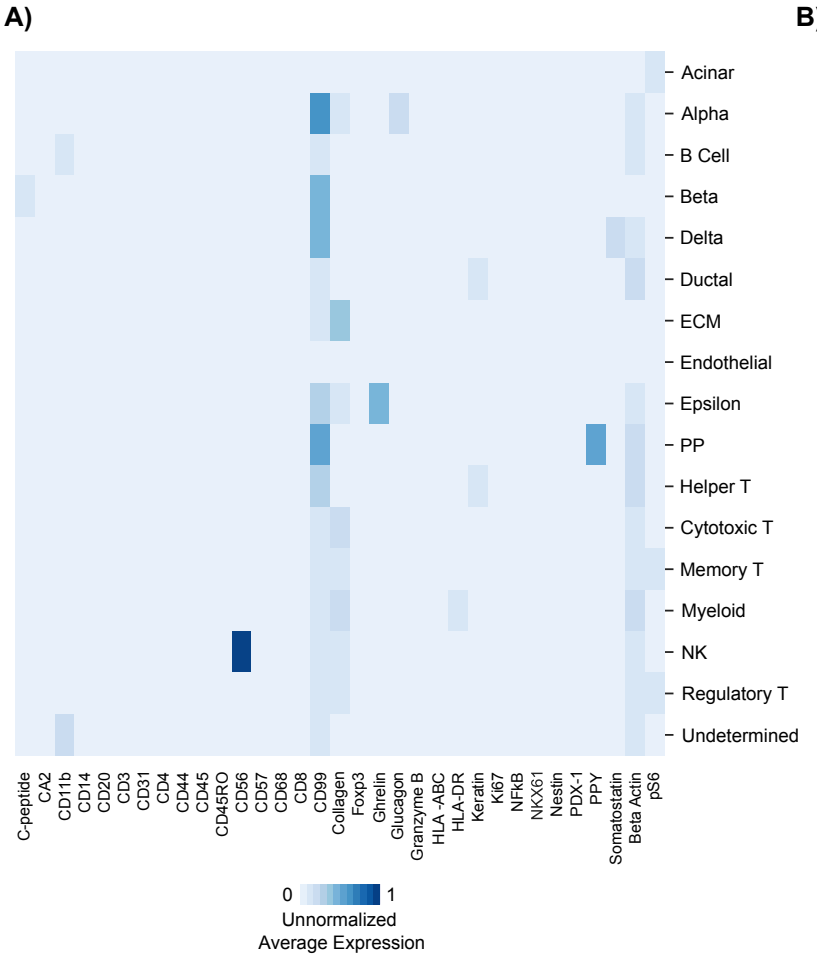
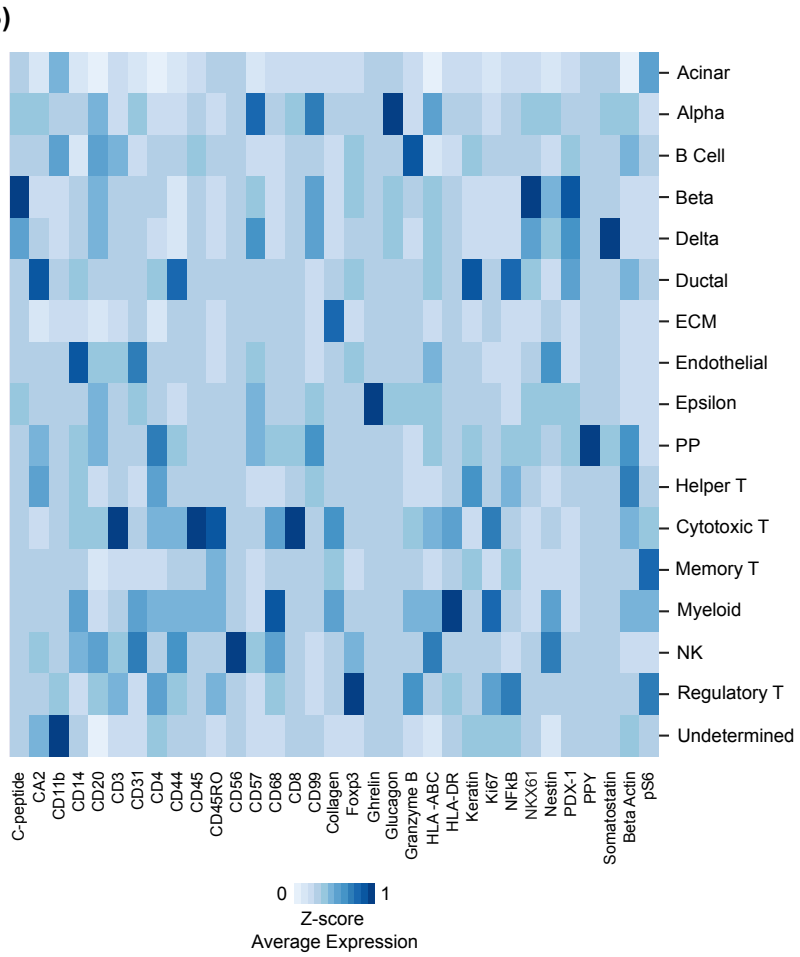
Supplementary Materials

