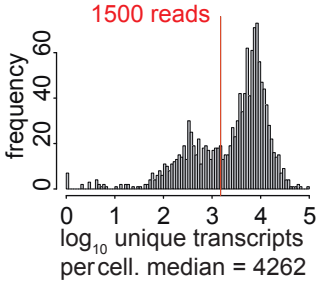
Cell Systems, Volume *3*

Supplemental Information

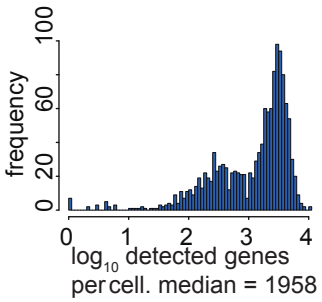
A Single-Cell Transcriptome Atlas of the Human Pancreas

Mauro J. Muraro, Gitanjali Dharmadhikari, Dominic Grün, Nathalie Groen, Tim Dielen, Erik Jansen, Leon van Gurp, Marten A. Engelse, Francoise Carlotti, Eelco J.P. de Koning, and Alexander van Oudenaarden

A



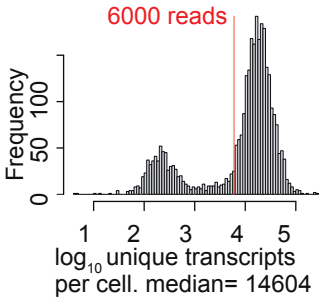
B



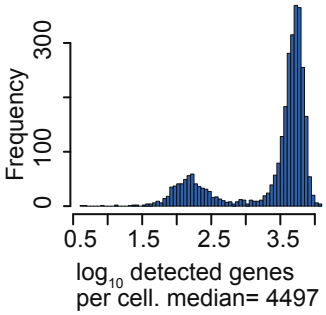
C

	alpha	beta	delta	PP
alpha		71	26	8
beta	6		30	12
delta	3	7		5
PP	2	1	1	

D



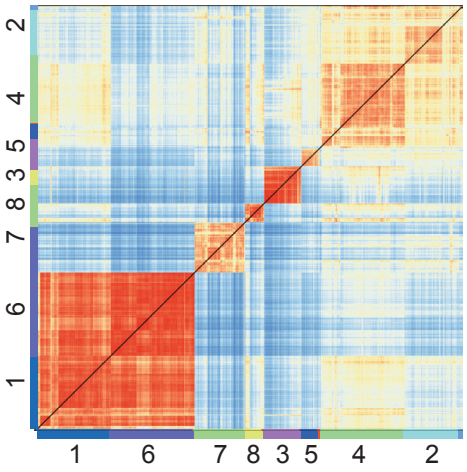
E



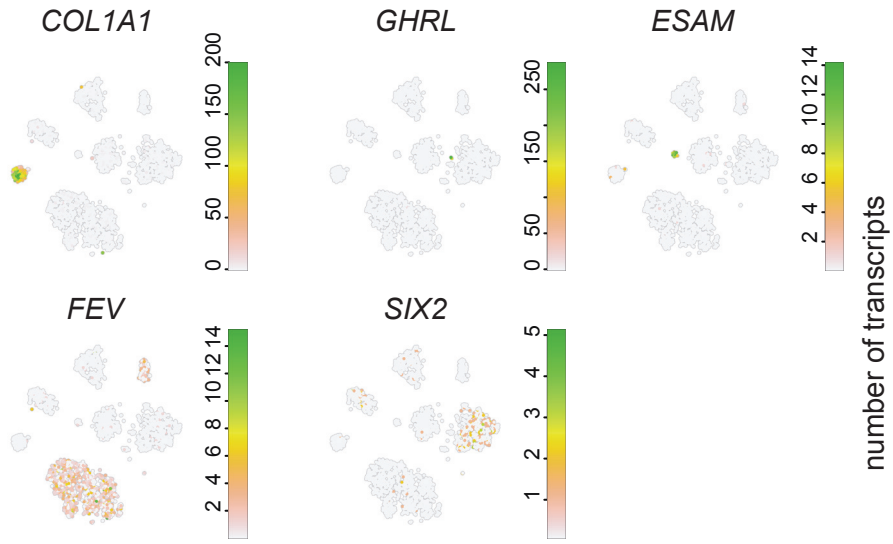
F

	alpha	beta	delta	PP
alpha		1376	533	259
beta	233		428	373
delta	71	74		160
PP	47	66	22	

