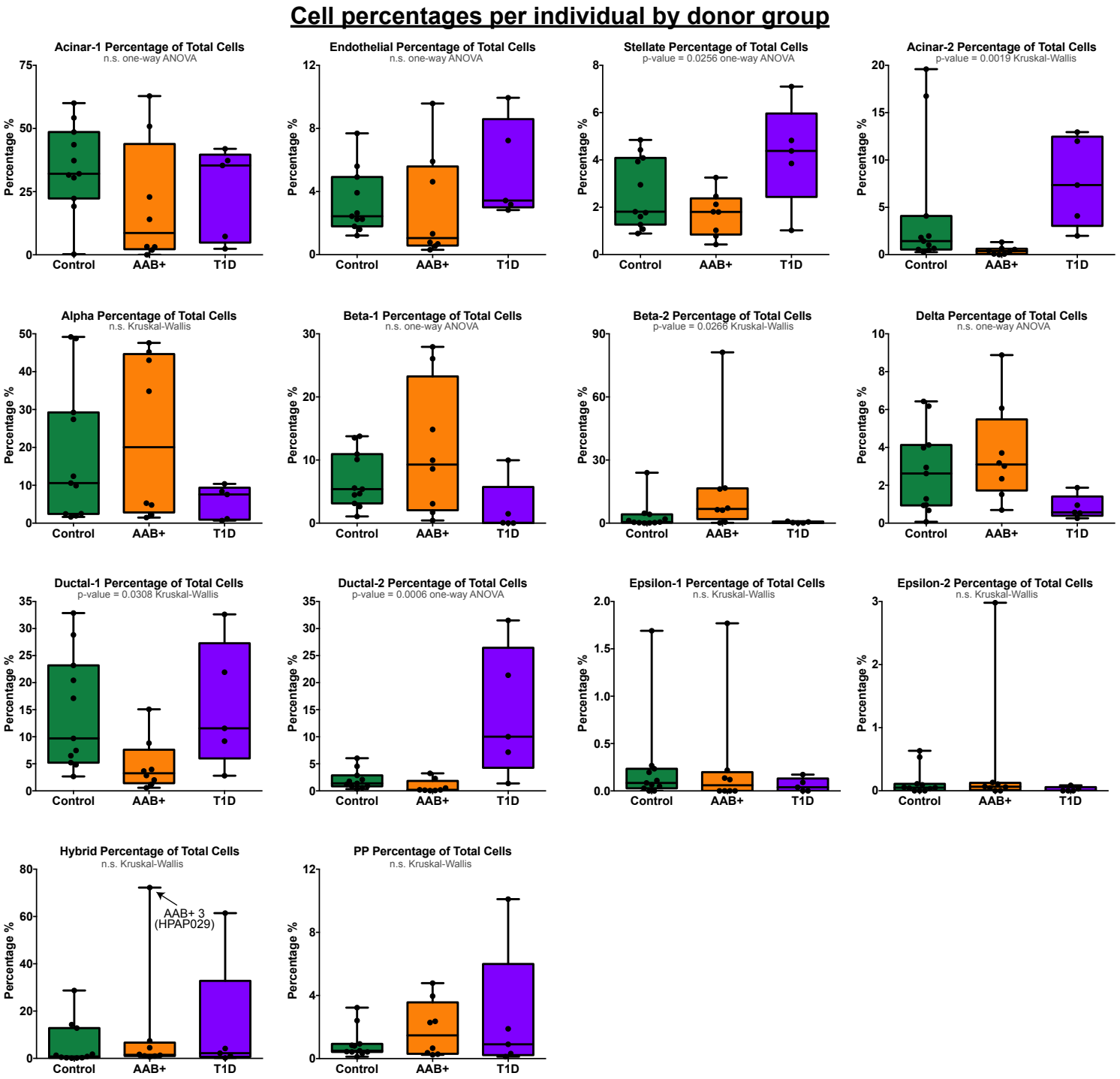

Cell percentages per individual by donor group
Acinar-1 Percentage of Total Cells
n.s. one-way ANOVA
Endothelial Percentage of Total Cells
n.s. one-way ANOVA
Stellate Percentage of Total Cells
p-value = 0.0256 one-way ANOVA
Acinar-2 Percentage of Total Cells
p-value = 0.0019 Kruskal-Wallis
Alpha Percentage of Total Cells
n.s. Kruskal-Wallis
Beta-1 Percentage of Total Cells
n.s. one-way ANOVA
Beta-2 Percentage of Total Cells
p-value = 0.0266 Kruskal-Wallis
Delta Percentage of Total Cells
n.s. one-way ANOVA
Ductal-1 Percentage of Total Cells
p-value = 0.0308 Kruskal-Wallis
Ductal-2 Percentage of Total Cells
p-value = 0.0006 one-way ANOVA
Epsilon-1 Percentage of Total Cells
n.s. Kruskal-Wallis
Epsilon-2 Percentage of Total Cells
n.s. Kruskal-Wallis
Hybrid Percentage of Total Cells
n.s. Kruskal-Wallis
AAB+ 3
(HPAP029)
PP Percentage of Total Cells
n.s. Kruskal-Wallis
Percentage %
Control
AAB+
T1D

Figure S10

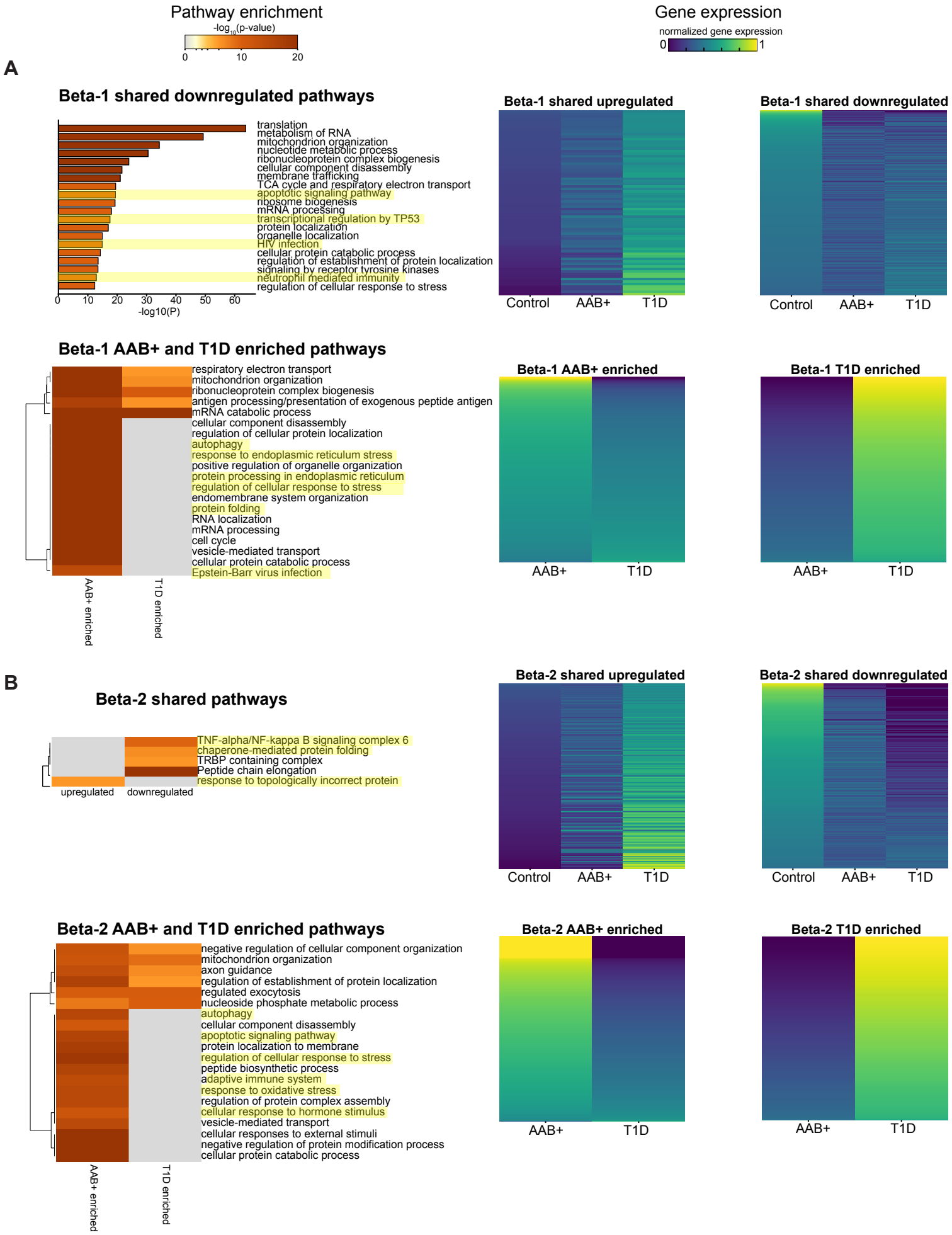

Pathway enrichment
-log10(p-value)
0
10
20
Gene expression
normalized gene expression
0
1
A
Beta-1 shared downregulated pathways
translation
metabolism of RNA
mitochondrion organization
nucleotide metabolic process
ribonucleoprotein complex biogenesis
cellular component disassembly
membrane trafficking
TCA cycle and respiratory electron transport
apoptotic signaling pathway
ribosome biogenesis
mRNA processing
transcriptional regulation by TP53
protein localization
organelle localization
HIV infection
cellular protein catabolic process
regulation of establishment of protein localization
signaling by receptor tyrosine kinases
neutrophil mediated immunity
regulation of cellular response to stress
0
10
20
30
40
50
60
-log10(P)
Beta-1 shared upregulated
Control
AAB+
T1D
Beta-1 shared downregulated
Control
AAB+
T1D
Beta-1 AAB+ and T1D enriched pathways
respiratory electron transport
mitochondrion organization
ribonucleoprotein complex biogenesis
antigen processing/presentation of exogenous peptide antigen
mRNA catabolic process
cellular component disassembly
regulation of cellular protein localization
autophagy
response to endoplasmic reticulum stress
positive regulation of organelle organization
protein processing in endoplasmic reticulum
regulation of cellular response to stress
endomembrane system organization
protein folding
RNA localization
mRNA processing
cell cycle
vesicle-mediated transport
cellular protein catabolic process
Epstein-Barr virus infection
AAB+ enriched
T1D enriched
Beta-1 AAB+ enriched
AAB+
T1D
Beta-1 T1D enriched
AAB+
T1D
B
Beta-2 shared pathways
TNF-alpha/NF-kappa B signaling complex 6
chaperone-mediated protein folding
TRBP containing complex
Peptide chain elongation
response to topologically incorrect protein
upregulated
downregulated
Beta-2 shared upregulated
Control
AAB+
T1D
Beta-2 shared downregulated
Control
AAB+
T1D
Beta-2 AAB+ and T1D enriched pathways
negative regulation of cellular component organization
mitochondrion organization
axon guidance
regulation of establishment of protein localization
regulated exocytosis
nucleoside phosphate metabolic process
autophagy
cellular component disassembly
apoptotic signaling pathway
protein localization to membrane
regulation of cellular response to stress
peptide biosynthetic process
adaptive immune system
response to oxidative stress
regulation of protein complex assembly
cellular response to hormone stimulus
vesicle-mediated transport
cellular responses to external stimuli
negative regulation of protein modification process
cellular protein catabolic process
AAB+ enriched
T1D enriched
Beta-2 AAB+ enriched
AAB+
T1D
Beta-2 T1D enriched
AAB+
T1D