Supplementary Figures

Figure S1





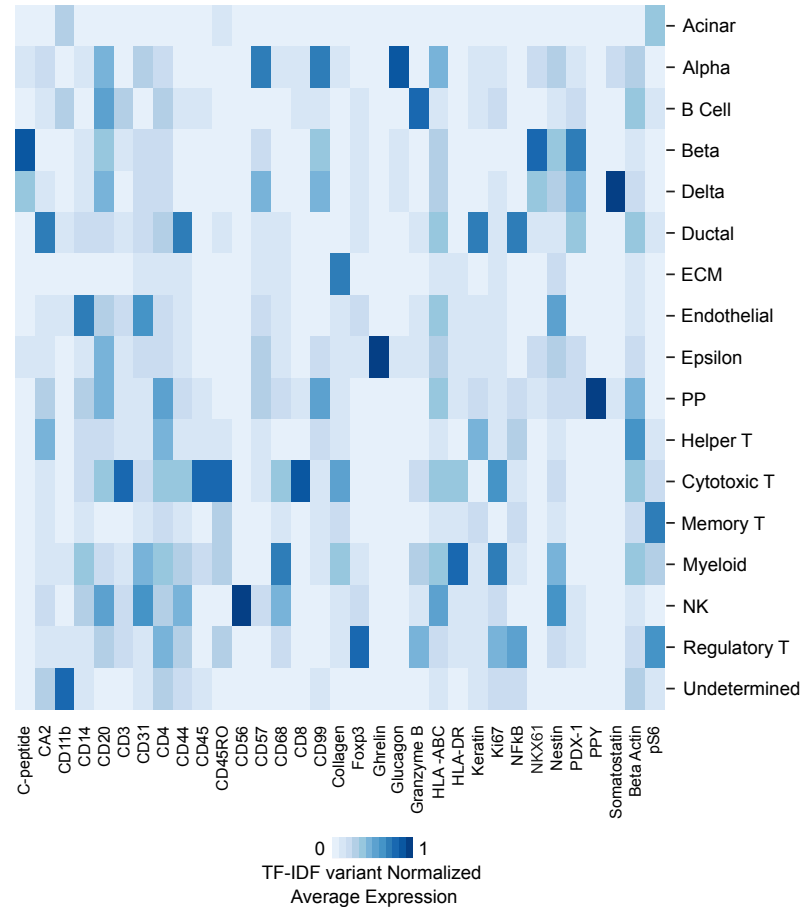
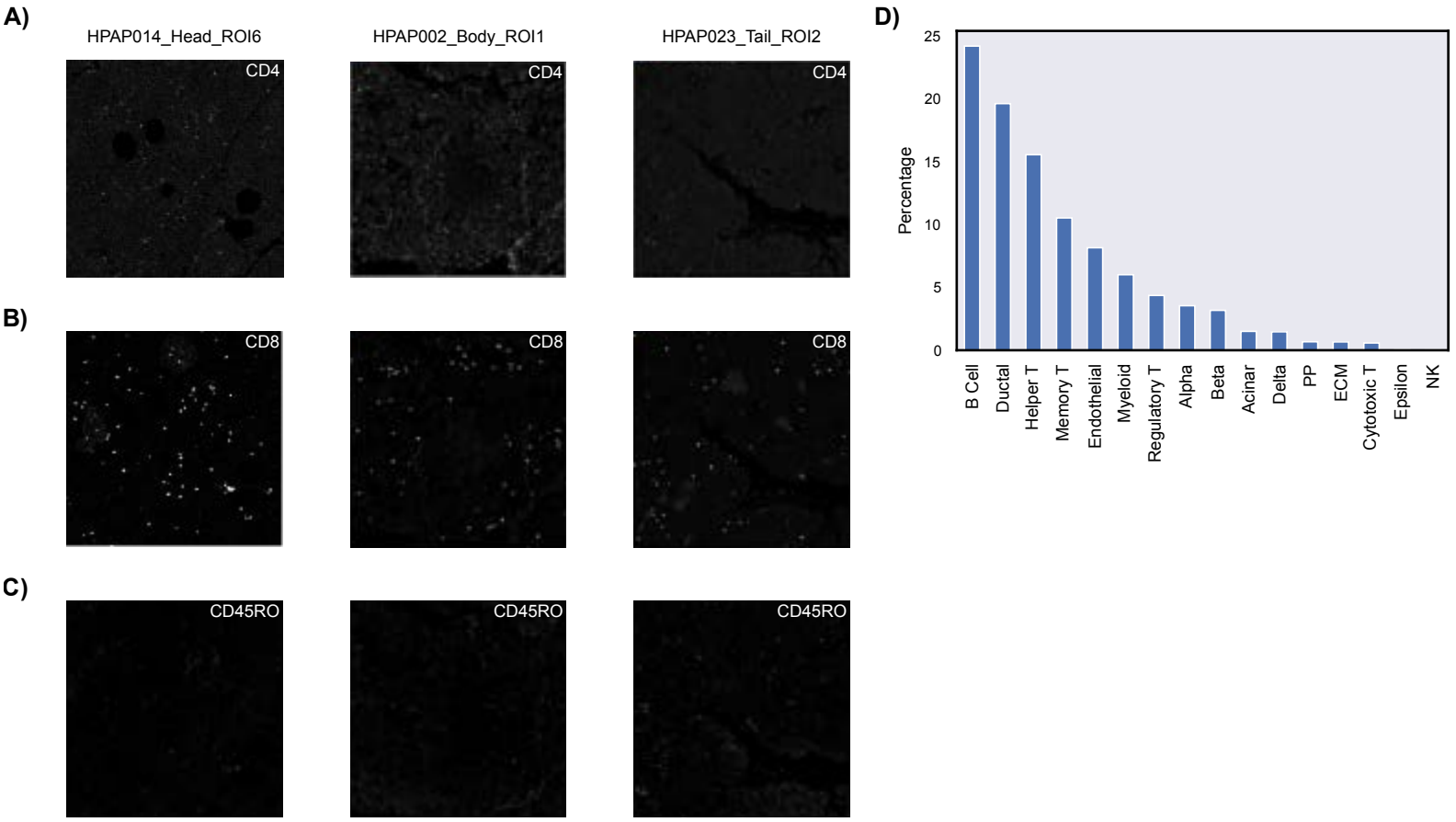