G



H



I

epsilon cells		duct cells		acinar cells		mesenchymal cells		endothelial cells	
genes	TF	genes	TF	genes	TF	genes	TF	genes	TF
GHRL	*VTN*	*SPP1*	*ONECUT2*	*PNLIP*	*GATA4*	*COL1A1*	*WNT5A*	*FLT1*	*SOX18*
ANXA13	*EBF1*	*CFTR*	*LITAF*	*REG1B*	*MECOM*	*COL1A2*	*SNAI2*	*KDR*	*RGCC*
PHGR1	*BMP7*	*AQP1*	*SOX4*	*PRSS1*	*NR5A2*	*COL3A1*	*NOTCH3*	*CD93*	*SMAD6*
ACSL1	*CDKN2A*	*ALDH1A3*	*DAB2*	*ALB*	*ZFP36L1*	*COL6A3*	*FBN1*	*ESAM*	*ERG*
FRZB	*PROX1*	*KRT19*	*CREB5*	*PRSS3P2*	*CSDA*	*FN1*	*HEYL*	*SOX18*	*PRDM1*
SPTSSB	*ARX*	*CRP*	*HLA-DQB1*	*CPA2*	*CEBPD*	*SFRP2*	*PRRX1*	*PECAM1*	*TCF4*
ASGR1	*ZKSCAN1*	*DEFB1*	*WWTR1*	*CTRB2*	*CREB3L1*	*COL5A1*	*UACA*	*ESM1*	*NOTCH4*
HEPACAM2		*CEACAM6*	*PPARGC1A*	*CEL*	*XBP1*	*SPARC*	*AEBP1*	*PASK*	*SNAI1*
VTN		*MMP7*	*PKHD1*	*PLA2G1B*	*LGR4*	*COL15A1*	*TBX3*	*SLCO2A1*	*NKX2-3*
SERPINA1		*TSPAN8*	*NFIB*	*CELA3A*	*NUPR1*	*SERPINE1*	*FOXF2*	*PLVAP*	*ETS1*

Figure S1

Figure S1. SORT-Seq allows for deep sequencing of human pancreas cells, Related to Figure 1.

(A) Histogram of the total detected transcripts per cell for cells of the five donors processed by manual CEL-Seq. On the X-axis are the $\log_{10}$ detected unique transcripts per cell. Y-axis is the frequency. The minimum number of unique transcripts per cell used as cutoff for downsampling and analysis is indicated in red (1500).

(B) Histogram of genes detected per cell for cells of the first five donors processed by manual CEL-Seq. X-axis are the genes detected per cell. Y-axis is the frequency.

(C) Table indicating the differentially expressed genes (blue) and transcription factors (green) when comparing across the different endocrine cell types from data prepared by manual CEL-Seq.

(D) Histogram of the total detected transcripts per cell for cells of the four donors (SORT-Seq) used in this study. On the X-axis are the $\log_{10}$ detected unique transcripts per cell. Y-axis is the frequency. . The minimum number of unique transcripts per cell used as cutoff for downsampling and analysis is indicated in red (6000).