Figure S11

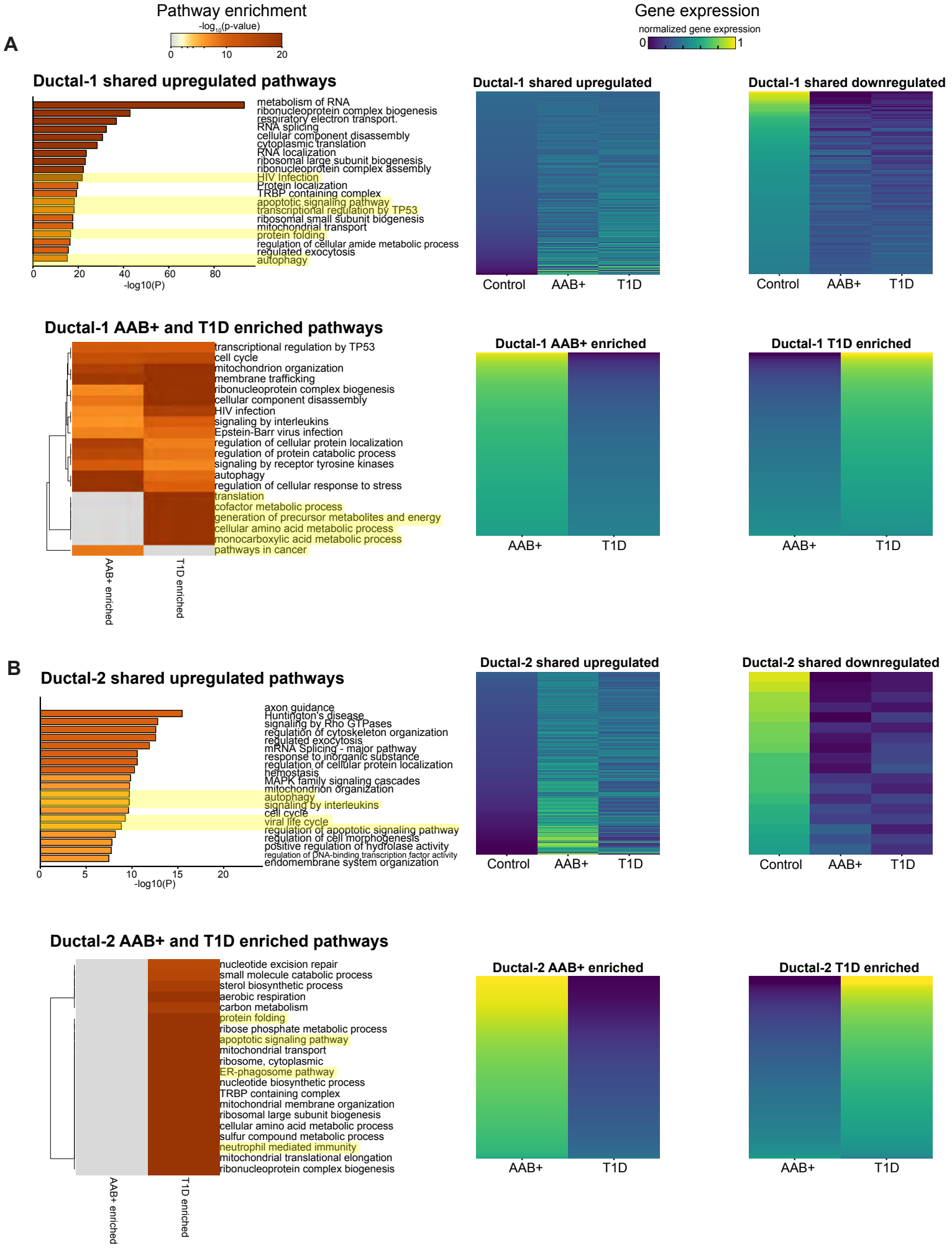

Pathway enrichment
-log10(p-value)
0
10
20
Gene expression
normalized gene expression
0
1
A
Ductal-1 shared upregulated pathways
metabolism of RNA
ribonucleoprotein complex biogenesis
respiratory electron transport.
RNA splicing
cellular component disassembly
cytoplasmic translation
RNA localization
ribosomal large subunit biogenesis
ribonucleoprotein complex assembly
HIV infection
Protein localization
TRBP containing complex
apoptotic signaling pathway
transcriptional regulation by TP53
ribosomal small subunit biogenesis
mitochondrial transport
protein folding
regulation of cellular amide metabolic process
regulated exocytosis
autophagy
0
20
40
60
80
-log10(P)
Ductal-1 shared upregulated
Control
AAB+
T1D
Ductal-1 shared downregulated
Control
AAB+
T1D
Ductal-1 AAB+ and T1D enriched pathways
transcriptional regulation by TP53
cell cycle
mitochondrion organization
membrane trafficking
ribonucleoprotein complex biogenesis
cellular component disassembly
HIV infection
signaling by interleukins
Epstein-Barr virus infection
regulation of cellular protein localization
regulation of protein catabolic process
signaling by receptor tyrosine kinases
autophagy
regulation of cellular response to stress
translation
cofactor metabolic process
generation of precursor metabolites and energy
cellular amino acid metabolic process
monocarboxylic acid metabolic process
pathways in cancer
AAB+ enriched
T1D enriched
Ductal-1 AAB+ enriched
AAB+
T1D
Ductal-1 T1D enriched
AAB+
T1D
B
Ductal-2 shared upregulated pathways
axon guidance
Huntington's disease
signaling by Rho GTPases
regulation of cytoskeleton organization
regulated exocytosis
mRNA Splicing - major pathway
response to inorganic substance
regulation of cellular protein localization
hemostasis
MAPK family signaling cascades
mitochondrion organization
autophagy
signaling by interleukins
cell cycle
viral life cycle
regulation of apoptotic signaling pathway
regulation of cell morphogenesis
positive regulation of hydrolase activity
regulation of DNA-binding transcription factor activity
endomembrane system organization
0
5
10
15
20
-log10(P)
Ductal-2 shared upregulated
Control
AAB+
T1D
Ductal-2 shared downregulated
Control
AAB+
T1D
Ductal-2 AAB+ and T1D enriched pathways
nucleotide excision repair
small molecule catabolic process
sterol biosynthetic process
aerobic respiration
carbon metabolism
protein folding
ribose phosphate metabolic process
apoptotic signaling pathway
mitochondrial transport
ribosome, cytoplasmic
ER-phagosome pathway
nucleotide biosynthetic process
TRBP containing complex
mitochondrial membrane organization
ribosomal large subunit biogenesis
cellular amino acid metabolic process
sulfur compound metabolic process
neutrophil mediated immunity
mitochondrial translational elongation
ribonucleoprotein complex biogenesis
AAB+ enriched
T1D enriched
Ductal-2 AAB+ enriched
AAB+
T1D
Ductal-2 T1D enriched
AAB+
T1D

Figure S12

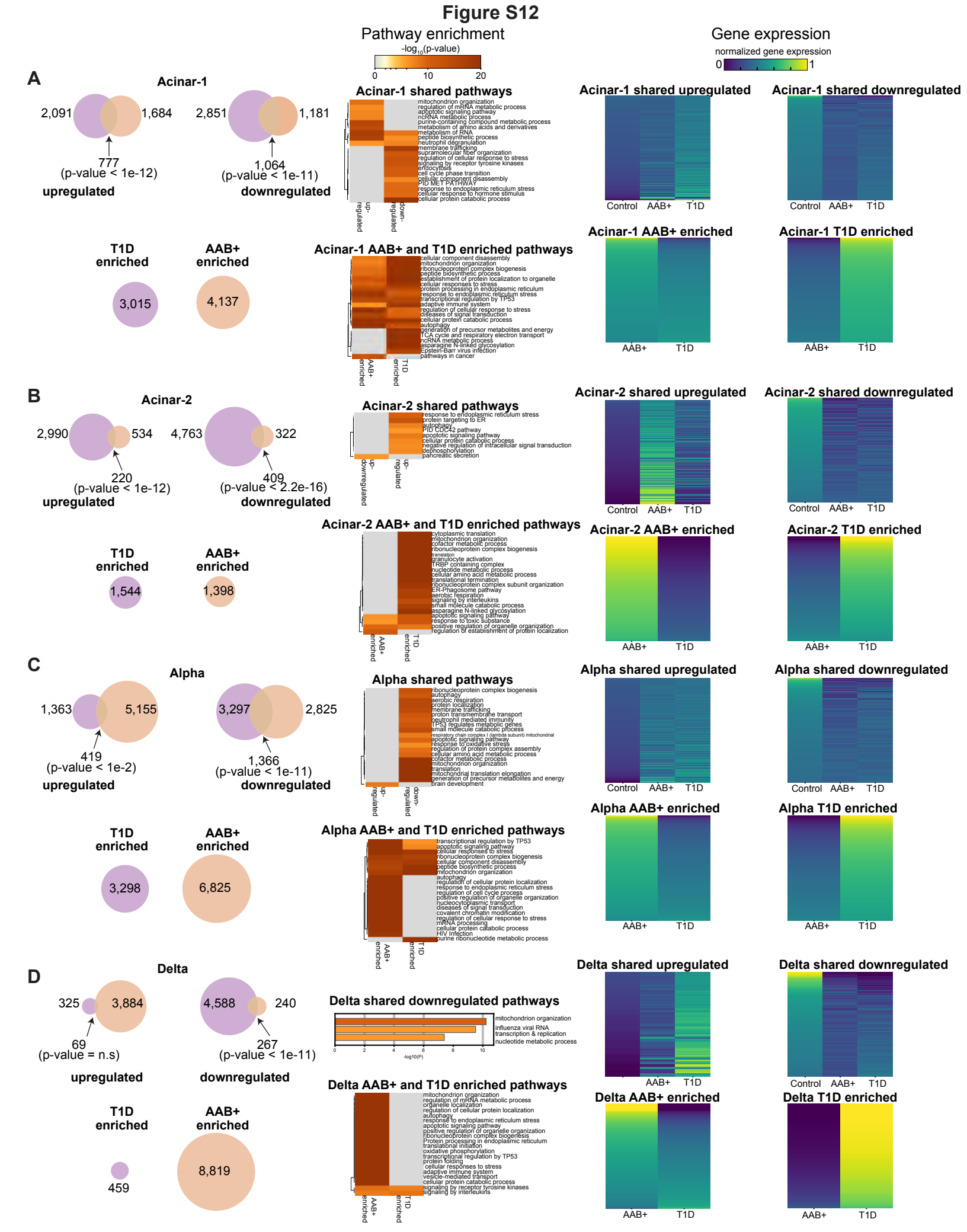

Pathway enrichment
-log10(p-value)
0
10
20
Gene expression
normalized gene expression
0
1
A
Acinar-1
2,091
1,684
777
(p-value < 1e-12)
upregulated
2,851
1,181
1,064
(p-value < 1e-11)
downregulated
Acinar-1 shared pathways
mitochondrion organization
regulation of mRNA metabolic process
apoptotic signaling pathway
ncRNA metabolic process
purine-containing compound metabolic process
metabolism of amino acids and derivatives
metabolism of RNA
peptide biosynthetic process
neutrophil degranulation
membrane trafficking
supramolecular fiber organization
regulation of cellular response to stress
signaling by receptor tyrosine kinases
endocytosis
cell cycle phase transition
cellular component disassembly
PID MET PATHWAY
response to endoplasmic reticulum stress
cellular response to hormone stimulus
cellular protein catabolic process
up-regulated
down-regulated
Acinar-1 shared upregulated
Acinar-1 shared downregulated
Control
AAB+
T1D
T1D enriched
AAB+ enriched
3,015
4,137
Acinar-1 AAB+ and T1D enriched pathways
cellular component disassembly
mitochondrion organization
ribonucleoprotein complex biogenesis
peptide biosynthetic process
establishment of protein localization to organelle
cellular responses to stress
protein processing in endoplasmic reticulum
response to endoplasmic reticulum stress
transcriptional regulation by TP53
adaptive immune system
regulation of cellular response to stress
diseases of signal transduction
cellular protein catabolic process
autophagy
generation of precursor metabolites and energy
TCA cycle and respiratory electron transport
ncRNA metabolic process
asparagine N-linked glycosylation
Epstein-Barr virus infection
pathways in cancer
AAB+ enriched
T1D enriched
Acinar-1 AAB+ enriched
Acinar-1 T1D enriched
B
Acinar-2
2,990
534
220
(p-value < 1e-12)
upregulated
4,763
322
409
(p-value < 2.2e-16)
downregulated
Acinar-2 shared pathways
response to endoplasmic reticulum stress
protein targeting to ER
autophagy
PID CDC42 pathway
apoptotic signaling pathway
cellular protein catabolic process
negative regulation of intracellular signal transduction
dephosphorylation
pancreatic secretion
up-downregulated
up-regulated
Acinar-2 shared upregulated
Acinar-2 shared downregulated
Control
AAB+
T1D
T1D enriched
AAB+ enriched
1,544
1,398
Acinar-2 AAB+ and T1D enriched pathways
cytoplasmic translation
mitochondrion organization
cofactor metabolic process
ribonucleoprotein complex biogenesis
translation
granulocyte activation
TRBP containing complex
nucleotide metabolic process
cellular amino acid metabolic process
translational termination
ribonucleoprotein complex subunit organization
ER-Phagosome pathway
aerobic respiration
signaling by interleukins
small molecule catabolic process
asparagine N-linked glycosylation
apoptotic signaling pathway
response to toxic substance
positive regulation of organelle organization
regulation of establishment of protein localization
AAB+ enriched
T1D enriched
Acinar-2 AAB+ enriched
Acinar-2 T1D enriched
C
Alpha
1,363
5,155
419
(p-value < 1e-2)
upregulated
3,297
2,825
1,366
(p-value < 1e-11)
downregulated
Alpha shared pathways
ribonucleoprotein complex biogenesis
autophagy
aerobic respiration
protein localization
membrane trafficking
proton transmembrane transport
neutrophil mediated immunity
TP53 regulates metabolic genes
small molecule catabolic process
respiratory chain complex I (lambda subunit) mitochondrial
apoptotic signaling pathway
response to oxidative stress
regulation of protein complex assembly
cellular amino acid metabolic process
cofactor metabolic process
mitochondrion organization
translation
mitochondrial translation elongation
generation of precursor metabolites and energy
brain development
up-regulated
down-regulated
Alpha shared upregulated
Alpha shared downregulated
Control
AAB+
T1D
T1D enriched
AAB+ enriched
3,298
6,825
Alpha AAB+ and T1D enriched pathways
transcriptional regulation by TP53
apoptotic signaling pathway
cellular responses to stress
ribonucleoprotein complex biogenesis
cellular component disassembly
peptide biosynthetic process
mitochondrion organization
autophagy
regulation of cellular protein localization
response to endoplasmic reticulum stress
regulation of cell cycle process
positive regulation of organelle organization
nucleocytoplasmic transport
diseases of signal transduction
covalent chromatin modification
regulation of cellular response to stress
mRNA processing
cellular protein catabolic process
HIV infection
purine ribonucleotide metabolic process
AAB+ enriched
T1D enriched
Alpha AAB+ enriched
Alpha T1D enriched
D
Delta
325
3,884
69
(p-value = n.s)
upregulated
4,588
240
267
(p-value < 1e-11)
downregulated
Delta shared downregulated pathways
mitochondrion organization
influenza viral RNA
transcription & replication
nucleotide metabolic process
0
2
4
6
8
10
-log10(P)
Delta shared upregulated
Delta shared downregulated
Control
AAB+
T1D
T1D enriched
AAB+ enriched
459
8,819
Delta AAB+ and T1D enriched pathways
mitochondrion organization
regulation of mRNA metabolic process
organelle localization
regulation of cellular protein localization
autophagy
response to endoplasmic reticulum stress
apoptotic signaling pathway
positive regulation of organelle organization
ribonucleoprotein complex biogenesis
Protein processing in endoplasmic reticulum
translational initiation
oxidative phosphorylation
transcriptional regulation by TP53
protein folding
cellular responses to stress
adaptive immune system
vesicle-mediated transport
cellular protein catabolic process
signaling by receptor tyrosine kinases
signaling by interleukins
AAB+ enriched
T1D enriched
Delta AAB+ enriched
Delta T1D enriched