Figure S1: Comparison of various IMC data normalization methods. Heatmaps showing average unnormalized **(A)**, protein-wise z-score normalized **(B)**, and a variant of TF-IDF normalized **(C)** expression levels of 33 IMC-measured proteins across AnnoSpat-annotated cell types. Heatmaps comparison indicates the benefit of a variant of TF-IDF for normalization in visualizing continuous protein expression readouts. Note: TF-IDF variant normalization is only used for data visualization and not cell-type annotation.

Figure S2



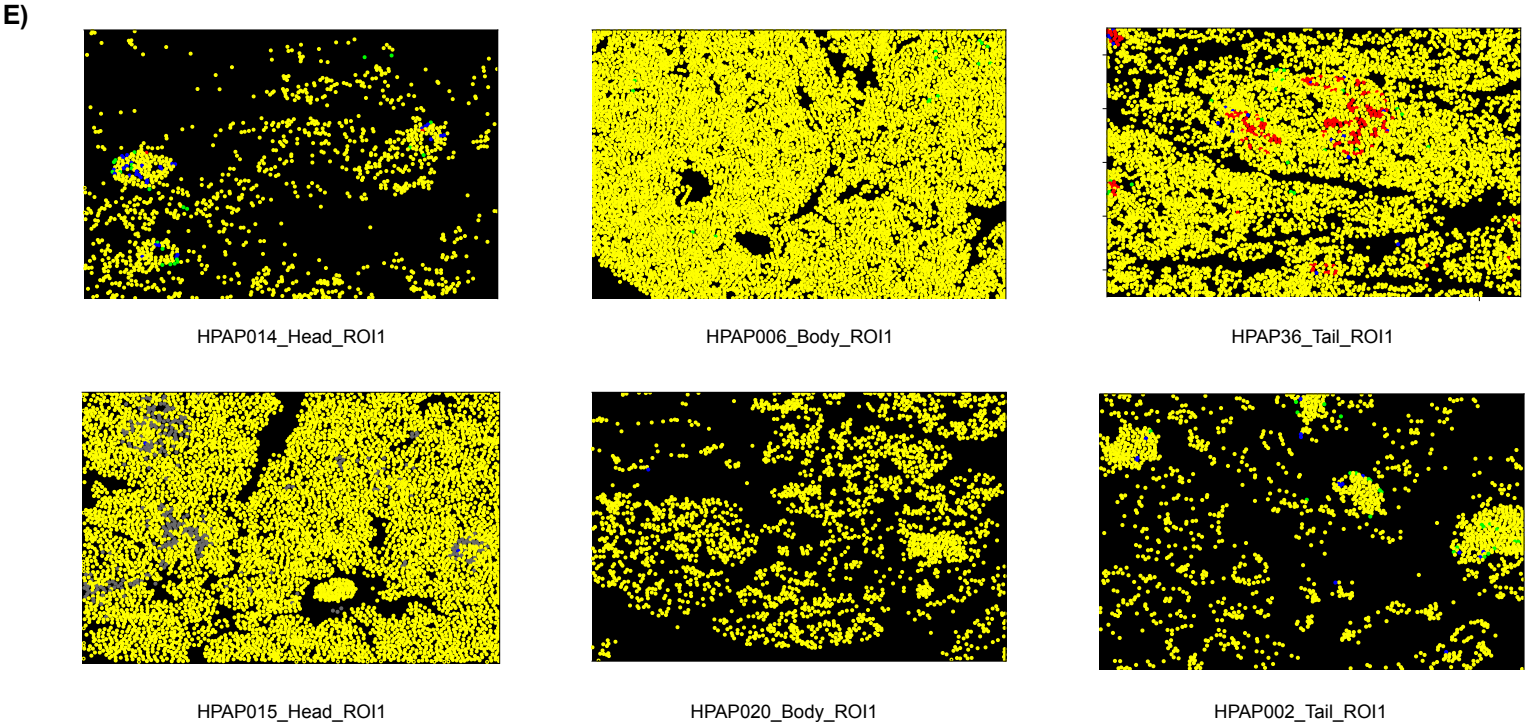


Figure S2: Comparison of AnnoSpat and AUCell cell-type annotation. **(A-C)** Randomly selected IMC images of ROIs from pancreas head, body, tail comparing CD4 (A), CD8 (B) and CD45RO (C) staining quality showing higher quality of CD8 compared to CD4 and CD45RO staining. **(D)** Bar plots showing percentage of AnnoSpat-annotated cell types that AUCell failed to annotate. **(E)** Yellow pseudo-color marking AnnoSpat-annotated cells that AUCell failed to annotate on randomly selected IMC images. Other cell types are colored as before (e.g. refer to Figure 3C).