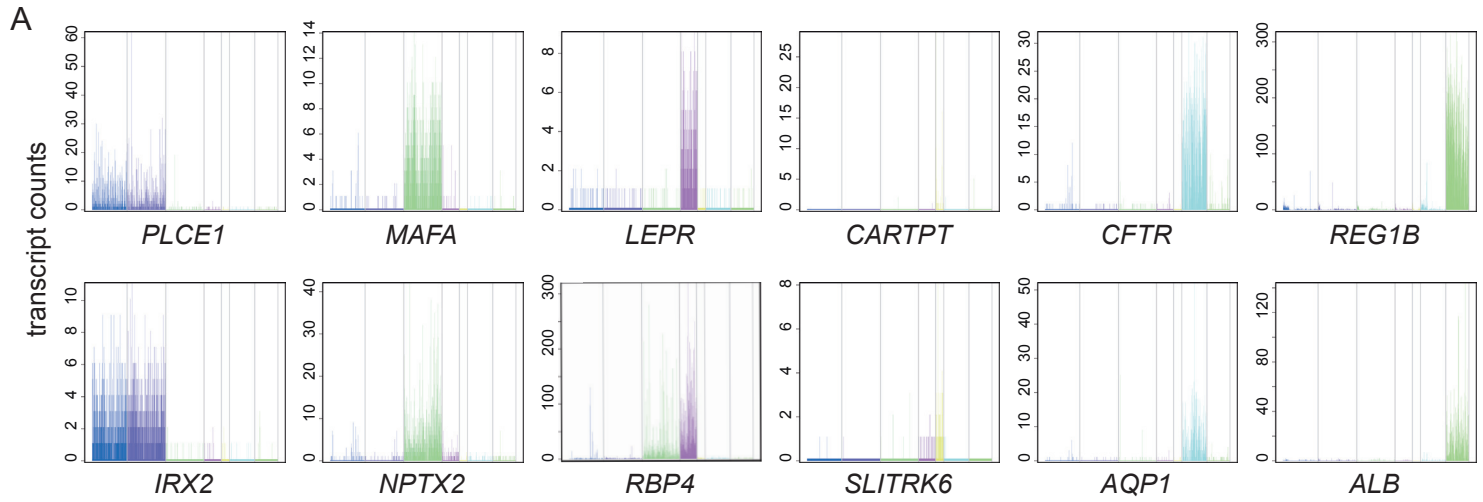
(E) Histogram of genes detected per cell for cells of the four donors (SORT-Seq) used in this study. X-axis are the genes detected per cell. Y-axis is the frequency. On average 1891 genes were detected.

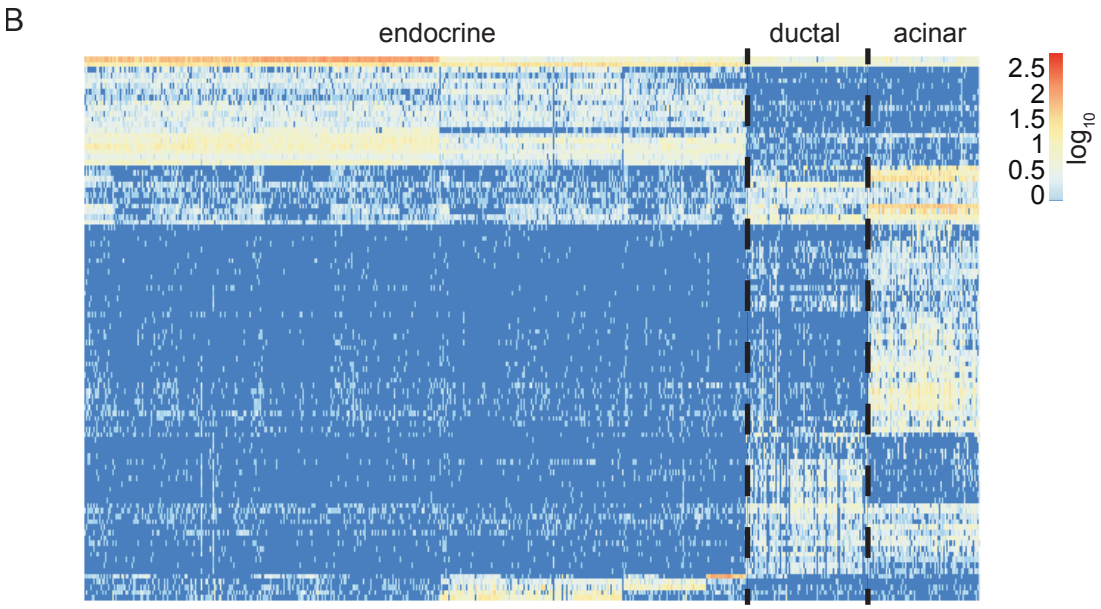
(F) Table indicating the differentially expressed genes (blue) and transcription factors (green) when comparing across the different endocrine cell types from data prepared by SORT-Seq.