Figure S13

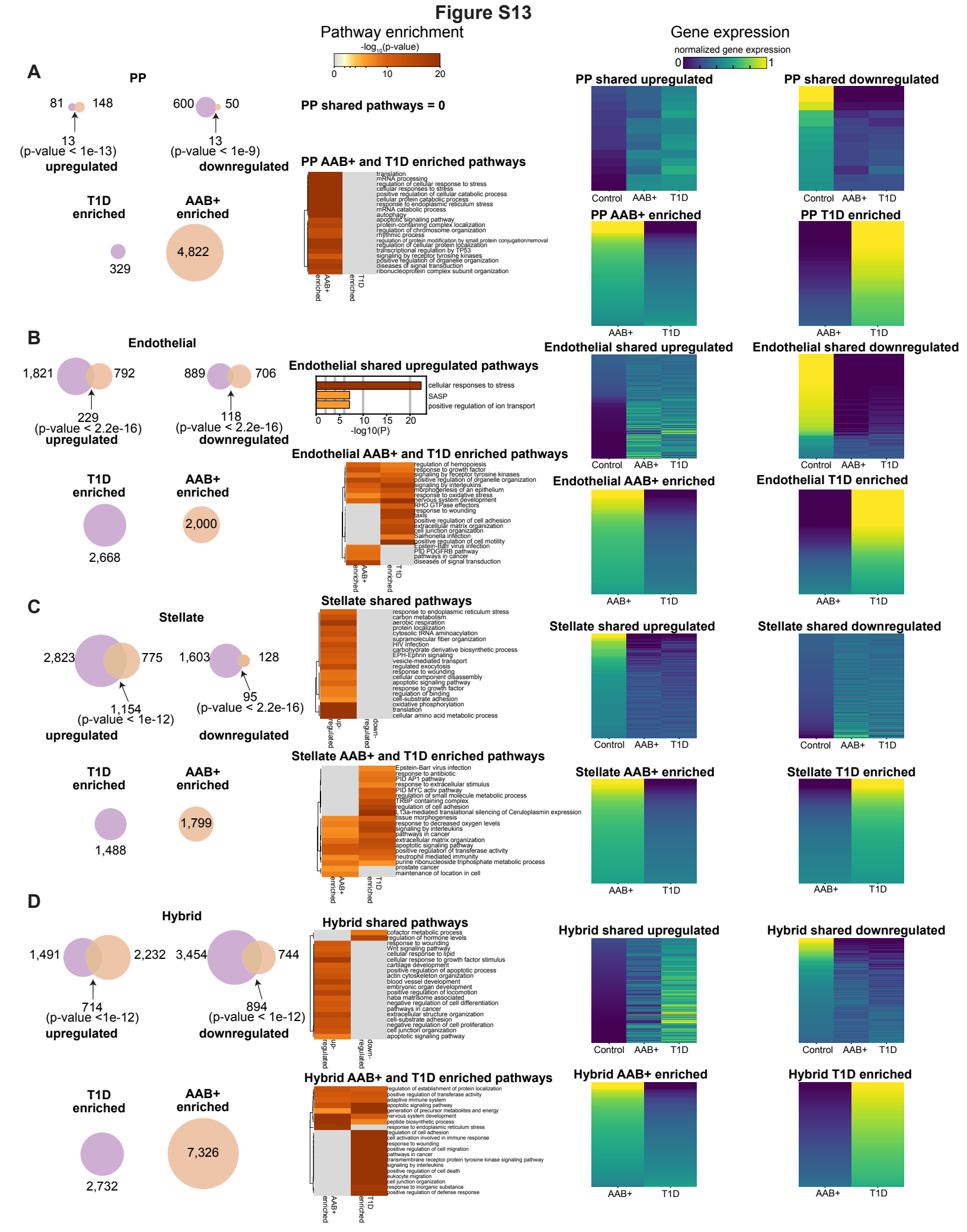

Pathway enrichment
-log10(p-value)
0
10
20
Gene expression
normalized gene expression
0
1
A
PP
81
148
13
(p-value < 1e-13)
upregulated
600
50
13
(p-value < 1e-9)
downregulated
T1D enriched
329
AAB+ enriched
4,822
PP shared pathways = 0
PP AAB+ and T1D enriched pathways
translation
mRNA processing
regulation of cellular response to stress
cellular responses to stress
positive regulation of cellular catabolic process
cellular protein catabolic process
response to endoplasmic reticulum stress
mRNA catabolic process
autophagy
apoptotic signaling pathway
protein-containing complex localization
regulation of chromosome organization
rhythmic process
regulation of protein modification by small protein conjugation/removal
regulation of cellular protein localization
transcriptional regulation by TP53
signaling by receptor tyrosine kinases
positive regulation of organelle organization
diseases of signal transduction
ribonucleoprotein complex subunit organization
AAB+ enriched
T1D enriched
PP shared upregulated
PP shared downregulated
Control
AAB+
T1D
PP AAB+ enriched
PP T1D enriched
B
Endothelial
1,821
792
229
(p-value < 2.2e-16)
upregulated
889
706
118
(p-value < 2.2e-16)
downregulated
T1D enriched
2,668
AAB+ enriched
2,000
Endothelial shared upregulated pathways
cellular responses to stress
SASP
positive regulation of ion transport
0
5
10
15
20
-log10(P)
Endothelial AAB+ and T1D enriched pathways
regulation of hemopoiesis
response to growth factor
signaling by receptor tyrosine kinases
positive regulation of organelle organization
signaling by interleukins
morphogenesis of an epithelium
response to oxidative stress
nervous system development
RHO GTPase effectors
response to wounding
taxis
positive regulation of cell adhesion
extracellular matrix organization
cell junction organization
Salmonella infection
positive regulation of cell motility
Epstein-Barr virus infection
PID PDGFRB pathway
pathways in cancer
diseases of signal transduction
AAB+ enriched
T1D enriched
Endothelial shared upregulated
Endothelial shared downregulated
Endothelial AAB+ enriched
Endothelial T1D enriched
C
Stellate
2,823
775
1,154
(p-value < 1e-12)
upregulated
1,603
128
95
(p-value < 2.2e-16)
downregulated
T1D enriched
1,488
AAB+ enriched
1,799
Stellate shared pathways
response to endoplasmic reticulum stress
carbon metabolism
aerobic respiration
protein localization
cytosolic tRNA aminoacylation
supramolecular fiber organization
HIV infection
carbohydrate derivative biosynthetic process
EPH-Ephrin signaling
vesicle-mediated transport
regulated exocytosis
response to wounding
cellular component disassembly
apoptotic signaling pathway
response to growth factor
regulation of binding
cell-substrate adhesion
oxidative phosphorylation
translation
cellular amino acid metabolic process
up-regulated
down-regulated
Stellate AAB+ and T1D enriched pathways
Epstein-Barr virus infection
response to antibiotic
PID AP1 pathway
response to extracellular stimulus
PID MYC activ pathway
regulation of small molecule metabolic process
TRBP containing complex
regulation of cell adhesion
L13a-mediated translational silencing of Ceruloplasmin expression
tissue morphogenesis
response to decreased oxygen levels
signaling by interleukins
pathways in cancer
extracellular matrix organization
apoptotic signaling pathway
positive regulation of transferase activity
neutrophil mediated immunity
purine ribonucleoside triphosphate metabolic process
prostate cancer
maintenance of location in cell
AAB+ enriched
T1D enriched
Stellate shared upregulated
Stellate shared downregulated
Stellate AAB+ enriched
Stellate T1D enriched
D
Hybrid
1,491
2,232
714
(p-value <1e-12)
upregulated
3,454
744
894
(p-value < 1e-12)
downregulated
T1D enriched
2,732
AAB+ enriched
7,326
Hybrid shared pathways
cofactor metabolic process
regulation of hormone levels
response to wounding
Wnt signaling pathway
cellular response to lipid
cellular response to growth factor stimulus
cartilage development
positive regulation of apoptotic process
actin cytoskeleton organization
blood vessel development
embryonic organ development
positive regulation of locomotion
naba matrisome associated
negative regulation of cell differentiation
pathways in cancer
extracellular structure organization
cell-substrate adhesion
negative regulation of cell proliferation
cell junction organization
apoptotic signaling pathway
up-regulated
down-regulated
Hybrid AAB+ and T1D enriched pathways
regulation of establishment of protein localization
positive regulation of transferase activity
adaptive immune system
apoptotic signaling pathway
generation of precursor metabolites and energy
nervous system development
peptide biosynthetic process
response to endoplasmic reticulum stress
regulation of cell adhesion
cell activation involved in immune response
response to wounding
positive regulation of cell migration
pathways in cancer
transmembrane receptor protein tyrosine kinase signaling pathway
signaling by interleukins
positive regulation of cell death
eukocyte migration
cell junction organization
response to inorganic substance
positive regulation of defense response
AAB+ enriched
T1D enriched
Hybrid shared upregulated
Hybrid shared downregulated
Hybrid AAB+ enriched
Hybrid T1D enriched

Figure S14

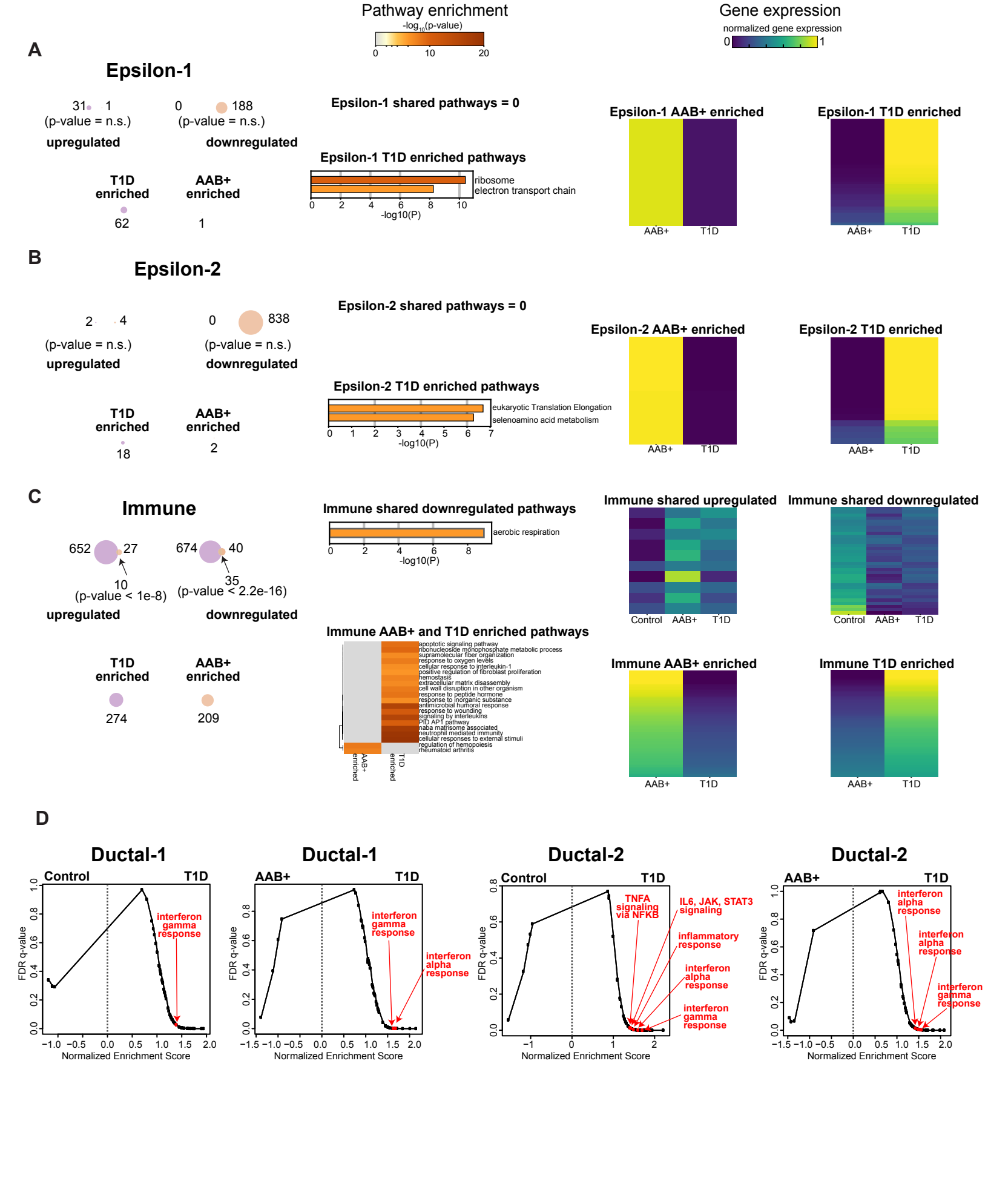

Pathway enrichment
-log10(p-value)
0
10
20
Gene expression
normalized gene expression
0
1
A
Epsilon-1
31
1
(p-value = n.s.)
upregulated
0
188
(p-value = n.s.)
downregulated
T1D enriched
62
AAB+ enriched
1
Epsilon-1 shared pathways = 0
Epsilon-1 T1D enriched pathways
ribosome
electron transport chain
-log10(P)
Epsilon-1 AAB+ enriched
Epsilon-1 T1D enriched
AAB+
T1D
B
Epsilon-2
2
4
(p-value = n.s.)
upregulated
0
838
(p-value = n.s.)
downregulated
T1D enriched
18
AAB+ enriched
2
Epsilon-2 shared pathways = 0
Epsilon-2 T1D enriched pathways
eukaryotic Translation Elongation
selenoamino acid metabolism
-log10(P)
Epsilon-2 AAB+ enriched
Epsilon-2 T1D enriched
AAB+
T1D
C
Immune
652
27
10
(p-value < 1e-8)
upregulated
674
40
35
(p-value < 2.2e-16)
downregulated
T1D enriched
274
AAB+ enriched
209
Immune shared downregulated pathways
aerobic respiration
-log10(P)
Immune AAB+ and T1D enriched pathways
apoptotic signaling pathway
ribonucleoside monophosphate metabolic process
supramolecular fiber organization
response to oxygen levels
cellular response to interleukin-1
positive regulation of fibroblast proliferation
hemostasis
extracellular matrix disassembly
cell wall disruption in other organism
response to peptide hormone
response to inorganic substance
antimicrobial humoral response
response to wounding
signaling by interleukins
PID AP1 pathway
naba matrisome associated
neutrophil mediated immunity
cellular responses to external stimuli
regulation of hemopoiesis
rheumatoid arthritis
AAB+ enriched
T1D enriched
Immune shared upregulated
Control
AAB+
T1D
Immune shared downregulated
Control
AAB+
T1D
Immune AAB+ enriched
Immune T1D enriched
AAB+
T1D
D
Ductal-1
Control
T1D
FDR q-value
interferon gamma response
Normalized Enrichment Score
Ductal-1
AAB+
T1D
interferon gamma response
interferon alpha response
Normalized Enrichment Score
Ductal-2
Control
T1D
TNFA signaling via NFKB
IL6, JAK, STAT3 signaling
inflammatory response
interferon alpha response
interferon gamma response
Normalized Enrichment Score
Ductal-2
AAB+
T1D
interferon alpha response
interferon alpha response
interferon gamma response
Normalized Enrichment Score

