

Supplementary materials and methods

Quantitative PCR (qPCR) probes

The used qPCR probes were from Life Technologies and listed as follow; Mm00440465_g1 for Neurod2, Mm00435565_m1 for Pdx1, Mm01950294_s1 for Insulin 1, Mm00731595_Gh for Insulin 2, Mm01269055_m1 for Glucagon, Mm00436671_m1 for Somatostatin, Mm01946604_s1 for Neurod1, Mm00545903_m1 for Arx, Mm01159036_m1 for Pax4, Mm00450550_m1 for Sult2b1, Mm01208248_m1 for Upk3bl, Mm01269371_m1 for Shf, Mm03053598_s1 for Mycl, Mm00459097_m1 for Steap1, Mm00803317_m1 for Rgs16, Mm04239638_m1 for Gm8773 and Mm00458153_g1 for Gng13.

Antibody list for immunohistochemistry

The following primary antibodies were used for staining: rabbit anti-Sox9 (Millipore, Cat. Nr: AB5535, Batch Nr: 3107073, dilution 1:300) (Carrasco et al., 2012), goat anti-Nkx6.1 (R&D systems, Cat. Nr: AF5857, dilution 1:300) (Bastidas-Ponce et al., 2017), rat anti-E-Cadherin (DECMA-1, a generous gift from Kremmer, dilution 1:500) (Bastidas-Ponce et al., 2017), rat anti-EpCAM (DSHB Hybridoma, Cat. Nr: 4G1, dilution 1:300), goat anti-Pancreatic polypeptide (Novus, Cat. Nr: NB100-1793, 1:300), hamster anti-Mucin-1 (Invitrogen, Cat. Nr: HM1630P0, Batch Nr: 022017, dilution 1:100), guinea pig anti-glucagon (TAKARA, Cat. Nr: M182, dilution 1:3500) (Bastidas-Ponce et al., 2017), chicken anti-GFP (to detect Venus; Aves Labs, Cat. Nr: 1020, dilution 1:1000) (Bastidas-Ponce et al., 2017), rabbit anti-insulin (Cell signaling, Cat. Nr: 3014, dilution 1:300) (Bastidas-Ponce et al., 2017), rabbit anti-Ngn3 (a generous gift from Helena Edlund, dilution 1:800), guinea pig anti-Ngn3 (a generous gift from Maïke Sander, dilution 1:1000) and rat anti-somatostatin (Invitrogen, Cat. Nr: MA5-16987, Batch Nr: 75739969, dilution 1:300). The secondary antibodies used were from Dianova and Invitrogen, (dilution 1:800) and Phalloidin Alexa Fluor 546 was from Invitrogen (Cat. Nr: A22283, dilution 1:200).

Islet isolation

The isolation of islets was performed using collagenase P (Roche) digestion of adult pancreas. First, 3 mL of collagenase P (1 mg/mL) was injected into the bile duct, then the perfused pancreas was dissected and placed into other 3 mL collagenase P for 15 min at 37 °C. To stop the digestion, 10 mL of G-solution (HBSS + 1% BSA) were added followed by centrifugation at 1600 rpm at 4 °C. After another washing step with G-solution, the pellets were re-suspended in 5.5 mL of gradient solution (5 mL 10% RPMI + 3 mL 40% Optiprep/per sample), and placed on top of 2.5 mL of the same solution. To form a 3-layers gradient, 6 mL of G-solution were added on top of the abovementioned layers. Samples were then centrifuged at 1700 rpm. Finally, the interphase between the upper and the middle

layer of the gradient was recovered and filtered through a 70 μm Nylon filter, washed with G-solution and the islets were handpicked under the microscope.

Analysis of single cell RNAseq data

Preprocessing of droplet-based single-cell RNAseq data

Demultiplexing of raw base call (BCL) files, alignment, read filtering, barcode and UMI counting were performed using the CellRanger analysis pipeline (Version 2.1.1) provided by 10X Genomics. High quality barcodes were selected based on the overall UMI distribution using the standard CellRanger cell detection algorithm. All further analyses were run using the python-based Scanpy API (Wolf et al., 2018) except stated otherwise (for software specifications and code availability see below). To further remove low quality cells, we filtered cells with a high fraction of counts from mitochondrial genes (20% or more) indicating stressed or dying cells, and cells expressing less than 1200 genes. In addition, genes with expression in less than 20 cells were excluded. Cell by gene count matrices of all samples were then concatenated to a single matrix and values log transformed. To account for differences in sequencing depth or cell size UMI counts were normalized using quantile normalization as previously described (Weinreb et al., 2018). Briefly, each cell is normalized by the total UMI count in the cell of genes that account for less than 5% of the total UMI counts across all cells. The top 4000 variable genes were selected based on normalized dispersion as described in (Zheng et al., 2017). This output matrix was input to all further analyses except for differential expression testing where all genes were used.