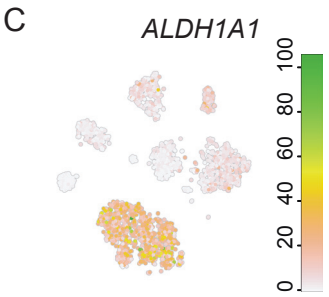
(G) Heat map showing distances between cellular transcriptomes obtained by sequencing. Clustering was performed by StemID (Grün et al, 2016). Distances are calculated as 1 – Pearson correlation and used as input for k-medoid clustering. Each line represents a cell and cells are grouped by cluster. Black lines indicate clusters, as do color bars and numbers on the axes, which match the colors and numbers in Figure 1B.

(H) t-SNE maps highlighting cell type-specific expression of pancreatic marker genes. Transcript counts are given in linear scale. Green indicates high expression.

(I) Tables denoting the top 10 differentially expressed genes and transcription factors (TF) when comparing one of the pancreatic cell types to all other cells in the dataset ($P<10^{-6}$). Continuation of Figure 1E.


A
transcript counts
PLCE1
MAFA
LEPR
CARTPT
CFTR
REG1B
IRX2
NPTX2
RBP4
SLITRK6
AQP1
ALB


B
endocrine
ductal
acinar
log10


C
ALDH1A1

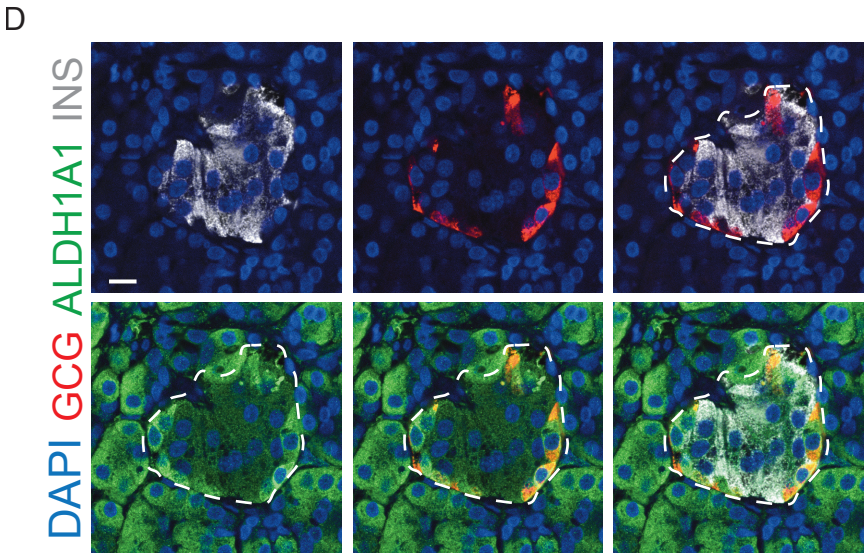

D
DAPI GCG ALDH1A1 INS

Figure S2

Figure S2. Cluster-restricted gene expression patterns and identification of new cell-type specific genes, Related to Figure 2.

(A) Expression of second (top) and third (bottom) most differentially expressed genes in each of six of the main pancreatic cell types. Down-sampled gene expression values are plotted on the Y-axis. Each bar represents a cell and cells are grouped by cluster with a specific color in the following order: alpha, beta, delta, PP, duct and acinar cells. If the most differentially expressed gene was also a canonical marker gene, the third and fourth most differentially expressed genes are shown.