Figure S15

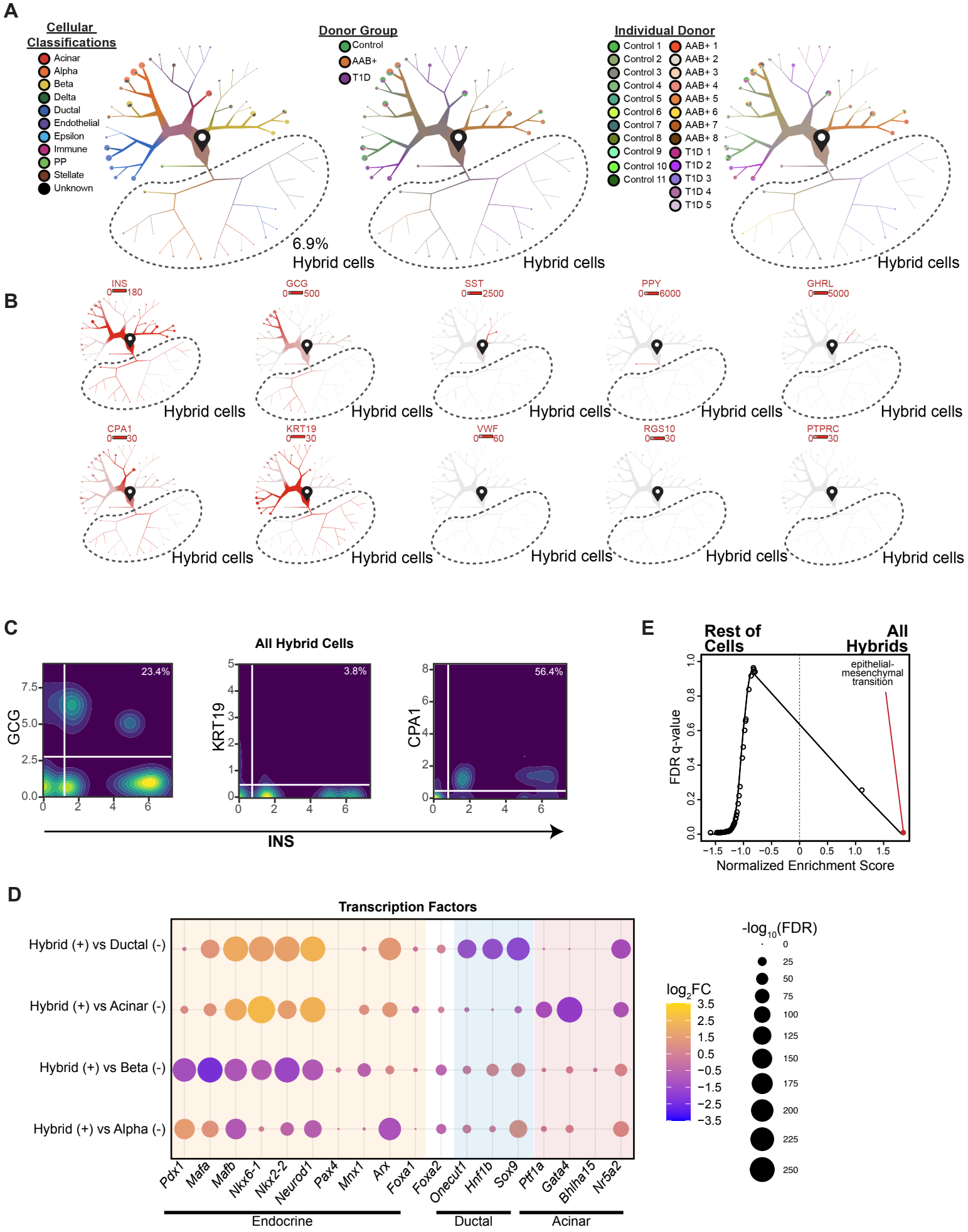
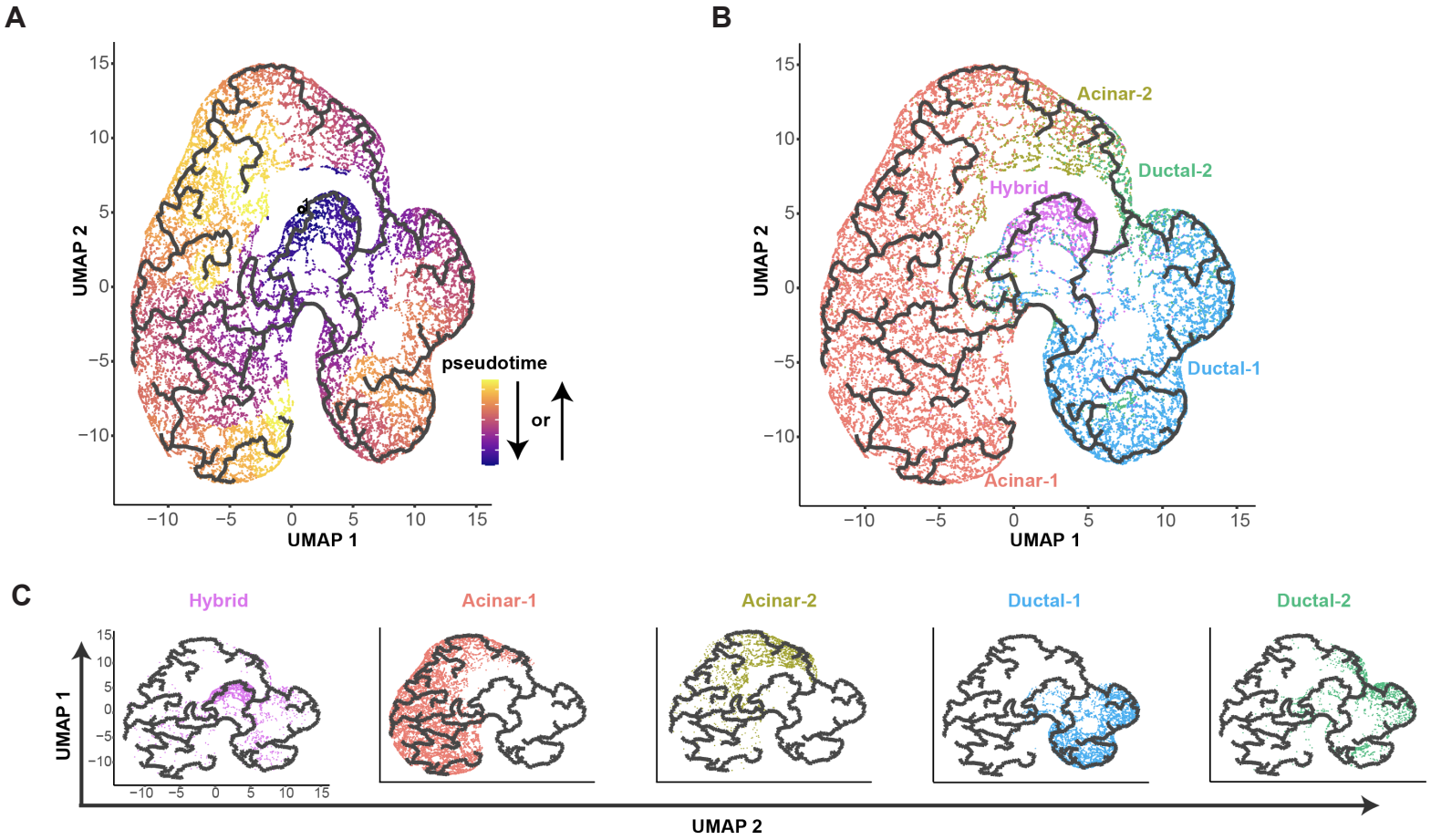

A
Cellular Classifications
Acinar
Alpha
Beta
Delta
Ductal
Endothelial
Epsilon
Immune
PP
Stellate
Unknown
6.9%
Hybrid cells
Donor Group
Control
AAB+
T1D
Hybrid cells
Individual Donor
Control 1
Control 2
Control 3
Control 4
Control 5
Control 6
Control 7
Control 8
Control 9
Control 10
Control 11
AAB+ 1
AAB+ 2
AAB+ 3
AAB+ 4
AAB+ 5
AAB+ 6
AAB+ 7
AAB+ 8
T1D 1
T1D 2
T1D 3
T1D 4
T1D 5
Hybrid cells
B
INS 0–180
GCG 0–500
SST 0–2500
PPY 0–6000
GHRL 0–5000
CPA1 0–30
KRT19 0–30
VWF 0–60
RGS10 0–30
PTPRC 0–30
Hybrid cells
C
All Hybrid Cells
23.4%
3.8%
56.4%
GCG
KRT19
CPA1
INS
E
Rest of Cells
All Hybrids
epithelial-mesenchymal transition
FDR q-value
Normalized Enrichment Score
D
Transcription Factors
Hybrid (+) vs Ductal (-)
Hybrid (+) vs Acinar (-)
Hybrid (+) vs Beta (-)
Hybrid (+) vs Alpha (-)
Pdx1
Mafa
Mafb
Nkx6-1
Nkx2-2
Neurod1
Pax4
Mnx1
Arx
Foxa1
Foxa2
Onecut1
Hnf1b
Sox9
Ptf1a
Gata4
Bhlha15
Nr5a2
Endocrine
Ductal
Acinar
log2FC
-log10(FDR)

Figure S16

Monocle Trajectory Analysis of Relevant Exocrine-related cell types: Hybrid, Acinar-1, Acinar-2, Ductal-1, and Ductal-2



Monocle Trajectory Analysis of Relevant Endocrine-related cell types: Hybrid, Alpha, Beta-1, Beta-2, and Ductal-1

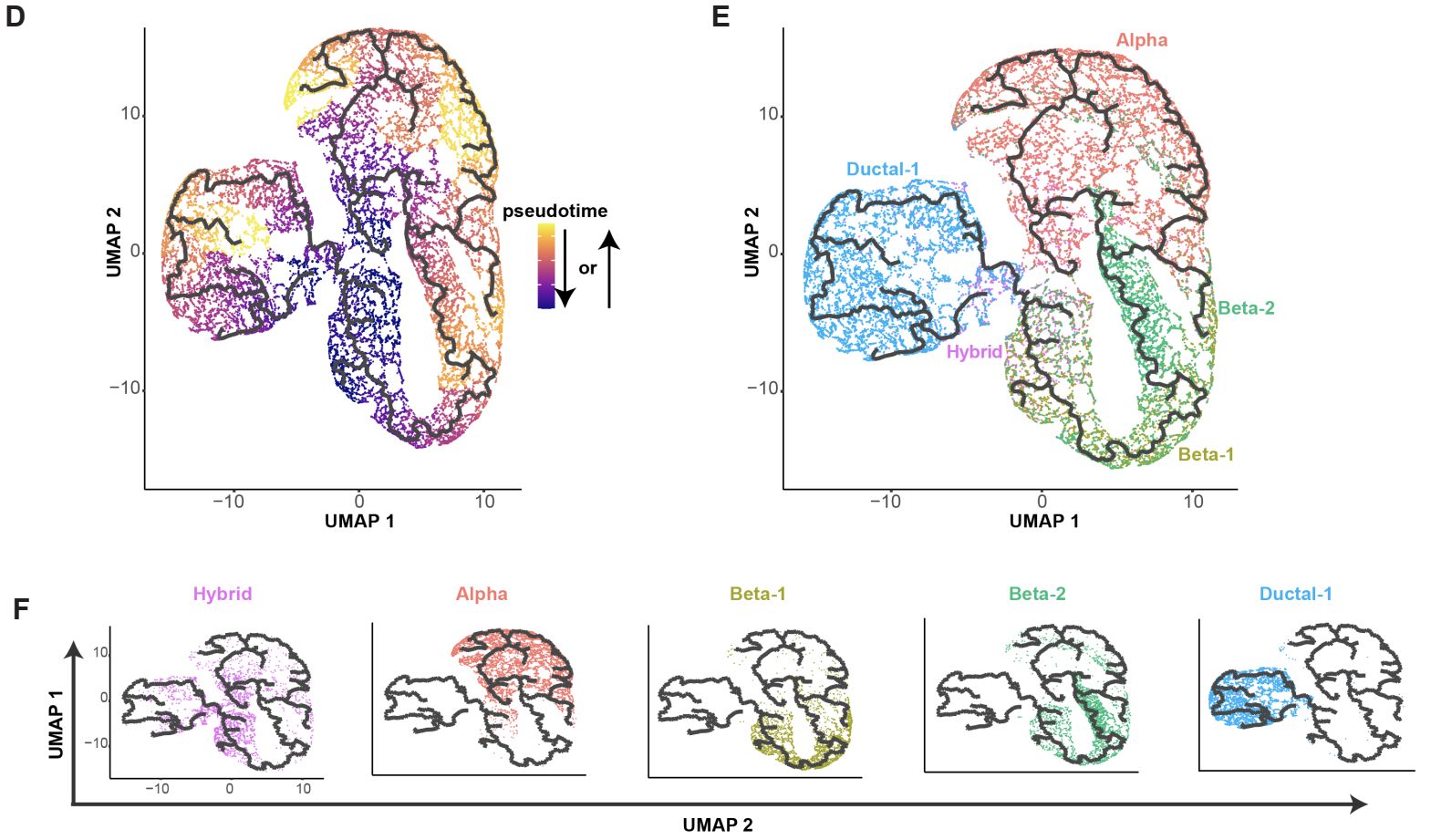


Figure S17

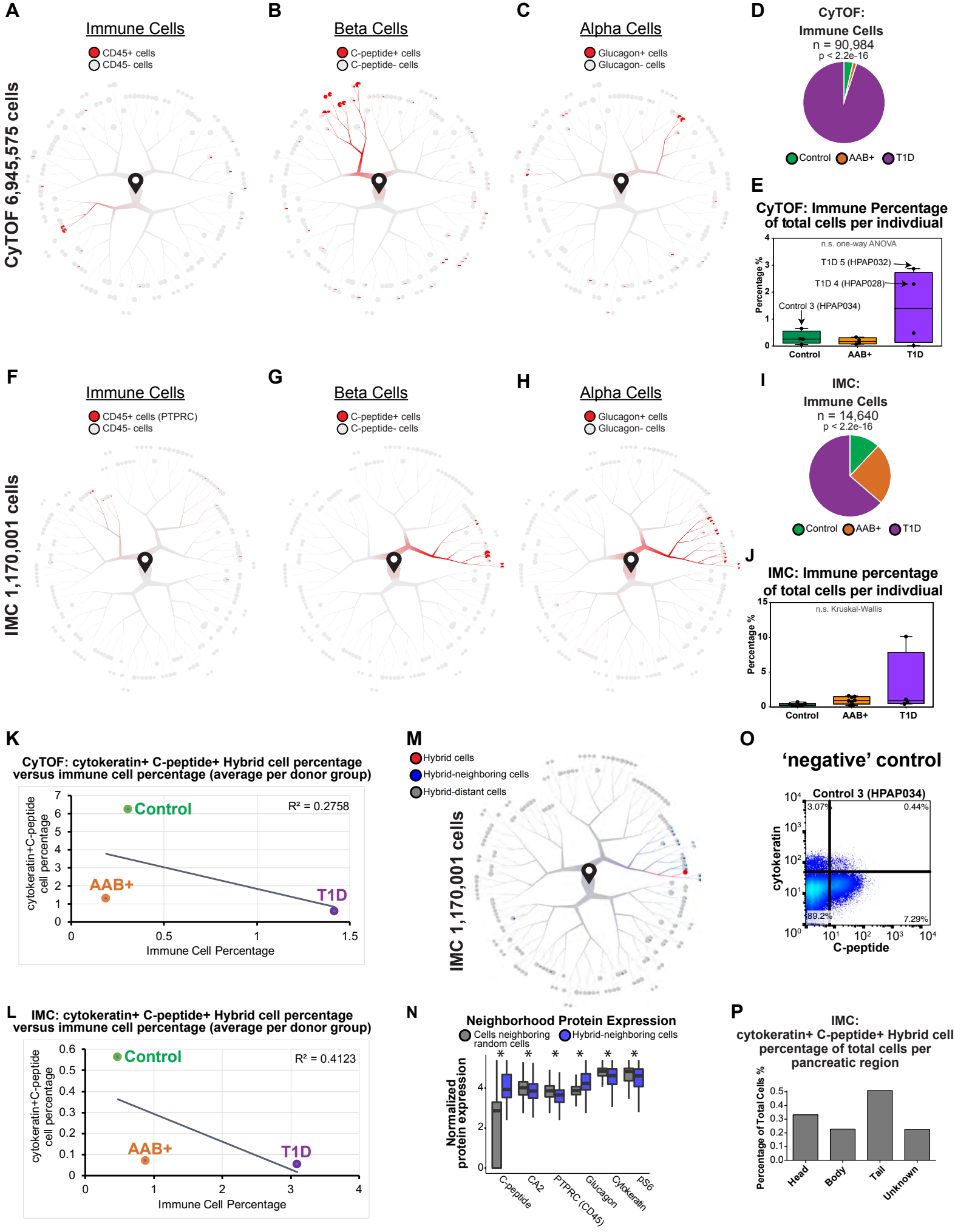

A
Immune Cells
CD45+ cells
CD45- cells
B
Beta Cells
C-peptide+ cells
C-peptide- cells
C
Alpha Cells
Glucagon+ cells
Glucagon- cells
CyTOF 6,945,575 cells
D
CyTOF:
Immune Cells
n = 90,984
p < 2.2e-16
Control
AAB+
T1D
E
CyTOF: Immune Percentage of total cells per indivdiual
n.s. one-way ANOVA
T1D 5 (HPAP032)
T1D 4 (HPAP028)
Control 3 (HPAP034)
Percentage %
Control
AAB+
T1D
F
Immune Cells
CD45+ cells (PTPRC)
CD45- cells
G
Beta Cells
C-peptide+ cells
C-peptide- cells
H
Alpha Cells
Glucagon+ cells
Glucagon- cells
IMC 1,170,001 cells
I
IMC:
Immune Cells
n = 14,640
p < 2.2e-16
Control
AAB+
T1D
J
IMC: Immune percentage of total cells per indivdiual
n.s. Kruskal-Wallis
Percentage %
Control
AAB+
T1D
K
CyTOF: cytokeratin+ C-peptide+ Hybrid cell percentage versus immune cell percentage (average per donor group)
Control
AAB+
T1D
R² = 0.2758
cytokeratin+C-peptide cell percentage
Immune Cell Percentage
M
Hybrid cells
Hybrid-neighboring cells
Hybrid-distant cells
IMC 1,170,001 cells
O
'negative' control
Control 3 (HPAP034)
3.07%
0.44%
89.2%
7.29%
cytokeratin
C-peptide
L
IMC: cytokeratin+ C-peptide+ Hybrid cell percentage versus immune cell percentage (average per donor group)
Control
AAB+
T1D
R² = 0.4123
cytokeratin+C-peptide cell percentage
Immune Cell Percentage
N
Neighborhood Protein Expression
Cells neighboring random cells
Hybrid-neighboring cells
Normalized protein expression
C-peptide
CA2
PTPRC (CD45)
Glucagon
Cytokeratin
pS6
P
IMC:
cytokeratin+ C-peptide+ Hybrid cell percentage of total cells per pancreatic region
Percentage of Total Cells %
Head
Body
Tail
Unknown