Figure S3

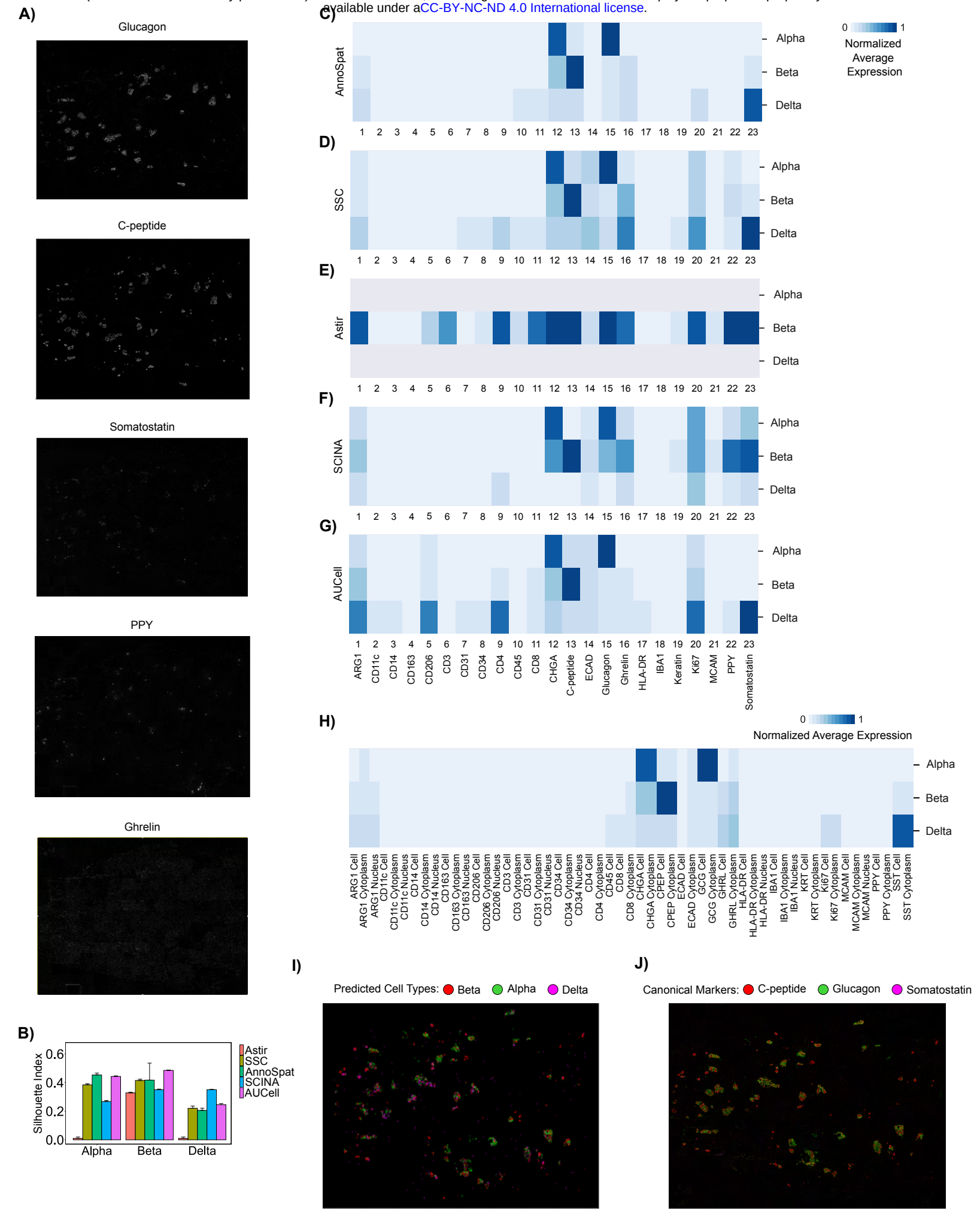
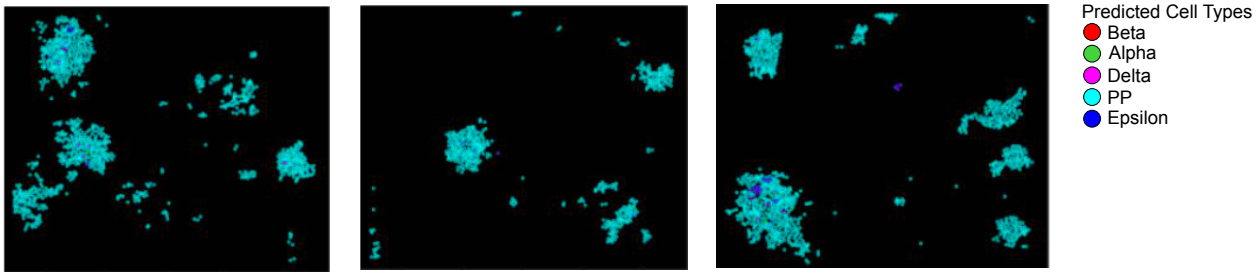