(B) Heat map of the top 100 differentially expressed genes between endocrine and exocrine cell types. Rows are genes, columns are cells. Dashed lines indicate separation between acinar, ductal and endocrine cells. Log_2 expression of transcript counts for genes is plotted where red is high expression. Genes are grouped based on hierarchical clustering.

(C) t-SNE map highlighting the expression *ALDH1A1*. Transcript counts are given in linear scale. Green indicates high expression.

(D) Immunohistochemistry for ALDH1A1 (green) glucagon (red) and insulin (gray) with counterstaining for DAPI (blue) on human pancreatic tissue sections. Co-staining for INS and GCG identifies an Islet of Langerhans (marked by white dashed line). Co-staining of ALDH1A1 with GCG and INS shows overlap in the alpha cells, but not the beta cells inside the islet of Langerhans. Surrounding acinar cells express ALDH1A1 as well. Scale bar is 25 μM.

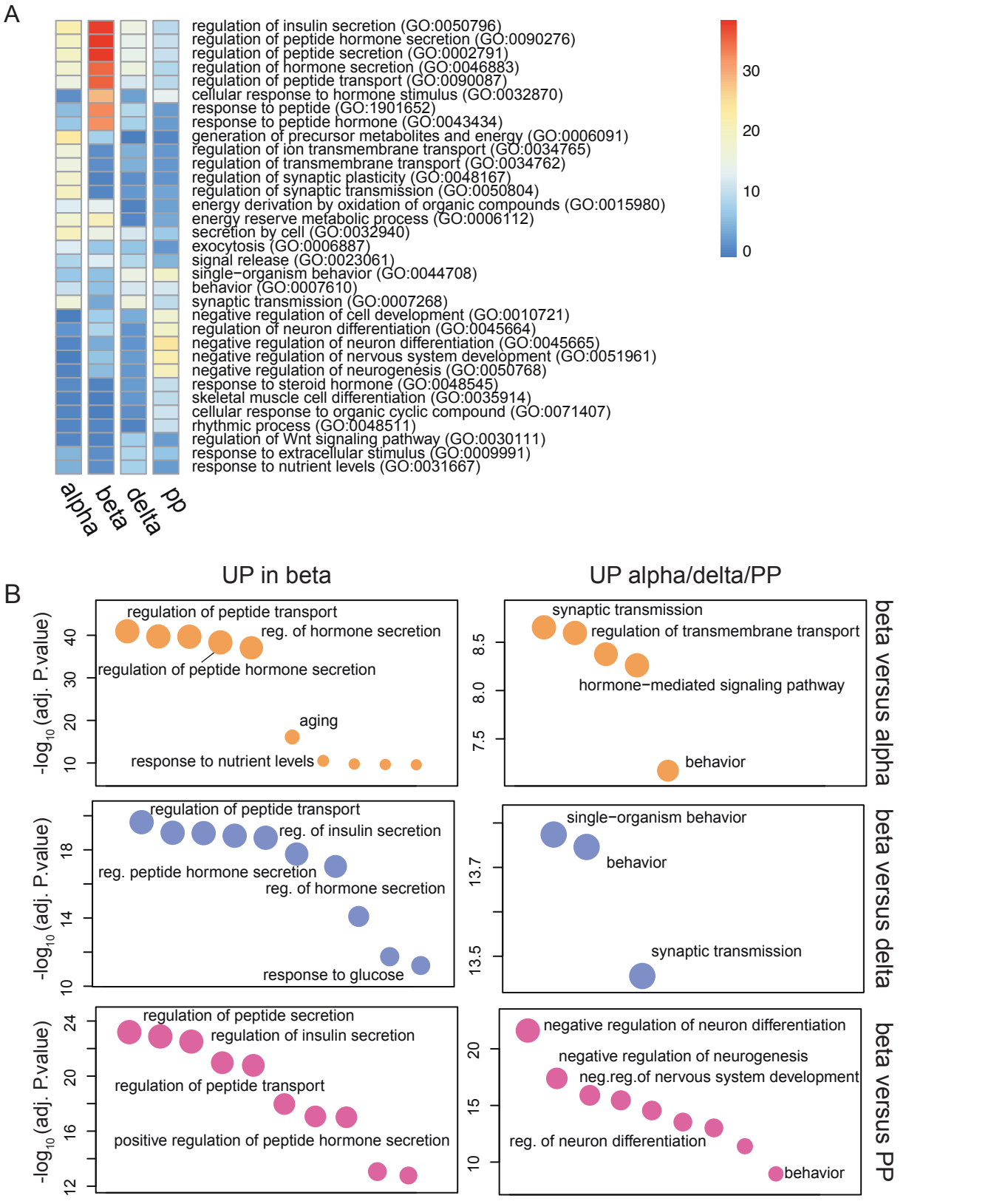

A
regulation of insulin secretion (GO:0050796)
regulation of peptide hormone secretion (GO:0090276)
regulation of peptide secretion (GO:0002791)
regulation of hormone secretion (GO:0046883)
regulation of peptide transport (GO:0090087)
cellular response to hormone stimulus (GO:0032870)
response to peptide (GO:1901652)
response to peptide hormone (GO:0043434)
generation of precursor metabolites and energy (GO:0006091)
regulation of ion transmembrane transport (GO:0034765)
regulation of transmembrane transport (GO:0034762)
regulation of synaptic plasticity (GO:0048167)
regulation of synaptic transmission (GO:0050804)
energy derivation by oxidation of organic compounds (GO:0015980)
energy reserve metabolic process (GO:0006112)
secretion by cell (GO:0032940)
exocytosis (GO:0006887)
signal release (GO:0023061)
single−organism behavior (GO:0044708)
behavior (GO:0007610)
synaptic transmission (GO:0007268)
negative regulation of cell development (GO:0010721)
regulation of neuron differentiation (GO:0045664)
negative regulation of neuron differentiation (GO:0045665)