Figure S18

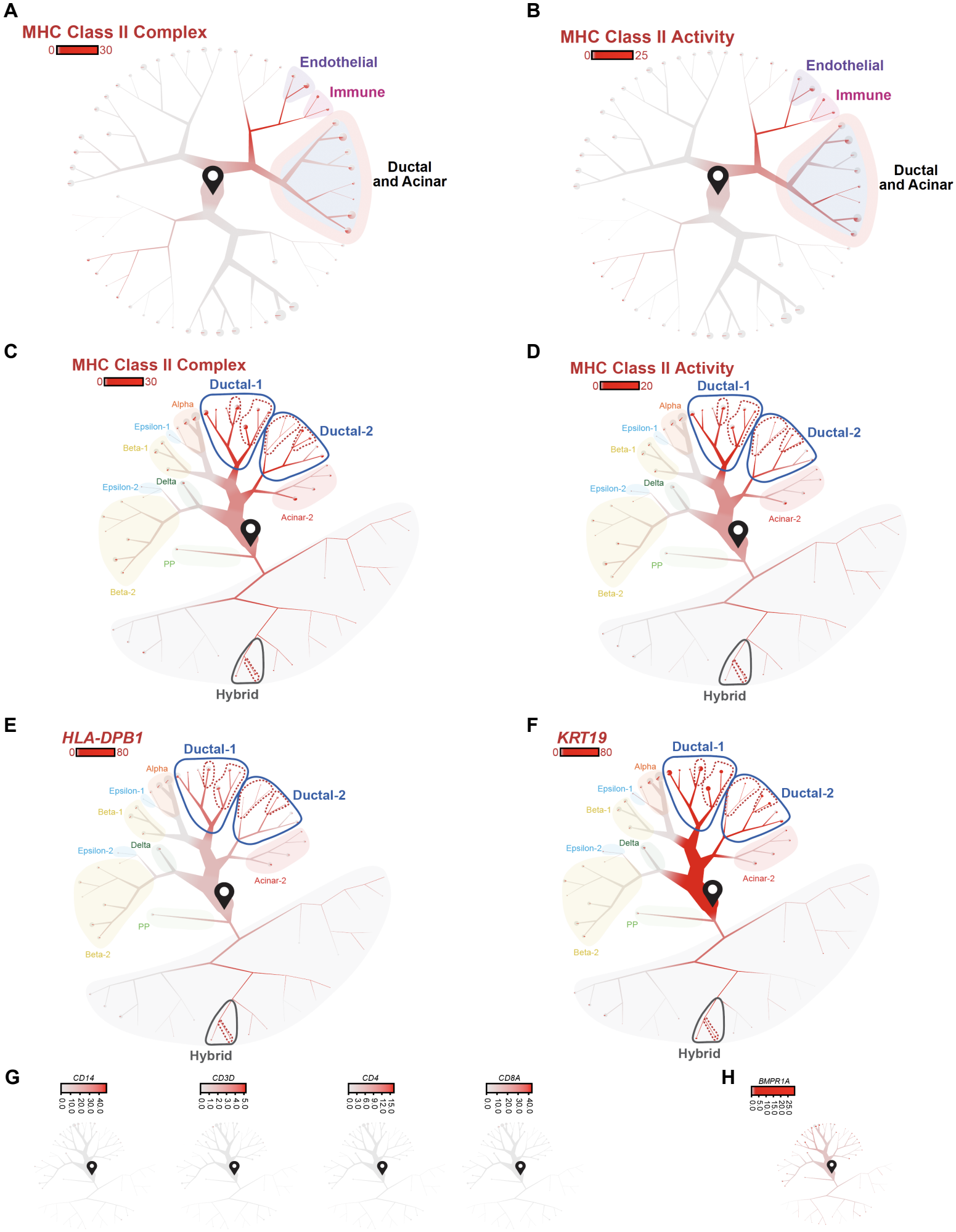

A
MHC Class II Complex
0
30
Endothelial
Immune
Ductal
and Acinar
B
MHC Class II Activity
0
25
Endothelial
Immune
Ductal
and Acinar
C
MHC Class II Complex
0
30
Ductal-1
Ductal-2
Alpha
Epsilon-1
Beta-1
Epsilon-2
Delta
Acinar-2
PP
Beta-2
Hybrid
D
MHC Class II Activity
0
20
Ductal-1
Ductal-2
Alpha
Epsilon-1
Beta-1
Epsilon-2
Delta
Acinar-2
PP
Beta-2
Hybrid
E
HLA-DPB1
0
80
Ductal-1
Ductal-2
Alpha
Epsilon-1
Beta-1
Epsilon-2
Delta
Acinar-2
PP
Beta-2
Hybrid
F
KRT19
0
80
Ductal-1
Ductal-2
Alpha
Epsilon-1
Beta-1
Epsilon-2
Delta
Acinar-2
PP
Beta-2
Hybrid
G
CD14
0.0 10.0 20.0 30.0 40.0
CD3D
0.0 1.0 2.0 3.0 4.0 5.0
CD4
0.0 3.0 6.0 9.0 12.0 15.0
CD8A
0.0 10.0 20.0 30.0 40.0
H
BMPR1A
0.0 5.0 10.0 15.0 20.0 25.0

Figure S19

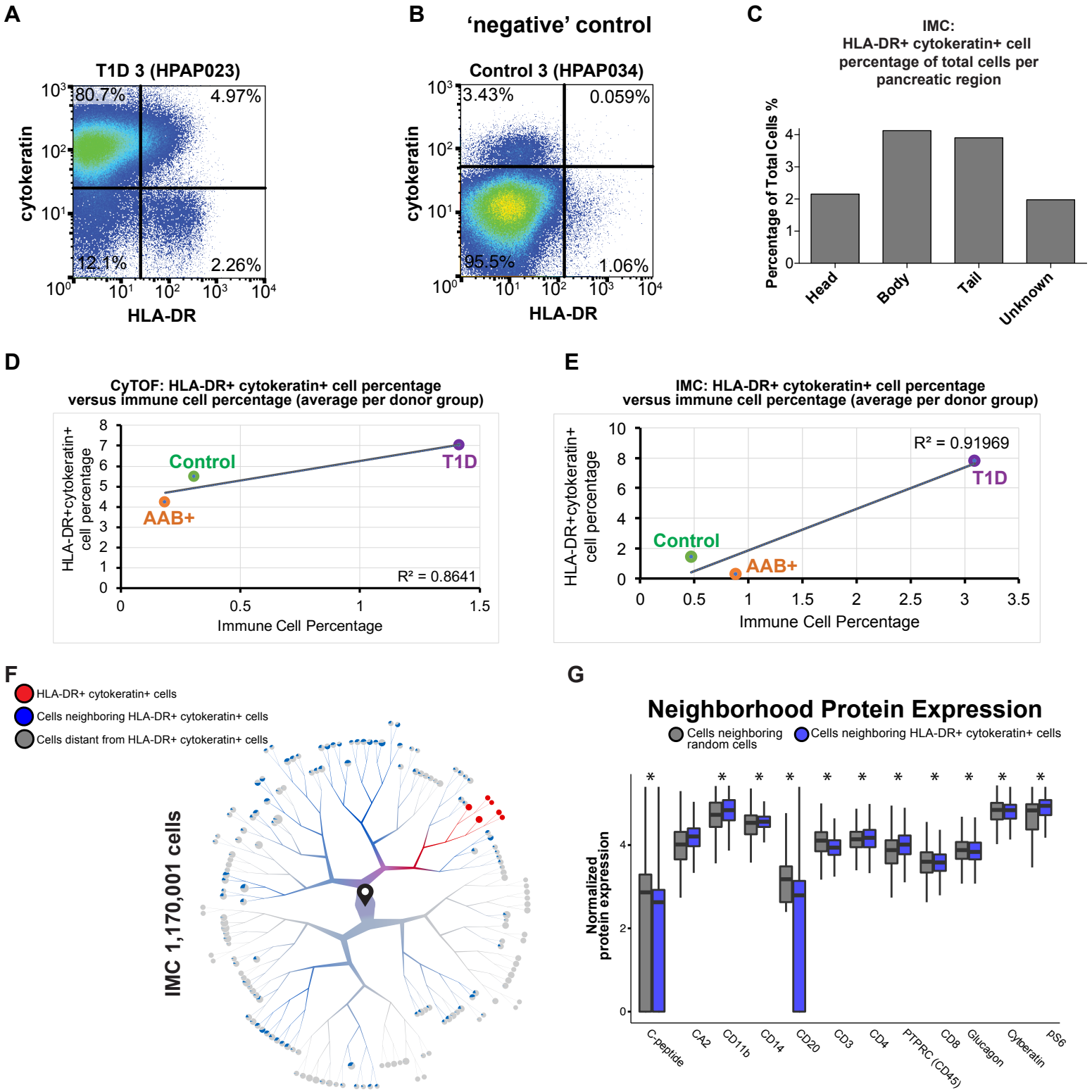

A
T1D 3 (HPAP023)
80.7%
4.97%
12.1%
2.26%
cytokeratin
HLA-DR
B
'negative' control
Control 3 (HPAP034)
3.43%
0.059%
95.5%
1.06%
cytokeratin
HLA-DR
C
IMC:
HLA-DR+ cytokeratin+ cell percentage of total cells per pancreatic region
Percentage of Total Cells %
Head
Body
Tail
Unknown
D
CyTOF: HLA-DR+ cytokeratin+ cell percentage versus immune cell percentage (average per donor group)
HLA-DR+cytokeratin+ cell percentage
Control
T1D
AAB+
R² = 0.8641
Immune Cell Percentage
E
IMC: HLA-DR+ cytokeratin+ cell percentage versus immune cell percentage (average per donor group)
HLA-DR+cytokeratin+ cell percentage
R² = 0.91969
T1D
Control
AAB+
Immune Cell Percentage
F
HLA-DR+ cytokeratin+ cells
Cells neighboring HLA-DR+ cytokeratin+ cells
Cells distant from HLA-DR+ cytokeratin+ cells
IMC 1,170,001 cells
G
Neighborhood Protein Expression
Cells neighboring random cells
Cells neighboring HLA-DR+ cytokeratin+ cells
Normalized protein expression
C-peptide
CA2
CD11b
CD14
CD20
CD3
CD4
PTPRC (CD45)
CD8
Glucagon
Cytokeratin
pS6

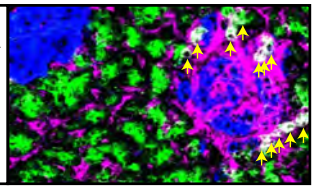
Figure S20

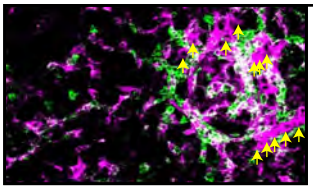
HLA-DR CD99 KRT | HLA-DR CD45 | MHC (HLA) Class I and II genetics

T1D

Donor A HPAP020 Tail ROI #1

Susceptible genotype: DR3

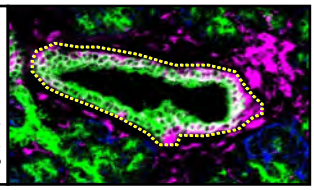


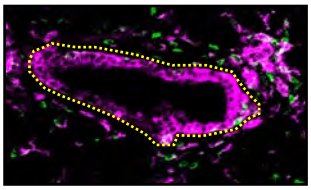


		allele 1	allele 2			allele 1	allele 2
Class I	A	01:01	02:01	Class II	DRB1	03:01	13:02
	B	45:01	53:01		DRB3	02:02	03:01
	C	06:02	04:01		DRB4	--------	--------
					DRB5	--------	--------
					DQB1	02:01	06:04
					DQA1	01:02	05:01
					DPB1	04:01	104:01
					DPA1	01:03	--------

Donor B HPAP002 Head ROI #2

*Susceptible genotype: DR4 & HLA-A*24:02*

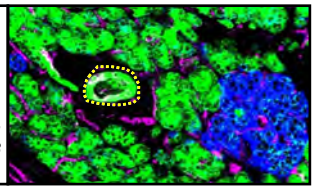


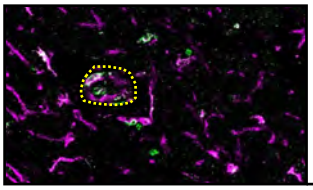


		allele 1	allele 2			allele 1	allele 2
Class I	A	02:01	24:02	Class II	DRB1	13:02	04:08
	B	35:17	41:01		DRB3	03:01	--------
	C	04:01	17:01		DRB4	--------	01:03
					DRB5	--------	--------
					DQB1	03:02	06:04
					DQA1	01:02	03:01
					DPB1	02:01	04:02
					DPA1	01:03	--------

Donor C HPAP023 Body ROI #2

Susceptible genotype: DR3



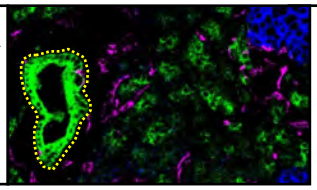


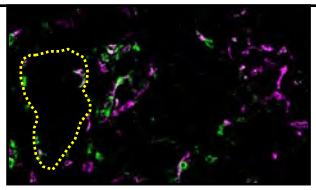
		allele 1	allele 2			allele 1	allele 2
Class I	A	02:01	33:01	Class II	DRB1	01:02	03:01
	B	08:01	14:02		DRB3	--------	01:01
	C	07:01	08:02		DRB4	--------	--------
					DRB5	--------	--------
					DQB1	02:01	05:01
					DQA1	01:01	05:01
					DPB1	04:01	34:01
					DPA1	01:03	--------