Figure S3: Comparative analysis of AnnoSpat cell-type annotation from CODEX data. **(A)** From top to bottom: raw images of glucagon, c-peptide, somatostatin, PPY, and ghrelin showing non-specificity of PP and epsilon markers in CODEX experiments. **(B)** Bar plots with error bars showing average and standard deviation Silhouette Index (SI) for cells annotated as alpha, beta, and delta by AnnoSpat, semi-supervised clustering (SSC), Astir, SCINA, and AUCell from non-diabetic pancreas CODEX data ($m = 10$ sets of $n = 50,000$ cells randomly selected from $n = 220,155$ measured cells). **(C-G)** Heatmaps showing marker proteins' normalized average expression levels for cells labeled as alpha, beta, and delta by AnnoSpat, SSC, Astir, SCINA, and AUCell from non-diabetic pancreas CODEX data ($n = 220,155$ measured cells). **(H)** Heatmap showing marker proteins' normalized average expression levels separately for the nucleus and cytoplasm of the cells annotated as alpha, beta, and delta by AnnoSpat based on protein intensity in the entire cells from non-diabetic pancreas CODEX data ($n = 220,155$ measured cells). **(I, J)** CODEX image is overlaid by AnnoSpat predicted cell types (I) or alpha (glucagon), beta (c-peptide), and delta (somatostatin) marker protein channels (J).

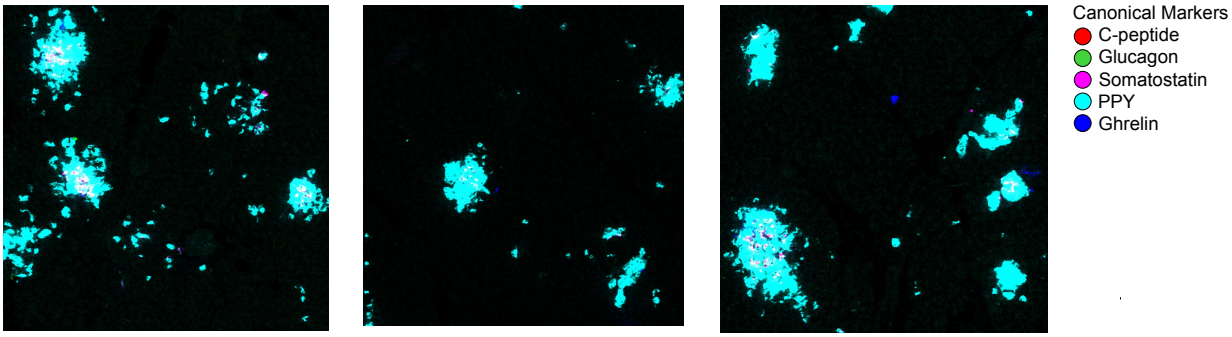
Figure S4

A)



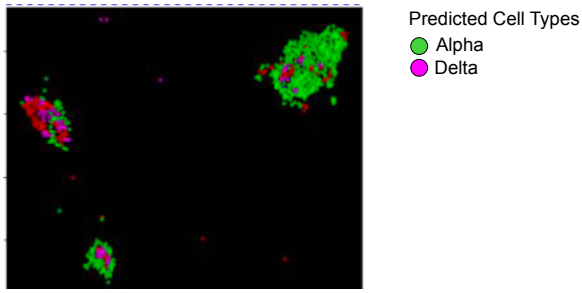
HPAP015_Head_ROI1 HPAP015_Head_ROI2 HPAP015_Head_ROI4

B)



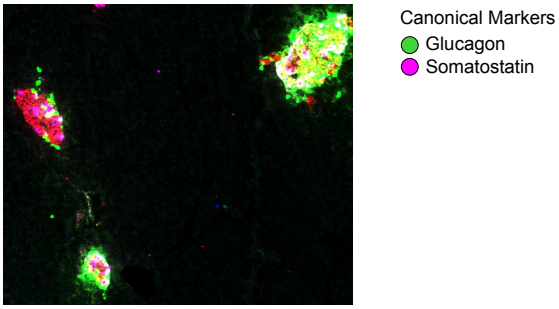
HPAP015_Head_ROI1 HPAP015_Head_ROI2 HPAP015_Head_ROI4

C)



HPAP006_Body_ROI1

D)

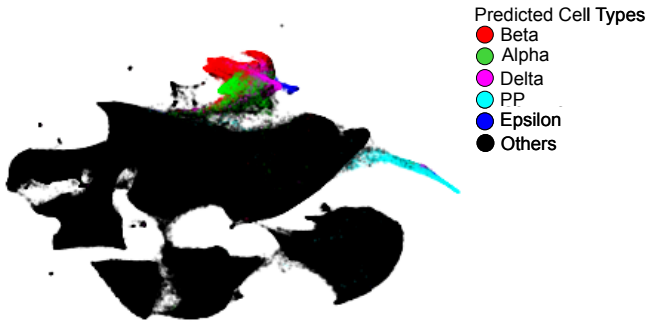


HPAP006_Body_ROI1

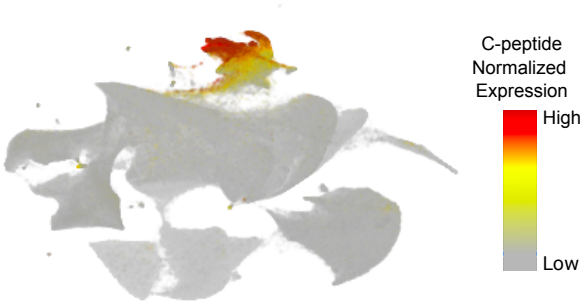
Figure S4: Comparison of AnnoSpat and expert annotation of pancreatic endocrine cell types. **(A-D)** Representative IMC images from donors with discordant expert and AnnoSpat cell-type annotation in Figures 3A and 3B are overlaid with AnnoSpat-predicted cell types (A and C) or endocrine canonical marker protein channels (B and D). C-peptide, glucagon, somatostatin, PPY, and ghrelin marking beta, alpha, delta, PP, and epsilon cells, respectively.