negative regulation of nervous system development (GO:0051961)
negative regulation of neurogenesis (GO:0050768)
response to steroid hormone (GO:0048545)
skeletal muscle cell differentiation (GO:0035914)
cellular response to organic cyclic compound (GO:0071407)
rhythmic process (GO:0048511)
regulation of Wnt signaling pathway (GO:0030111)
response to extracellular stimulus (GO:0009991)
response to nutrient levels (GO:0031667)
alpha
beta
delta
pp
30
20
10
0
B
UP in beta
UP alpha/delta/PP
-log10 (adj. P.value)
regulation of peptide transport
reg. of hormone secretion
regulation of peptide hormone secretion
aging
response to nutrient levels
synaptic transmission
regulation of transmembrane transport
hormone−mediated signaling pathway
behavior
beta versus alpha
reg. of insulin secretion
reg. peptide hormone secretion
response to glucose
single−organism behavior
synaptic transmission
beta versus delta
regulation of peptide secretion
regulation of insulin secretion
positive regulation of peptide hormone secretion
negative regulation of neuron differentiation
negative regulation of neurogenesis
neg.reg.of nervous system development
reg. of neuron differentiation
beta versus PP
C
ITPR1
SLC6A4
GHSR

Figure S3

Figure S3. GO-term analysis reveals cell-type specific gene expression patterns relevant to endocrine biology and glucose metabolism.

(A) Heatmap showing the combined list of top 15 enriched GO terms for genes differentially expressed in endocrine cell types. Color indicates 1/p-value value so that red indicates a high score.

(B) Plot showing top 10 enriched GO terms for genes differentially expressed between beta cells compared to the three other endocrine cell types. The left column shows GO terms for genes with higher expression in beta cells, the right column shows GO terms of genes with higher expression in each of the other endocrine cell types. Terms are ordered on p-value on the x-axis, with the most significant on the left. Names of relevant terms are highlighted.