AAB+

Donor A HPAP019 Tail ROI #1

Susceptible genotype: DR4

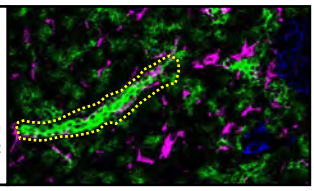


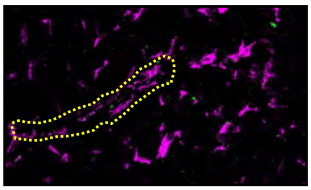


		allele 1	allele 2			allele 1	allele 2
Class I	A	03:01	11:01	Class II	DRB1	01:01	04:04
	B	07:02	27:05		DRB3	--------	--------
	C	07:02	01:02		DRB4	01:03	--------
					DRB5	--------	--------
					DQB1	03:02	05:01
					DQA1	01:01	03:01
					DPB1	02:01	04:01
					DPA1	01:03	--------

Donor B HPAP029 Tail ROI #1

Susceptible genotype: DR4

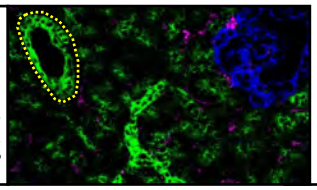


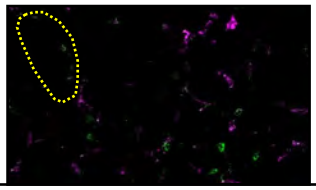


		allele 1	allele 2			allele 1	allele 2
Class I	A	03:01	32:01	Class II	DRB1	04:01	04:08
	B	15:01	51:01		DRB3	--------	--------
	C	03:04	15:02		DRB4	01:03	--------
					DRB5	--------	--------
					DQB1	03:01	03:02
					DQA1	03:01	03:03
					DPB1	03:01	04:01
					DPA1	01:03	--------

Donor C HPAP003 Body ROI #6

*Susceptible genotype: DR4 & HLA-A*24:02*



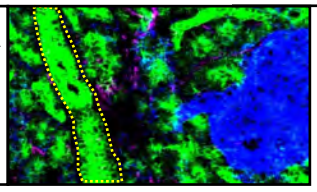


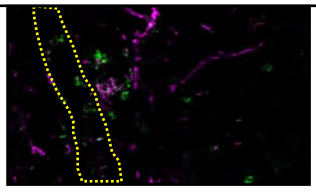
		allele 1	allele 2			allele 1	allele 2
Class I	A	02:01	24:02	Class II	DRB1	04:01	09:01
	B	44:02	51:01		DRB3	--------	--------
	C	03:03	05:01		DRB4	01:03	--------
					DRB5	--------	--------
					DQB1	03:02	03:03
					DQA1	03:01	03:02
					DPB1	02:01	04:01
					DPA1	01:03	--------

Control

Donor A HPAP006 Body ROI #4

*Susceptible genotype: HLA-A*24:02 (no Class II susceptibility)*

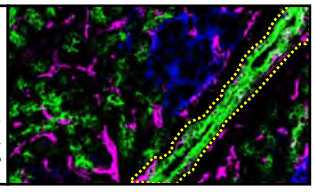


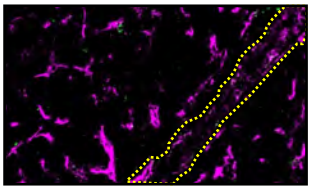


		allele 1	allele 2			allele 1	allele 2
Class I	A	24:02	25:01	Class II	DRB1	11:03	13:01
	B	15:01	18:01		DRB3	01:01	02:02
	C	03:03	12:03		DRB4	--------	--------
					DRB5	--------	--------
					DQB1	03:01	06:03
					DQA1	01:03	05:05
					DPB1	02:01	04:01
					DPA1	01:03	--------

Donor B HPAP026 Head ROI #2

*Protective genotype: HLA-DRB1*15:01-HLA-DQB1*06:02*

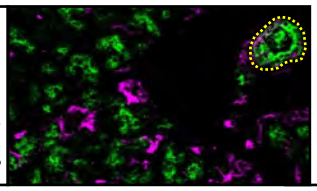


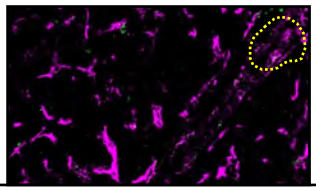


		allele 1	allele 2			allele 1	allele 2
Class I	A	02:01	-------	Class II	DRB1	01:01	15:01
	B	15:01	44:02		DRB3	--------	--------
	C	03:03	05:01		DRB4	--------	--------
					DRB5	--------	01:01
					DQB1	05:01	06:02
					DQA1	01:01	01:02
					DPB1	04:01	04:02
					DPA1	01:03	--------

Donor C HPAP014 Head ROI #1

*Susceptible genotype: DR4 & HLA-A*24:02*





		allele 1	allele 2			allele 1	allele 2
Class I	A	01:01	24:02	Class II	DRB1	01:01	04:01
	B	27:05	35:01		DRB3	--------	--------
	C	04:01	05:01		DRB4	--------	01:03
					DRB5	--------	--------
					DQB1	03:02	05:01
					DQA1	01:01	03:01
					DPB1	04:01	04:02
					DPA1	01:03	--------

Figure S1: Cell numbers and clustering before complete filtering

(A) Pie chart displaying the cell numbers/proportions of each individual donor per donor type.

(B) Box plot displaying the average gene number per cell per donor type.

(C) UMAP visualization of cell clusters for all cells.

(D) Doublets and singlets, as identified using DoubletFinder, across cell clusters visualized by UMAP.

(E) UMAP visualization of the normalized gene expression counts of each canonical gene marker of each major cell type.

Figure S2: Doublet removal and UMI counts

(A) Doublets and singlets, as identified using Scrublet, across cell clusters visualized by UMAP per individual.

(B) Venn diagram indicating the number of cells deemed doublets by DoubletFinder and Scrublet, as well as cells that were commonly identified by both approaches.

(C) Table indicating the number of cells removed and the resulting total cell number for each step of filtering.

(D) Unique molecular identifier (UMI) counts per cell projected across the dendrogram visualization and clustering of all cells from Figure 1C. Pie charts at the end of the branches display the breakdown of UMI counts per cell within that terminal cluster. Cells begin at the start pin symbol, and from there are partitioned based on similarities and differences in gene expression.

(E) UMI counts per cell projected across the dendrogram visualization and clustering of ductal and endocrine cells from Figure 1D. Pie charts at the end of the branches display the breakdown of UMI counts per cell within that terminal cluster. Cells begin at the start pin symbol, and from there are partitioned based on similarities and differences in gene expression.

Figure S3: Cell numbers and clustering after complete filtering

(A) Pie chart displaying the cell numbers/proportions of each individual donor per donor type.

(B) UMAP visualization of cell clusters for all cells.

(C) UMAP visualization donor groups across clusters for all cells.

(D) UMAP visualization of Garnett cellular classifications across clusters for all cells.

(E) UMAP visualization of the normalized gene expression counts of each canonical gene marker of each major cell type.

Figure S4: Marker gene expression confirms canonical cell types

(A) Dendrograms highlighting the expression of each canonical gene marker of each major cell type across the dendrogram of all cells in Figure 1C. Scale bars represent normalized transcript numbers. Cells begin at the start pin symbol, and from there are partitioned based on similarities and differences in gene expression.

(B) Dendrograms highlighting the expression of each canonical gene marker of each major cell type across the dendrogram of ductal and endocrine cells in Figure 1D. Scale bars represent normalized transcript numbers. Cells begin at the start pin symbol, and from there are partitioned based on similarities and differences in gene expression.

(C) Schematic of the human pancreatic islet anatomy and major cell types.

(D) Pie chart displaying the cell numbers/proportions of each cell type defined in Figure 1, C and D.

Figure S5: Technical factors or biological factors do not drive cell segregation

(A) Box plot displaying clumpiness measure across biological and technical factors.

(B-K) Cells begin at the start pin symbol, and from there are partitioned based on similarities and differences in gene expression. (Top) Technical or biological factors projected across the dendrogram visualization and clustering of all cells from Figure 1C. Pie charts at the end of the branches display the breakdown of technical or biological factors within that terminal cluster. (Bottom) Quantification of the relationship between different technical or biological factors using the clumpiness measure as determined from the dendrograms. Clumpiness is a measure of aggregation of labels within a dendrogram. Here, each leaf of the dendrogram contains a collection of labels (technical or biological covariates). The more labels

group together within the dendrogram, the higher the clumpiness value. When using hierarchy in biological systems to elucidate relationships between metadata, the distribution of metadata labels within the hierarchy may exhibit different levels of aggregation—uniform, random, or clumped. Our strategy for estimating the impact of technical and biological variance on substructures beyond cell type identity was to use a measure called “clumpiness”. Clumpiness is a measure for finding the level of aggregation between labels distributed among the leaves of a hierarchical tree and extensively measures the relationships between metadata. We have shown that our measure is robust to the cell number, data size, and label size, and thus is scale invariant. In addition, our measure is generalizable to more than two labels and is efficiently computable and maintainable. In the revised manuscript Supplemental Figures S5 and S6, we demonstrate that although cells from individual donors tend to cluster with each other as expected, no particular feature affected cell type identity definition in our analytical approach.

Figure S6: Technical factors or biological factors do not drive cell segregation

(A-I) Cells begin at the start pin symbol, and from there are partitioned based on similarities and differences in gene expression. (Top) Technical or biological factors projected across the dendrogram visualization and clustering of all cells from Figure 1C. Pie charts at the end of the branches display the breakdown of technical or biological factors within that terminal cluster. (Bottom) Quantification of the relationship between different technical or biological factors using the clumpiness measure as determined from the dendrograms. Clumpiness is a measure of aggregation of labels within a dendrogram. Here, each leaf of the dendrogram contains a collection of labels (technical or biological covariates). The more labels group together within the dendrogram, the higher the clumpiness value. When using hierarchy in biological systems to elucidate relationships between metadata, the distribution of metadata labels within the hierarchy may exhibit different levels of aggregation—uniform, random, or clumped. Our strategy for estimating the impact of technical and biological variance on substructures beyond cell type identity was to use a measure called “clumpiness”. Clumpiness is

a measure for finding the level of aggregation between labels distributed among the leaves of a hierarchical tree and extensively measures the relationships between metadata. We have shown that our measure is robust to the cell number, data size, and label size, and thus is scale invariant. In addition, our measure is generalizable to more than two labels and is efficiently computable and maintainable. In the revised manuscript Supplemental Figures S5 and S6, we demonstrate that although cells from individual donors tend to cluster with each other as expected, no particular feature affected cell type identity definition in our analytical approach.

Figure S7: Discernment of cell types when clustering cells from each disease state independently

(A-C) Cells begin at the start pin symbol, and from there are partitioned based on similarities and differences in gene expression. Dendrogram visualization and clustering of all cells from each of the donor states independently. Cells begin at the start pin symbol. Garnett cellular classifications (Top Left) and individual donor (Bottom Left) labels across the dendrogram. (Right) Pie chart displaying the cell proportions of each cell type.

(D-F) Dendrogram visualization of the expression of each canonical gene marker of each major cell type across the dendrograms from A-C. Scale bars represent normalized transcript numbers.

Figure S8: Discernment of cell types when clustering cells from each disease state independently

(A) Cells begin at the start pin symbol, and from there are partitioned based on similarities and differences in gene expression. (Inset) Dendrogram visualization and clustering of all cells from Figure 1C. Highlighted in pink is the immune cell cluster, which was subsetted and re-clustered here. (Main) Dendrogram visualization of clustered immune cells (along with small clusters of Stellate and Schwann cells) with Garnett cellular classification (left), donor group (middle), or individual donor (right) designations across the tree.

(B) Dendrograms highlighting the expression of each canonical gene marker of each relevant cell type across the dendrogram of immune cells in Figure S8A. Scale bars represent normalized transcript numbers.

(C) Gene set enrichment of ImmGen cell types for the genes upregulated in the *PTPRC* high immune cell cluster (pink in fig. S8A) when compared to the closely transcriptionally related stellate (brown in fig. S8A) plus Schwann cells (grey in fig. S8A).

(D) Bar graph displaying the proportion of immune cells from each donor group. P-value presented is the result of the Chi-squared test.

(E) Box plots displaying immune cell percentage of total cells per individual across donor groups.

Figure S9: Discernment of cell types when clustering cells from each disease state independently

(A) Bar graph displaying the proportion of cells from each donor group for all major pancreatic cell types. P-values presented are the result of the Chi-squared test.

(B) Box plots displaying cell type percentage of total cells per individual across donor groups for all major pancreatic cell types.

Figure S10: Gene and gene ontology pathways that are shared and different across disease states in beta cells

(A-B) (Left) For each cell type, displayed are gene ontology pathways that are shared across T1D and AAB+ cells when compared to Control cells (top) or pathways that are differently enriched in T1D cells vs AAB+ cells (bottom). The top 20 clusters are displayed and a stringent cut-off of $p \leq 1e\text{-}6$ was applied to determine significant gene ontology pathways. Pathways highlighted in yellow are displayed in Figure 2. (Right) Heatmaps displaying the degree of gene expression changes of genes (rows) that are shared (top) or differential (bottom) across AAB+ and T1D disease states.

Figure S11: Gene and gene ontology pathways that are shared and different across disease states in ductal cells

(A-B) (Left) For each cell type, displayed are gene ontology pathways that are shared across T1D and AAB+ cells when compared to Control cells (top) or pathways that are differently enriched in T1D cells vs AAB+ cells (bottom). The top 20 clusters are displayed and a stringent cut-off of $p \leq 1e\text{-}6$ was applied to determine significant gene ontology pathways. Pathways highlighted in yellow are displayed in Figure 2. (Right) Heatmaps displaying the degree of gene expression changes of genes (rows) that are shared (top) or differential (bottom) across AAB+ and T1D disease states.

Figure S12: Gene and gene ontology pathways that are shared and different across disease states in Acinar-1, Acinar-2, Alpha, and Delta cells

(A-D) (Left) For each cell type, Venn diagrams indicate the numbers of upregulated and downregulated genes, as well as overlapping genes, across the two donor states. Circles indicate the numbers of genes that are 'T1D enriched' or 'AAB enriched'. P-values presented are the results of hypergeometric CDF tests (one-tailed test for overrepresentation). (Middle) For each cell type, displayed are gene ontology pathways that are shared across T1D and AAB+ cells when compared to Control cells (top) or pathways that are differently enriched in T1D cells vs AAB+ cells (bottom). The top 20 clusters are displayed and a stringent cut-off of $p \leq 1e\text{-}6$ was applied to determine significant gene ontology pathways. (Right) Heatmaps displaying the degree of gene expression changes of genes (rows) that are shared (top) or differential (bottom) across AAB+ and T1D disease states.