Figure S5

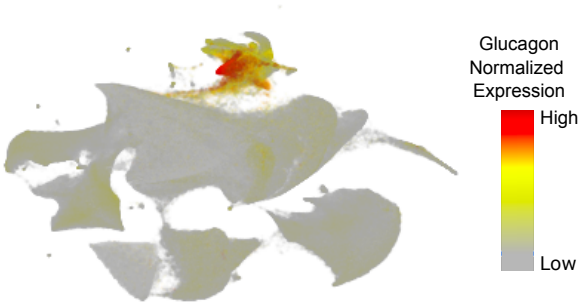
A)



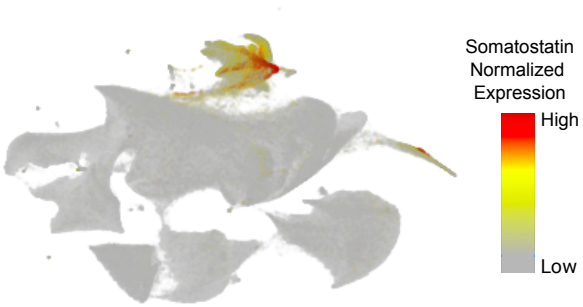
B)



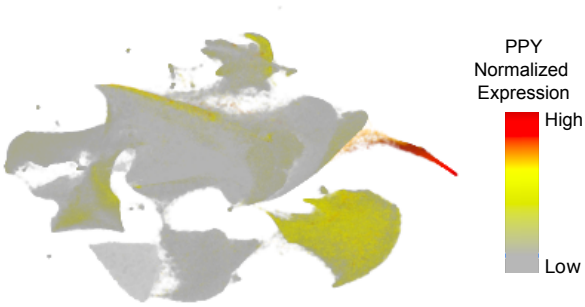
C)



D)



E)



F)

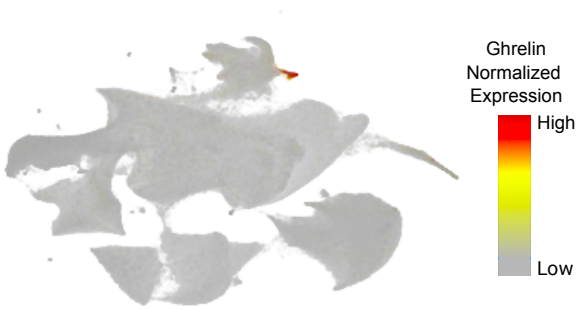
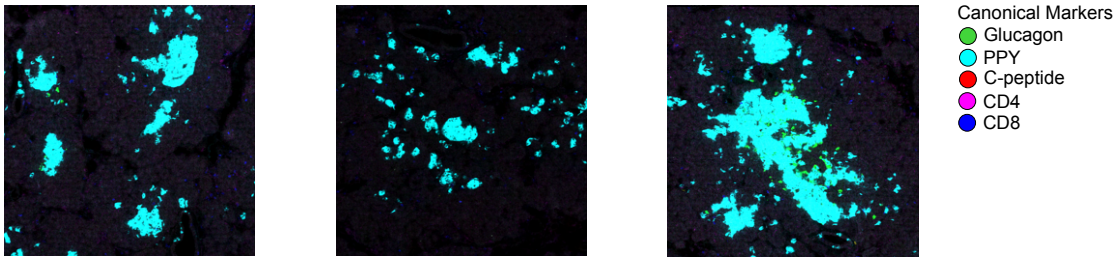


Figure S5: Comparison of protein marker expression levels and AnnoSpat annotations across pancreatic endocrine cell types. **(A-F)** UMAP plots overlaid by AnnoSpat-predicted cell types (A), and expression levels of c-peptide (B), glucagon (C), somatostatin (D), PPY (E), and ghrelin (F) in $n = 65,643$ cells across $m = 141$ slides of 16 pancreas donors.

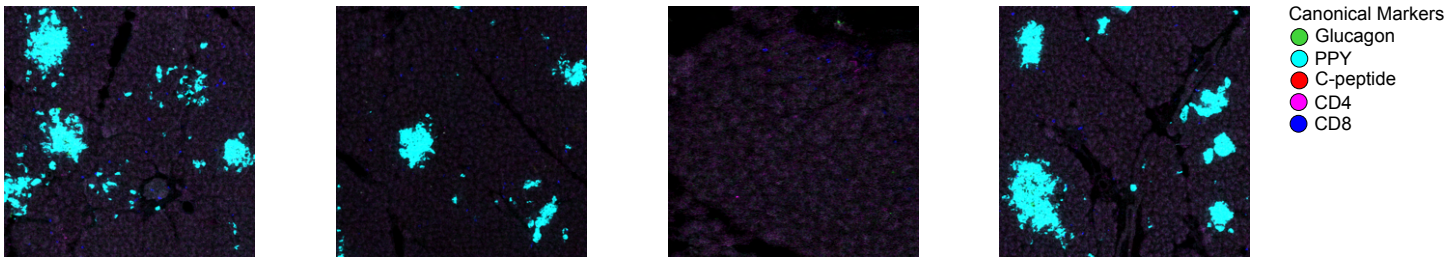
Figure S6

A)