(C) t-SNE map genes found upon GO-term analysis with alpha, beta or delta specific expression. Green indicates high expression.

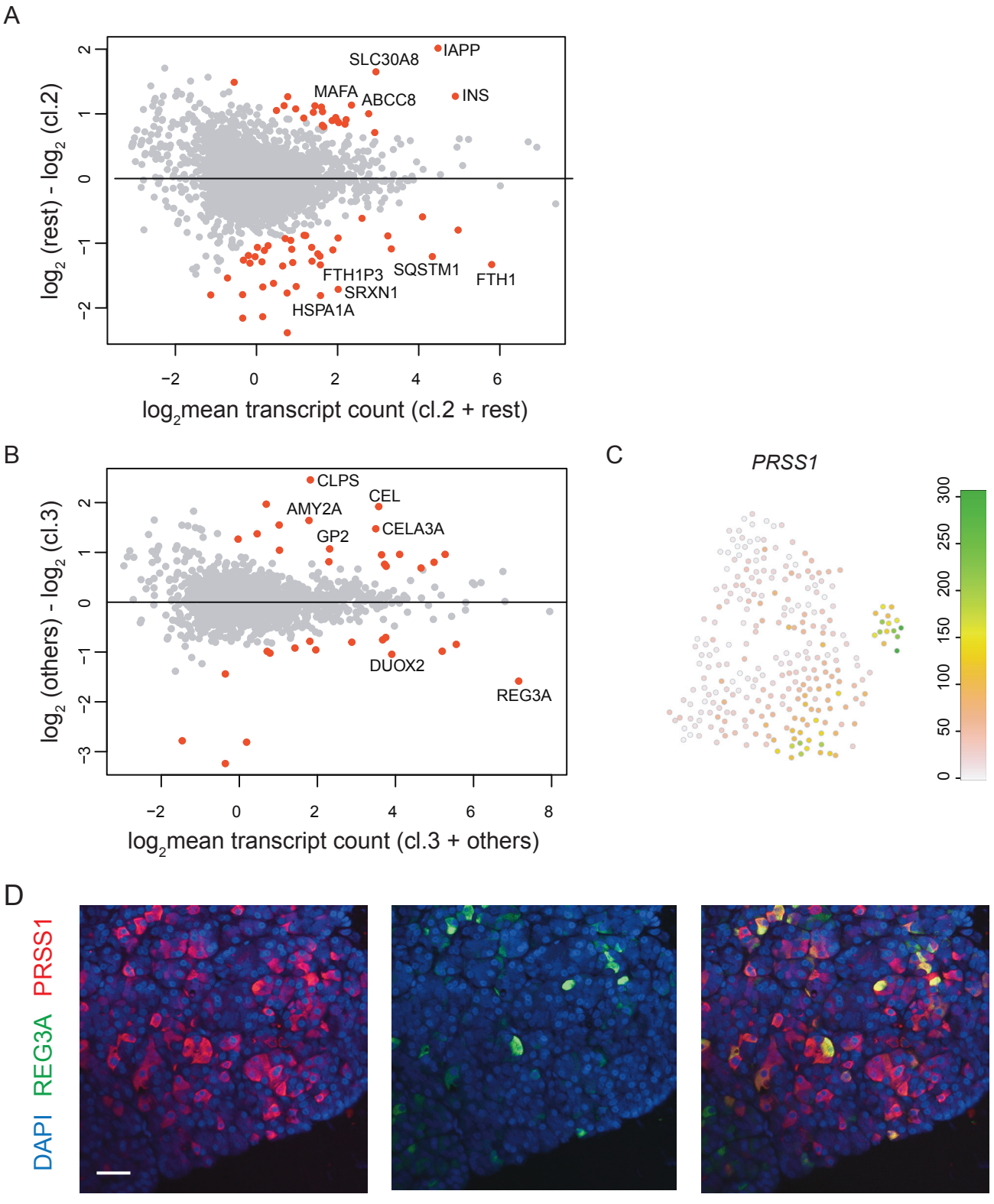

A
SLC30A8
IAPP
MAFA
ABCC8
INS
FTH1P3
SQSTM1
FTH1
SRXN1
HSPA1A
log2 (rest) - log2 (cl.2)
log2mean transcript count (cl.2 + rest)
B
CLPS
CEL
AMY2A
CELA3A
GP2
DUOX2
REG3A
log2 (others) - log2 (cl.3)
log2mean transcript count (cl.3 + others)
C
PRSS1
D
DAPI
REG3A
PRSS1