Figure S13: Gene and gene ontology pathways that are shared and different across disease states in PP, Endothelial, Stellate, and Hybrid cells

(A-D) (Left) For each cell type, Venn diagrams indicate the numbers of upregulated and downregulated genes, as well as overlapping genes, across the two disease states. Circles indicate the numbers of genes that are 'T1D enriched' or 'AAB enriched'. P-values presented are the results of hypergeometric CDF tests (one-

tailed test for overrepresentation). (Middle) For each cell type, displayed are gene ontology pathways that are shared across T1D and AAB+ cells when compared to Control cells (top) or pathways that are differently enriched in T1D cells vs AAB+ cells (bottom). The top 20 clusters are displayed and a stringent cut-off of $p \leq 1e-6$ was applied to determine significant gene ontology pathways. (Right) Heatmaps displaying the degree of gene expression changes of genes (rows) that are shared (top) or differential (bottom) across AAB+ and T1D disease states.

Figure S14: Gene and gene ontology pathways that are shared and different across disease states in Epsilon-1, Epsilon-2, and Immune cells

(A-C) (Left) For each cell type, Venn diagrams indicate the numbers of upregulated and downregulated genes, as well as overlapping genes, across the two disease states. Circles indicate the numbers of genes that are 'T1D enriched' or 'AAB enriched'. P-values presented are the results of hypergeometric CDF tests (one-tailed test for overrepresentation). (Middle) For each cell type, displayed are gene ontology pathways that are shared across T1D and AAB+ cells when compared to Control cells (top) or pathways that are differently enriched in T1D cells vs AAB+ cells (bottom). The top 20 clusters are displayed and a stringent cut-off of 1e-6 was applied to determine significant gene ontology pathways. (Right) Heatmaps displaying the degree of gene expression changes of genes (rows) that are shared (top) or differential (bottom) across AAB+ and T1D disease states.

(D) GSEA analysis plots of FDR q-value vs Normalized Enrichment Score. For both ductal populations, Ductal-1 and Ductal-2, T1D cells were compared to AAB+ or Control cells to determine differentially enriched gene sets. Demarcated in red and labeled are signatures of interest.

Figure S15: The hybrid cellular state exists across multiple donors and has unique transcriptional features

(A) Cells begin at the start pin symbol, and from there are partitioned based on similarities and differences in gene expression. Dendrogram visualization and clustering of ductal and endocrine cells, without the dominant donor that

contributes to Hybrid cells, AAB+ donor #3. Cells begin at the start pin symbol. Pie charts at the end of the branches display the breakdown of Garnett cellular classification (left), the donor group (middle), or individual donors (right) within that terminal cluster. Circled in black is the Hybrid cluster.

(B) Dendrograms highlight the expression of each canonical gene marker of each major cell type across the dendrogram in fig. S15A. Scale bars represent normalized transcript numbers.

(C) Co-expression analysis of *INS* and the relevant gene indicative of each subpopulation of Hybrid cells on an individual cell level across all Hybrid cells (Ductal-like Hybrids plus Endocrine-like Hybrids plus Acinar-like Hybrids). For each gene, values represent normalized transcript numbers. Percentages in the upper right-hand corner indicate the percentage of co-expressing cells.

(D) Dot plot displaying the $\log_2$(Fold Change (FC)) (circle color) and $-\log_{10}$(FDR) (circle size) of transcription factors (TFs) that are important for endocrine, ductal, and acinar cell fates across four comparisons: Hybrid versus Ductal (Hybrid/(Ductal-1 plus Ductal-2)), Hybrid versus Acinar (Hybrid/Acinar-2), Hybrid versus Beta (Hybrid/(Beta-1 plus Beta-2)), and Hybrid versus Alpha (Hybrid/Alpha). *FOXA2* is a transcription factor important in both endocrine and ductal cell fates.

(E) GSEA analysis plots of FDR q-value vs Normalized Enrichment Score. All Hybrid cells (Ductal-like Hybrids plus Endocrine-like Hybrids plus Acinar-like Hybrids) were compared to all other cell populations combined in the ductal and endocrine dendrogram to determine differentially enriched gene sets. Demarcated in red and labeled is a signature of interest.

Figure S16: Monocle trajectory analysis

(A-C) Monocle trajectory analysis of exocrine-related cell types: Hybrid, Acinar-1, Acinar-2, Ductal-1, and Ductal-2. Dark grey line indicates a possible cellular trajectory. Cells are labeled by their pseudotime (A), all cellular types together (B), or individual cellular types separately (C).

(D-E) Monocle trajectory analysis of endocrine-related cell types: Hybrid, Alpha, Beta-1, Beta-2, and Ductal-1. Dark grey line indicates a possible cellular trajectory. Cells

are labeled by their pseudotime (A), all cellular types together (B), or individual cellular types separately (C).

Figure S17: CyTOF and IMC validation of Hybrid cells

(A) Cells begin at the start pin symbol, and from there are partitioned based on similarities and differences in protein expression. Dendrogram visualization of the immune cell cluster, CD45 positive (+) cells, as determined by the analysis of the flow cytometry by time-of-flight (CyTOF) data.

(B) Dendrogram visualization of the beta cell cluster, C-peptide positive (+) cells, as determined by the analysis of the CyTOF data.

(C) Dendrogram visualization of the alpha cell cluster, Glucagon positive (+) cells, as determined by the analysis of the CyTOF data.

(D) Pie chart displaying immune cells and the relative proportions of each donor group from the CyTOF data. P-value presented is the result of the Chi-squared test.

(E) Box plots displaying immune cell percentage of total cells per individual across donor groups from the CyTOF data.

(F) Dendrogram visualization of the immune cell cluster, CD45 positive (+) cells, as determined by the analysis of the imaging mass cytometry (IMC) data.

(G) Dendrogram visualization of the beta cell cluster, C-peptide positive (+) cells, as determined by the analysis of the IMC data.

(H) Dendrogram visualization of the alpha cell cluster, Glucagon positive (+) cells, as determined by the analysis of the IMC data.

(I) Pie chart displaying immune cells and the relative proportions of each donor group from the IMC data. P-value presented is the result of the Chi-squared test.

(J) Box plots displaying immune cell percentage of total cells per individual across donor groups from the IMC data.

(K-L) Cytokeratin+ C-peptide+ Hybrid cells versus immune percentage of total cells. For each donor group, the mean immune cell percentage of total cells and the mean Hybrid cell percentage of total cells across all individual donors per donor group was computed to generate one data point per donor group for CyTOF (K) or IMC (L) data.

(M) Dendrogram visualization of Hybrid cell (red), Hybrid-Neighboring cell (blue), and Hybrid-Distant cell (grey) enriched clusters as determined by leveraging the spatial architecture provided by IMC data. Dendrograms provided in fig. S17F-H showing immune, beta, and alpha clusters support that Hybrid-neighboring cells are predominately alpha and beta cells, which further supports that Hybrid cells are located within the islets (endocrine compartment) rather than the exocrine compartment of the pancreas.

(N) Boxplots showing the normalized protein expression of different canonical markers in Hybrid-Neighboring cells (blue) versus cells neighboring random cells (grey). The number of random cells evaluated was equal to the number of Hybrid cells. Differential marker expression significance for neighbors in the IMC analysis was determined using permutation tests. For each marker, the distribution of that marker value for each of the designated n neighbors was compared against 100 distributions derived from n random cells across the entire IMC tree. * indicates p-value < 0.01.

(O) Two parameter CyTOF analysis of cytokeratin and C-peptide expression in single cells from Control donor #3 (HPAP034), a donor with a very low percentage of Hybrid cells as determined by TooManyCells analysis of the CyTOF data.

(P) Bar plot displaying the proportion of Hybrid cells from each pancreatic region from the IMC data.

Figure S18: Corroboration of HLA-DR+ Ductal cells

(A-B) Cells begin at the start pin symbol, and from there are partitioned based on similarities and differences in protein expression. Dendrograms highlighting the expression of the MHC class II complex (A) or MHC class II activity (B) across the dendrogram of all cells in Figure 1C. Scale bars represent normalized transcript numbers (mean across all MHC class II complex genes (A) or MHC class II activity genes (B)).

(C-D) Dendrograms highlighting the expression of the MHC class II complex (C) or MHC class II activity (D) across the dendrogram of ductal and endocrine cells in

Figure 1D. Scale bars represent normalized transcript numbers (mean across all MHC class II complex genes (C) or MHC class II activity genes (D)).

(E-F) Dendrograms highlighting the expression of the *HLA-DPB1* (E) or *KRT19* (F) across the dendrogram of ductal and endocrine cells in Figure 1D. Scale bars represent normalized transcript numbers.

(G) Dendrograms highlighting the expression of the immune-related genes across the dendrogram of ductal and endocrine cells in Figure 1D. Scale bars represent normalized transcript numbers.

(H) Dendrograms highlighting the expression of the *BMPR1A* across the dendrogram of ductal and endocrine cells in Figure 1D. Scale bars represent normalized transcript numbers.

Figure S19: CyTOF and IMC validation of HLA-DR+ Ductal cells

(A) Two parameter CyTOF analysis of HLA-DR and cytokeratin protein expression in single cells from T1D donor #3 (HPAP023).

(B) Two parameter CyTOF analysis of HLA-DR and cytokeratin protein expression in single cells from Control donor #3 (HPAP034), a donor with a very low percentage of HLA-DR+ ductal cells as determined by unbiased analysis of CyTOF data with TooManyCells.

(C) Bar plot displaying the proportion of HLA-DR+ cytokeratin+ cells from each pancreatic region from the IMC data.

(D-E) HLA-DR+ cytokeratin+ cells versus immune percentage of total cells. For each donor group, the mean immune cell percentage of total cells and the mean HLA-DR+ ductal cell percentage of total cells across all individual donors per donor group was computed to generate one data point per donor group for CyTOF (D) or IMC (E) data.

(F) Dendrogram visualization of the clusters of HLA-DR+ cytokeratin+ cells (red), cells neighboring HLA-DR+ cytokeratin+ (blue), and cells distant from HLA-DR+ cytokeratin+ cells (grey) as determined by leveraging the spatial architecture provided by IMC data. Cells begin at the start pin symbol, and from there are partitioned based on similarities and differences in protein expression.

(G) Boxplots showing the normalized protein expression of different canonical markers in cells neighboring HLA-DR+ cytokeratin+ cells (blue) versus cells neighboring random cells (grey). The number of random cells evaluated was equal to the number of HLA-DR+ cytokeratin+ cells. Differential marker expression significance for neighbors in the IMC analysis was determined using permutation tests. For each marker, the distribution of that marker value for each of the designated n neighbors was compared against 100 distributions derived from n random cells across the entire IMC tree. * indicates p-value < 0.01.

Figure S20: MHC Class II positive ductal cells are present in islets of all donor types, and a high percentage of MHC Class II positive ductal cells in T1D donors appears to be correlated with an inflammatory environment

(Left) Imaging mass cytometry (IMC) in a region of interest (ROI) in pancreatic tissue from three representative individual donors for each donor group type (T1D, AAB+, and Control). HLA-DR is a general marker of MHC Class II (HLA-DR) expression, CD99 is a general islet marker, KRT (pan-keratin) is a ductal cell marker, and CD45 (PTPRC) is a general immune cell marker. Notably, HLA-DR+ ductal cells were primarily located in large ductal structures (outlined in yellow). The images presented here are publicly available at https://www.pancreatlas.org/datasets/508. (Right) HLA typing performed by next generation sequencing. Comprehensive clinical information about each donor is provided in PANC-DB: https://hpap.pmacs.upenn.edu/. Highlighted in yellow are the particular HLA alleles contributing to the susceptible or protective genotypes, which are abbreviated for each donor on the left side of the figure as follows. The four susceptible genotypes assessed were (1) HLA-DRB1*03:01-HLA-DQA1*05:01-HLA-DQB1*02:01 (abbreviated as 'DR3', referring to the haplotype bearing the DRB1*03 allele); (2) HLA-DRB1*04:01/02/04/05/08-HLA-DQA1*03:01-HLA-DQB1*03:02/04 (or HLA-DQB1*02) (abbreviated as 'DR4', referring to the haplotype bearing the DRB1*04 allele); (3) HLA-A*24:02; and (4) HLA-B*39:06. The two protective genotypes assessed were (1) HLA-DRB1*15:01-HLA-DQB1*06:02 and (2) HLA-DRB1*07:01-HLA-DQB1*03:03 (Howson et al., 2012; Howson et al., 2009; Inshaw et al., 2020; Nejentsev et al., 2007; Noble and Valdes,

2011; Valdes et al., 2012). Notably, HLA-DR+ ductal cells were found across all HLA genotypes, including both susceptible and protective genotypes.

Tables

Table 1: Donor clinical information

Table 2: Differential gene list between Beta-1 cells (positive fold change (FC)) and Beta-2 cells (negative FC)

Table 3: Differential gene list between Epsilon-1 cells (positive fold change (FC)) and Epsilon-2 cells (negative FC)

Table 4: Differential gene list between Ductal-1 cells (positive fold change (FC)) and Ductal-2 cells (negative FC)

Table 5: Differential gene list between Ductal-2 cells (positive fold change (FC)) and Acinar-2 cells (negative FC)

Table 6: Differential gene list between AAB+ Beta-1 cells (positive fold change (FC)) and Control Beta-1 cells (negative FC)

Table 7: Differential gene list between T1D Beta-1 cells (positive fold change (FC)) and Control Beta-1 cells (negative FC)

Table 8: Differential gene list between AAB+ Beta-2 cells (positive fold change (FC)) and Control Beta-2 cells (negative FC)

Table 9: Differential gene list between T1D Beta-2 cells (positive fold change (FC)) and Control Beta-2 cells (negative FC)

Table 10: Differential gene list between AAB+ Ductal-1 cells (positive fold change (FC)) and Control Ductal-1 cells (negative FC)

Table 11: Differential gene list between T1D Ductal-1 cells (positive fold change (FC)) and Control Ductal-1 cells (negative FC)

Table 12: Differential gene list between AAB+ Ductal-2 cells (positive fold change (FC)) and Control Ductal-1 cells (negative FC)

Table 13: Differential gene list between T1D Ductal-2 cells (positive fold change (FC)) and Control Ductal-1 cells (negative FC)

Table 14: Differential gene list between AAB+ Acinar-1 cells (positive fold change (FC)) and Control Acinar-1 cells (negative FC)

Table 15: Differential gene list between T1D Acinar-1 cells (positive fold change (FC)) and Control Acinar-1 cells (negative FC)

Table 16: Differential gene list between AAB+ Acinar-2 cells (positive fold change (FC)) and Control Acinar-2 cells (negative FC)

Table 17: Differential gene list between T1D Acinar-2 cells (positive fold change (FC)) and Control Acinar-2 cells (negative FC)

Table 18: Differential gene list between AAB+ Alpha cells (positive fold change (FC)) and Control Alpha cells (negative FC)

Table 19: Differential gene list between T1D Alpha cells (positive fold change (FC)) and Control Alpha cells (negative FC)