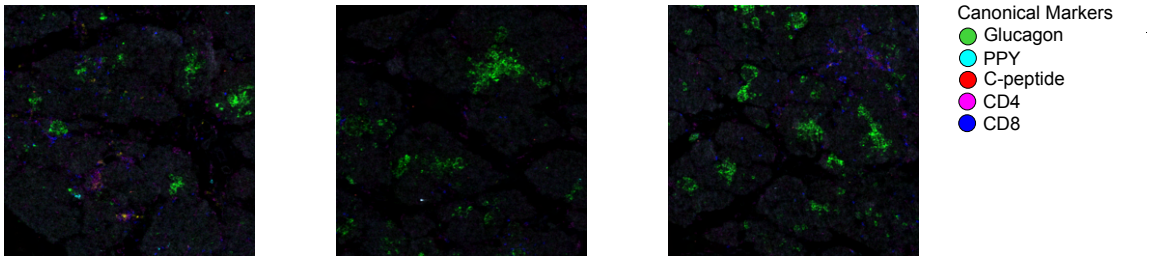
HPAP002_Head_ROI1 HPAP002_Head_ROI2 HPAP002_Head_ROI4

B)



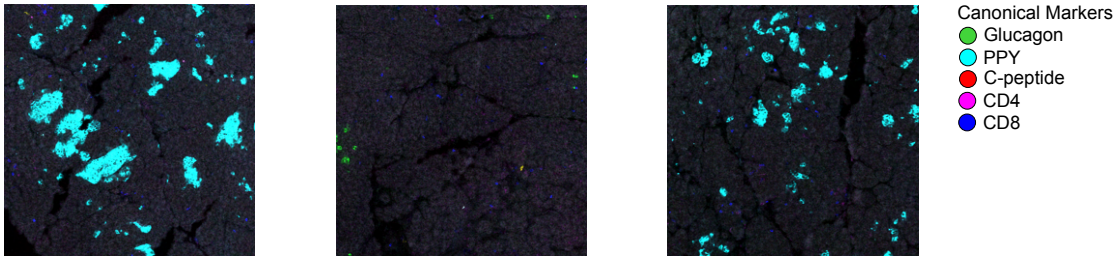
HPAP015_Head_ROI1 HPAP015_Head_ROI2 HPAP015_Head_ROI3 HPAP015_Head_ROI4

C)



HPAP020_Head_ROI2 HPAP020_Head_ROI3 HPAP020_Head_ROI4

D)



HPAP023_Head_ROI1 HPAP023_Head_ROI2 HPAP023_Head_ROI3

Figure S6: PP cell count increases in the pancreas head during T1D progression. **(A-D)** IMC images from pancreatic head ROIs overlaid with expression levels of canonical protein markers of alpha (glucagon), beta (c-peptide), PP (PPY), helper T (CD4), and cytotoxic T (CD8) cells.