Figure S4

Figure S4. Outlier identification shows heterogeneity within acinar and beta cells, Related to Figure 3.

(A) Differential gene expression analysis of beta subclusters (*FTH1*-high cluster 2 versus the rest of the cells). Grey dots indicate genes, red dots indicate significant genes (*P*<10ˆ-6).

(B) Differential gene expression analysis between the acinar subclusters (*REG3A*-high cluster 1 versus the rest of the cells). Grey dots indicate genes, red dots indicate significant genes (*P*<10ˆ-6).

(B) t-SNE map highlighting the expression *PRSS1* across all acinar cells. Transcript counts are given in linear scale. Green indicates high expression.

(D) Immunohistochemistry showing protein expression for REG3A (green), and PRSS1 (red) with counterstaining for DAPI (blue). Scale bar is 25 μM.


A
Live
CD24 - / CD44 -
alpha
PP
endothelial
duct
delta
acinar
beta
B
GCG
INS
SST
PPY
PRSS1
KRT19
ESM1
C
GCG
INS
SST
PPY
TM4SF4

Figure S5

Figure S5. FACS Enrichment of endocrine and alpha cells based on novel cell-surface markers, Related to Figure 4

(A) t-SNE map highlighting the cells coming from the different FACS gating strategies. Each strategy is one color. Names of cell types are indicated next to their corresponding clusters. Cells sorted on only live (DAPI) marker are pink. Cells sorted against CD24 and CD44 expression are green.

(B) t-SNE maps highlighting the expression of the main pancreatic marker genes in libraries obtained by sorting for live or CD24/CD44 negative cells. Green indicates high expression.

(C) t-SNE map highlighting the expression of the main pancreatic marker genes in libraries obtained by sorting for live or CD24-/TM4SF4+/- cells. Green means high expression.

Supplemental tables and data

Donor	Sex	Age	BMI	Purity	Cell #	Type of sort
D28	Male	54	26	65	768	Live
D29	Male	23	22	95	768	Live
D30	Female	48	26	60	768	Live
D31	Male	59	25	60	768	Live
D16	Male	53	25	90	576	CD24/CD44
D25	male	30	18	55	768	CD24/TM4SF4

Table S1: Donor information (related to Figure 1)
Donor sex, age, BMI, purity of islet isolation procedure, number of processed cells and FACS gating parameters are indicated.

Table S2: **D30 Donor comparison (related to Figure 1)**
List of differentially expressed genes in each cell type between donor D30 and the other 3 donors.

Table S3: **Differential gene expression (related to figure 1)**
Differentially expressed genes between each cell type compared to all others (across all donors)

Table S4: **Differential transcription factor expression (related to figure 1)**
Differentially expressed transcription factors between each cell type compared to all others (across all donors)

Table S5: **GO-term analysis for endocrine cell types**
(Related to main text section: GO-term analysis reveals cell-type specific gene expression patterns relevant to diabetes and glucose metabolism")
GO-term analysis for alpha, beta, delta and PP cells compared to all others (across al donors)

Table S6: **Mean expression of subpopulations (related to figure 3)**
Average of gene expression across all cells of acinar and beta subpopulations.

Table S7: **Differential cell-surface marker expression (related to figure 4)**
Differentially expressed cell-surface markers factors between each cell type compared to all others (across all donors)

Data S1: (Related to Figures 1, 2, 3 and 4)
Data analysis script detailing StemID parameters and differential gene expression analysis between one cell type and all others.