Table 20: Differential gene list between AAB+ Delta cells (positive fold change (FC)) and Control Delta cells (negative FC)

Table 21: Differential gene list between T1D Delta cells (positive fold change (FC)) and Control Delta cells (negative FC)

Table 22: Differential gene list between AAB+ PP cells (positive fold change (FC)) and Control PP cells (negative FC)

Table 23: Differential gene list between T1D PP cells (positive fold change (FC)) and Control PP cells (negative FC)

Table 24: Differential gene list between AAB+ Endothelial cells (positive fold change (FC)) and Control Endothelial cells (negative FC)

Table 25: Differential gene list between T1D Endothelial cells (positive fold change (FC)) and Control Endothelial cells (negative FC)

Table 26: Differential gene list between AAB+ Stellate cells (positive fold change (FC)) and Control Stellate cells (negative FC)

Table 27: Differential gene list between T1D Stellate cells (positive fold change (FC)) and Control Stellate cells (negative FC)

Table 28: Differential gene list between AAB+ Hybrid cells (positive fold change (FC)) and Control Hybrid cells (negative FC)

Table 29: Differential gene list between T1D Hybrid cells (positive fold change (FC)) and Control Hybrid cells (negative FC)

Table 30: Differential gene list between AAB+ Epsilon-1 cells (positive fold change (FC)) and Control Epsilon-1 cells (negative FC)

Table 31: Differential gene list between T1D Epsilon-1 cells (positive fold change (FC)) and Control Epsilon-1 cells (negative FC)

Table 32: Differential gene list between AAB+ Epsilon-2 cells (positive fold change (FC)) and Control Epsilon-2 cells (negative FC)

Table 33: Differential gene list between T1D Epsilon-2 cells (positive fold change (FC)) and Control Epsilon-2 cells (negative FC)

Table 34: Differential gene list between AAB+ Immune cells (positive fold change (FC)) and Control Immune cells (negative FC)

Table 35: Differential gene list between T1D Immune cells (positive fold change (FC)) and Control Immune cells (negative FC)

Table 36: Differential gene list between T1D Beta-1 cells (positive fold change (FC)) and AAB+ Beta-1 cells (negative FC)

Table 37: Differential gene list between T1D Beta-2 cells (positive fold change (FC)) and AAB+ Beta-2 cells (negative FC)

Table 38: Differential gene list between T1D Ductal-1 cells (positive fold change (FC)) and AAB+ Ductal-1 cells (negative FC)

Table 39: Differential gene list between T1D Ductal-2 cells (positive fold change (FC)) and AAB+ Ductal-2 cells (negative FC)

Table 40: Differential gene list between T1D Acinar-1 cells (positive fold change (FC)) and AAB+ Acinar-1 cells (negative FC)

Table 41: Differential gene list between T1D Acinar-2 cells (positive fold change (FC)) and AAB+ Acinar-2 cells (negative FC)

Table 42: Differential gene list between T1D Alpha cells (positive fold change (FC)) and AAB+ Alpha cells (negative FC)

Table 43: Differential gene list between T1D Delta cells (positive fold change (FC)) and AAB+ Delta cells (negative FC)

Table 44: Differential gene list between T1D PP cells (positive fold change (FC)) and AAB+ PP cells (negative FC)

Table 45: Differential gene list between T1D Endothelial cells (positive fold change (FC)) and AAB+ Endothelial cells (negative FC)

Table 46: Differential gene list between T1D Stellate cells (positive fold change (FC)) and AAB+ Stellate cells (negative FC)

Table 47: Differential gene list between T1D Hybrid cells (positive fold change (FC)) and AAB+ Hybrid cells (negative FC)

Table 48: Differential gene list between T1D Epsilon-1 cells (positive fold change (FC)) and AAB+ Epsilon-1 cells (negative FC)

Table 49: Differential gene list between T1D Epsilon-2 cells (positive fold change (FC)) and AAB+ Epsilon-2 cells (negative FC)

Table 50: Differential gene list between T1D Immune cells (positive fold change (FC)) and AAB+ Immune cells (negative FC)

Table 51: Differential gene list between Endocrine-like Hybrid cells (positive fold change (FC)) and beta cells (both Beta-1 and Beta-2 cells; negative FC)

Table 52: Differential gene list between Endocrine-like Hybrid cells (positive fold change (FC)) and Alpha cells (negative FC)

Table 53: Differential gene list between Ductal-like Hybrid cells (positive fold change (FC)) and beta cells (both Beta-1 and Beta-2 cells; negative FC)

Table 54: Differential gene list between Ductal-like Hybrid cells (positive fold change (FC)) and ductal cells (both Ductal-1 and Ductal-2 cells; negative FC)

Table 55: Differential gene list between Acinar-like Hybrid cells (positive fold change (FC)) and beta cells (both Beta-1 and Beta-2 cells; negative FC)

Table 56: Differential gene list between Acinar-like Hybrid cells (positive fold change (FC)) and Acinar-2 cells (negative FC)

Table 57: Differential gene list between all Hybrid cells (positive fold change (FC)) and all remaining cells in the ductal and endocrine tree (negative FC)

Table 58: CyTOF Panel

Table 59: IMC Panel

Table 60: Garnett cell type marker file (*37*)

References

1. J. Boldison, F. S. Wong, Immune and Pancreatic beta Cell Interactions in Type 1 Diabetes. *Trends Endocrinol Metab* **27**, 856-867 (2016).
2. E. R. Unanue, X. Wan, The Immunoreactive Platform of the Pancreatic Islets Influences the Development of Autoreactivity. *Diabetes* **68**, 1544-1551 (2019).
3. M. A. Atkinson, Pancreatic biopsies in type 1 diabetes: revisiting the myth of Pandora's box. *Diabetologia* **57**, 656-659 (2014).
4. M. A. Atkinson, G. S. Eisenbarth, Type 1 diabetes: new perspectives on disease pathogenesis and treatment. *Lancet* **358**, 221-229 (2001).
5. Y. J. Wang *et al.*, Multiplexed In Situ Imaging Mass Cytometry Analysis of the Human Endocrine Pancreas and Immune System in Type 1 Diabetes. *Cell Metab* **29**, 769-783 e764 (2019).
6. J. C. Barrett *et al.*, Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat Genet* **41**, 703-707 (2009).
7. M. Campbell-Thompson *et al.*, Network for Pancreatic Organ Donors with Diabetes (nPOD): developing a tissue biobank for type 1 diabetes. *Diabetes Metab Res Rev* **28**, 608-617 (2012).
8. K. H. Kaestner, A. C. Powers, A. Naji, H. Consortium, M. A. Atkinson, 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 (2019).
9. H. A. Pliner, J. Shendure, C. Trapnell, Supervised classification enables rapid annotation of cell atlases. *Nat Methods* **16**, 983-986 (2019).
10. G. W. Schwartz *et al.*, TooManyCells identifies and visualizes relationships of single-cell clades. *Nat Methods*, (2020).
11. S. Bonner-Weir *et al.*, Transdifferentiation of pancreatic ductal cells to endocrine beta-cells. *Biochem Soc Trans* **36**, 353-356 (2008).
12. L. Bouwens, G. Kloppel, Islet cell neogenesis in the pancreas. *Virchows Arch* **427**, 553-560 (1996).
13. H. Heimberg *et al.*, Adult human pancreatic duct and islet cells exhibit similarities in expression and differences in phosphorylation and complex formation of the homeodomain protein Ipf-1. *Diabetes* **49**, 571-579 (2000).
14. H. S. Kim, M. K. Lee, beta-Cell regeneration through the transdifferentiation of pancreatic cells: Pancreatic progenitor cells in the pancreas. *J Diabetes Investig* **7**, 286-296 (2016).
15. W. J. Rutter, The development of the endocrine and exocrine pancreas. *Monogr Pathol* **21**, 30-38 (1980).
16. F. Zhong, Y. Jiang, Endogenous Pancreatic beta Cell Regeneration: A Potential Strategy for the Recovery of beta Cell Deficiency in Diabetes. *Front Endocrinol (Lausanne)* **10**, 101 (2019).
17. T. S. Heng, M. W. Painter, C. Immunological Genome Project, The Immunological Genome Project: networks of gene expression in immune cells. *Nat Immunol* **9**, 1091-1094 (2008).
18. M. Baron *et al.*, A Single-Cell Transcriptomic Map of the Human and Mouse Pancreas Reveals Inter- and Intra-cell Population Structure. *Cell Syst* **3**, 346-360 e344 (2016).

19. A. Butler, P. Hoffman, P. Smibert, E. Papalexi, R. Satija, Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* **36**, 411-420 (2018).
20. M. J. Muraro *et al.*, A Single-Cell Transcriptome Atlas of the Human Pancreas. *Cell Syst* **3**, 385-394 e383 (2016).
21. A. Segerstolpe *et al.*, Single-Cell Transcriptome Profiling of Human Pancreatic Islets in Health and Type 2 Diabetes. *Cell Metab* **24**, 593-607 (2016).
22. Y. J. Wang *et al.*, Single-Cell Transcriptomics of the Human Endocrine Pancreas. *Diabetes* **65**, 3028-3038 (2016).
23. U. Christen *et al.*, A dual role for TNF-alpha in type 1 diabetes: islet-specific expression abrogates the ongoing autoimmune process when induced late but not early during pathogenesis. *J Immunol* **166**, 7023-7032 (2001).
24. Q. Chen, J. Kang, C. Fu, The independence of and associations among apoptosis, autophagy, and necrosis. *Signal Transduct Target Ther* **3**, 18 (2018).
25. A. Lombardi, E. Tsomos, S. S. Hammerstad, Y. Tomer, Interferon alpha: The key trigger of type 1 diabetes. *J Autoimmun* **94**, 7-15 (2018).
26. B. S. Marro *et al.*, Progression of type 1 diabetes from the prediabetic stage is controlled by interferon-alpha signaling. *Proc Natl Acad Sci U S A* **114**, 3708-3713 (2017).
27. K. C. Osum *et al.*, Interferon-gamma drives programmed death-ligand 1 expression on islet beta cells to limit T cell function during autoimmune diabetes. *Sci Rep* **8**, 8295 (2018).
28. D. Avrahami *et al.*, Single-cell transcriptomics of human islet ontogeny defines the molecular basis of beta-cell dedifferentiation in T2D. *Mol Metab* **42**, 101057 (2020).
29. M. Roig, A. Macaya, F. Munell, A. Capdevila, Acute neurologic dysfunction associated with destructive lesions of the basal ganglia: a benign form of infantile bilateral striatal necrosis. *J Pediatr* **117**, 578-581 (1990).
30. X. Qiu *et al.*, Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods* **14**, 979-982 (2017).
31. M. Fanjul *et al.*, Evidence for epithelial-mesenchymal transition in adult human pancreatic exocrine cells. *J Histochem Cytochem* **58**, 807-823 (2010).
32. I. A. Valdez *et al.*, Proinflammatory Cytokines Induce Endocrine Differentiation in Pancreatic Ductal Cells via STAT3-Dependent NGN3 Activation. *Cell Rep* **15**, 460-470 (2016).
33. E. W. Newell, N. Sigal, S. C. Bendall, G. P. Nolan, M. M. Davis, Cytometry by time-of-flight shows combinatorial cytokine expression and virus-specific cell niches within a continuum of CD8+ T cell phenotypes. *Immunity* **36**, 142-152 (2012).
34. Y. J. Wang *et al.*, Single-Cell Mass Cytometry Analysis of the Human Endocrine Pancreas. *Cell Metab* **24**, 616-626 (2016).
35. J. A. Noble, A. M. Valdes, Genetics of the HLA region in the prediction of type 1 diabetes. *Curr Diab Rep* **11**, 533-542 (2011).
36. S. Nejentsev *et al.*, Localization of type 1 diabetes susceptibility to the MHC class I genes HLA-B and HLA-A. *Nature* **450**, 887-892 (2007).
37. M. Fasolino, G. Vahedi, Web portal for scRNA-seq in our study. *http://faryabi05.med.upenn.edu:8050*.
38. A. K. K. Teo *et al.*, Single-cell analyses of human islet cells reveal de-differentiation signatures. *Cell Death Discov* **4**, 14 (2018).

39. R. R. De Krijger *et al.*, The midgestational human fetal pancreas contains cells coexpressing islet hormones. *Dev Biol* **153**, 368-375 (1992).
40. M. Polak, L. Bouchareb-Banaei, R. Scharfmann, P. Czernichow, Early pattern of differentiation in the human pancreas. *Diabetes* **49**, 225-232 (2000).
41. M. K. Chiang, D. A. Melton, Single-cell transcript analysis of pancreas development. *Dev Cell* **4**, 383-393 (2003).
42. H. Katsuta *et al.*, Single pancreatic beta cells co-express multiple islet hormone genes in mice. *Diabetologia* **53**, 128-138 (2010).
43. M. J. Riedel *et al.*, Immunohistochemical characterisation of cells co-producing insulin and glucagon in the developing human pancreas. *Diabetologia* **55**, 372-381 (2012).
44. M. M. F. Qadir *et al.*, Single-cell resolution analysis of the human pancreatic ductal progenitor cell niche. *Proc Natl Acad Sci U S A* **117**, 10876-10887 (2020).
45. T. Rodriguez-Calvo, O. Ekwall, N. Amirian, J. Zapardiel-Gonzalo, M. G. von Herrath, Increased immune cell infiltration of the exocrine pancreas: a possible contribution to the pathogenesis of type 1 diabetes. *Diabetes* **63**, 3880-3890 (2014).
46. V. Steimle, C. A. Siegrist, A. Mottet, B. Lisowska-Grospierre, B. Mach, Regulation of MHC class II expression by interferon-gamma mediated by the transactivator gene CIITA. *Science* **265**, 106-109 (1994).
47. J. E. Wosen, D. Mukhopadhyay, C. Macaubas, E. D. Mellins, Epithelial MHC Class II Expression and Its Role in Antigen Presentation in the Gastrointestinal and Respiratory Tracts. *Front Immunol* **9**, 2144 (2018).
48. M. A. Russell *et al.*, HLA Class II Antigen Processing and Presentation Pathway Components Demonstrated by Transcriptome and Protein Analyses of Islet beta-Cells From Donors With Type 1 Diabetes. *Diabetes* **68**, 988-1001 (2019).
49. C. Dorrell *et al.*, Human islets contain four distinct subtypes of beta cells. *Nat Commun* **7**, 11756 (2016).
50. T. Stuart *et al.*, Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902 e1821 (2019).
51. C. S. McGinnis, L. M. Murrow, Z. J. Gartner, DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst* **8**, 329-337 e324 (2019).
52. Y. Zhou *et al.*, Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* **10**, 1523 (2019).
53. A. Subramanian *et al.*, Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550 (2005).
54. A. Liberzon *et al.*, Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739-1740 (2011).
55. R. Dassaye, S. Naidoo, M. E. Cerf, Transcription factor regulation of pancreatic organogenesis, differentiation and maturation. *Islets* **8**, 13-34 (2016).
56. X. Qiu *et al.*, Single-cell mRNA quantification and differential analysis with Census. *Nat Methods* **14**, 309-315 (2017).
57. C. Trapnell *et al.*, The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* **32**, 381-386 (2014).