Figure S7

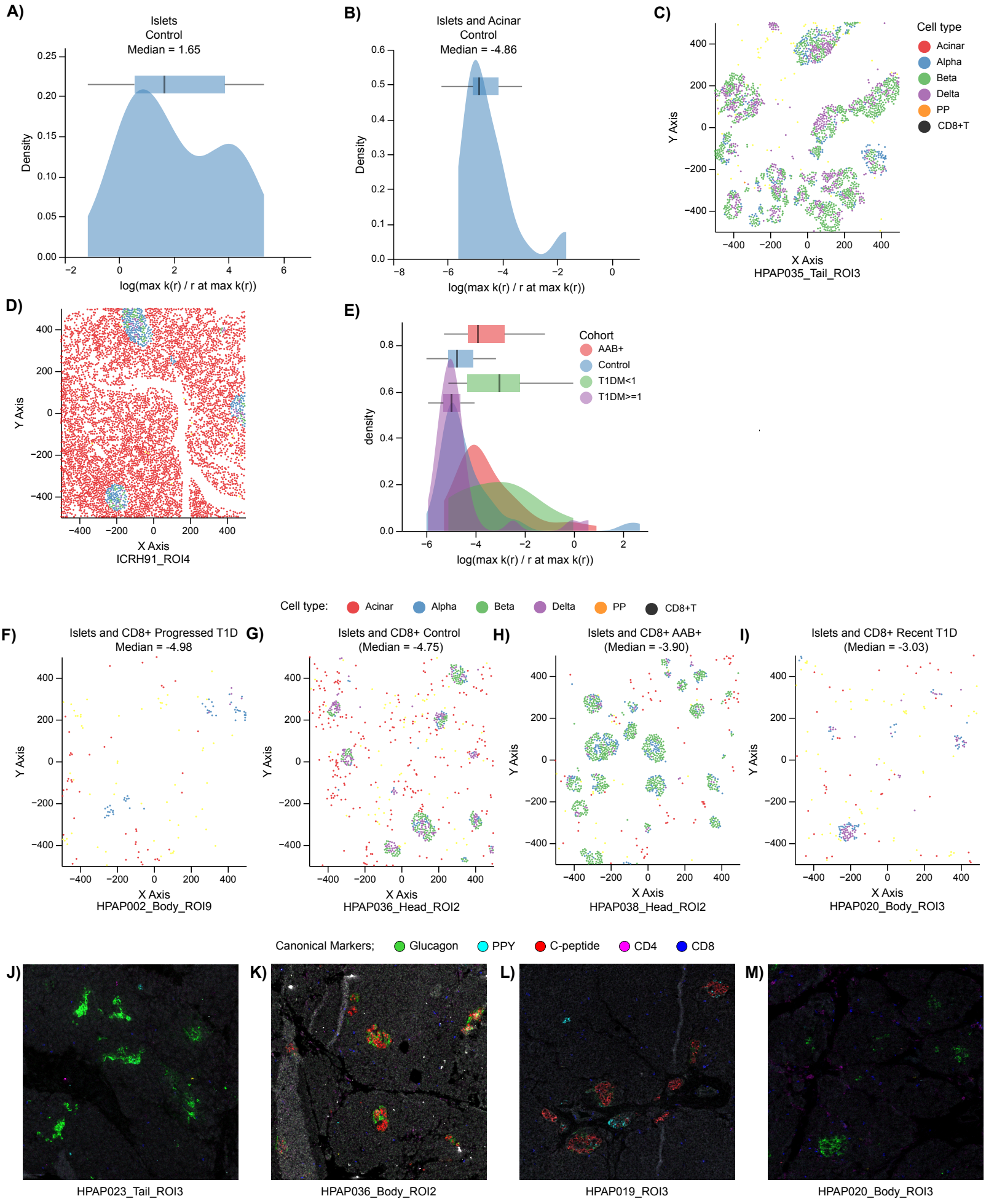


Figure S7: The extent of CD8$^+$ T cell infiltration in islets changes during T1D progression. Analysis corresponding to Figure 6 but with an alternative summarization measure of mark cross-correlation function, which takes into account correlation value as well as distance (r): $\omega(r) = \log \frac{\max_r k_{mm}(r)}{\arg\max_r k_{mm}(r)}$. **(A, B)** Distributions with box-and-whisker plot overlays of $\omega(r)$ across all ROIs for endocrine cells with respect to themselves (A) or with respect to acinar cells (B). **(C, D)** Scatter plots showing location of cells within ROIs at the median of (A) and (B) distributions are plotted in (C) and (D), respectively. Cells are colored by AnnoSpat-predicted cell types. Endocrine cells tend to aggregate around themselves more often than with acinar cells. **(E)** The distributions with box-and-whisker plot overlays of $\omega(r)$ across control, AAb$^+$, recent T1D, and prolonged T1D. AAb$^+$ and recent T1D tend to have greater aggregation of islets with CD8$^+$ T cells than control and prolonged T1D cohorts. **(F-I)** Scatter plots showing location of cells within ROIs at the median of each cohort in (E). From lowest to highest aggregation: prolonged T1D (F), control (G), AAb$^+$ (H), and recent T1D (I). Cells are colored by AnnoSpat-predicted cell types. **(J-M)** IMC images from pancreatic ROIs overlaid with expression levels of canonical protein markers of alpha (glucagon), beta (c-peptide), PP (PPY), helper T (CD4), and cytotoxic T (CD8) cells confirming changes in the CD8$^+$ T cell infiltration in islets during T1D progression. Images in (J) to (M) correspond to scatter plots in Figures 6F to 6I, respectively. Box-and-whisker plots: center line, median; box limits, upper (75th) and lower (25th) percentiles; whiskers, 1.5 · interquartile range; points, outliers.

Supplementary Tables

Table S1

IMC antibody panel.

Table S2

HPAP pancreas donor information.

Table S3

AnnoSpat's marker file input for annotating the listed cell types from IMC antibodies.

Table S4

SI and DB scores for labeling endocrine cells from IMC samples of T1D donors' pancreata. Numbers in parentheses: standard deviation. NA: no cell was annotated.

Table S5

SI and DB scores for labeling endocrine cells from IMC samples of non-diabetic (control) donors' pancreata. Numbers in parentheses: standard deviation. NA: no cell was annotated.

Table S6

SI and DB scores for labeling endocrine cells from IMC samples of combined non-diabetic and T1D (Combined) donors' pancreata. Numbers in parentheses: standard deviation. NA: no cell was annotated.

Table S7

Fraction of endocrine cells labeled by each algorithm from IMC samples of T1D, non-diabetic (control), and combined T1D and control (Combined) donors' pancreata.

Table S8

Mean and standard deviation of run-time for listed algorithms to annotate cells from IMC samples of T1D, non-diabetic (control), and combined T1D and control (Combined) donors' pancreata. Each algorithm was run three times on data sets of $n = 374,397$, $n = 795,604$, and $n = 1,170,001$ cells from IMC samples of T1D, control, and combined T1D and control donors

using a machine with Ubuntu 20.04, 1.05TB Memory, Intel Xeon Gold CPU 6230R @ 2.1GHz, 2 physical processors 52 cores, and 104 threads.

Table S9

CODEX antibody panel.

Table S10

AnnoSpat's marker file input for annotating the listed cell types from CODEX antibodies.

Table S11

SI and DB scores for labeling alpha, beta, and delta cells from a non-diabetic donor pancreas CODEX. Numbers in parentheses: standard deviation. NA: no cell was annotated.

Table S12

Fraction of expert-annotated endocrine cell types in different regions of pancreata from donors studied in[16].