Low dimensional embedding, visualization and clustering

A single-cell neighborhood graph was computed on the 50 first principal components that sufficiently explain the variation in the data using 15 nearest neighbors. Uniform Manifold Approximation and Projection (UMAP) (McInnes and Healy, 2018) was run for visualization. For clustering and cell type identification louvain-based clustering (Blondel et al., 2008) at varying resolution in different parts of the data manifold was used as implemented in louvain-igraph (<https://github.com/vtraag/louvain-igraph>) and adopted by Scanpy. Cell types were annotated based on the expression of known marker genes. Clusters were merged if only reflecting further heterogeneity within a cell type not discussed in this manuscript. We removed clearly distinct clusters of cells highly expressing immune cell or mesenchymal marker genes, respectively. In addition, we found a major cluster and a small endocrine subcluster expressing a mixture of gene expression profiles of other cell types but no specific marker genes suggesting that these cells are doublets. We further checked this hypothesis and calculated a doublet score signifying the probability of being a doublet for each cell using the doublet detection package **Scrublet** (Wolock et al., 2019) using `scr.compute_doublet_scores` function with parameters as recommended in the Scrublet tutorial (`min_counts = 2`, `min_cells = 3`, `vscore_percentile = 85`, `n_pc = 50`, `expected_doublet_rate = 0.08`, `sim_doublet_ratio = 3`, `n_neighbors = 15`). Indeed, the majority of

these cells were predicted to be doublets even when setting a high threshold for the doublet score and we therefore removed these two clusters from further analyses. In all other subtypes we found only a small proportion of predicted doublet cells that did not form a separate cluster or misleadingly connect two clusters. We did not remove these cells as downstream analyses were not strongly influenced by them. In different figures varying resolutions of subtype clustering are shown. EP, Pre-endocrine and α -cell subtypes were annotated based on marker gene expression as well as the developmental stages they existed. For the specification of $Ngn3^{low}$ progenitors we referred to the expression of *Neurog3* and picked all cells with expression > 0 in the multipotent, tip, trunk and ductal cluster. Proliferating subtypes were annotated using the cell cycle classification described below.

Marker gene identification and subtype characterization

Characteristic gene signatures were identified by testing for differential expression of a subgroup against all other cells or between two subgroups as outlined in the text using a t-test with overestimated variance implemented in the *tl.rank_genes_groups* function of Scanpy. Testing was performed on the log-transformed quantile normalized data to account for confounding differences in sequencing depth between samples. It has been shown that a simple t-test performs comparable to differential expression tools adopting more complex models in a simple set-up as here (no additional confounding factors) (Soneson and Robinson, 2018). Gene set enrichment was performed with the *gseapy* (v0.9.3) implementation of EnrichR. To identify genes enriched in EPs, $Ngn3^+$ cells ($Ngn3^{low}$ progenitors and $Ngn3^{high}$ precursors) were compared against all other cells except for acinar cells. Top 250 upregulated genes were considered (Table S2). From this list, EP signature genes that are transiently expressed in EPs similar to *Ngn3* but not or only lowly expressed in other populations were extracted by computing the mean expression of all 250 genes in each cell type after scaling the data from 0 to 1. All genes that showed a mean expression > 0.04 in $Ngn3^{high}$ precursors, < 0.04 in all other non-endocrine progenitor-precursor clusters and < 0.18 in Fev^{high} cells were defined as EP signature genes. Expression thresholds were chosen based on the mean expression levels of *Neurog3*. For characterization of the EP subtypes we compared ductal cells against $Ngn3^{low}$ progenitors, $Ngn3^{low}$ progenitors against ductal cells, $Ngn3^{high}$ precursors against $Ngn3^{low}$ progenitors and Fev^{high} against $Ngn3^{high}$ precursors to identify genes changing along the differentiation trajectory.

Cell cycle classification

To classify cells into cell cycle phases we used a cell scoring function as described by (Satija et al., 2015) and implemented in the *tl.score_genes_cell_cycle* function in Scanpy with default parameters. Shortly, the score is the average expression of the gene set subtracted with the average expression of a randomly sampled background set with expression values within the same range. Cell cycle genes as defined in (Tirosh et al., 2016) were used to distinguish between S, G2/M and G1 phase cells.

Reconstruction of lineage relationships and differentiation trajectories

To infer lineage relationships between clusters and predict potential routes of differentiation partition-based graph abstraction (PAGA) was performed (Wolf et al., 2019) using the *tl.paga* function of Scanpy with a threshold on edge significance of 0.11 for all subtypes and 0.1 for the endocrine lineage. Edge weights represent the confidence of a connection calculated based on a measure for connectivity. Paths in the PAGA graph signify cluster relationships indicating potential differentiation paths. To infer a pseudotemporal ordering of the cells along the predicted routes in the PAGA graph, diffusion pseudotime (dpt) (Haghverdi et al., 2016) was used as implemented in Scanpy (*tl.dpt*) setting a root cell within the starting population. Dpt is designed to learn such continuous cellular differentiation trajectories and projects the differentiating cells to a one-dimensional developmental trajectory. Before dpt inference the single cell neighborhood-graph and diffusion map was recalculated for the represented subset of cells. To test whether cell cycle effects mask other biological processes and affect the pseudotemporal ordering, dpt was recomputed after linearly regressing out S and G2M cell scores (using *sc.pp.regress_out*). Removal of cell cycle effects did not considerably change the ordering of the cells. Also, the expression of genes described in the manuscript figures was not affected except for, as expected, genes involved in DNA replication and cell cycle.

Directionality of endocrine differentiation using RNA velocity estimation

To infer directionality of endocrine differentiation we performed RNA velocity estimation (La Manno et al., 2018) implemented as a stochastic version in the scVelo python package (<https://github.com/theislab/scvelo>). We extracted spliced and unspliced reads using the velocityto pipeline (<http://velocityto.org>), which generated a loompy file. The file was read into an AnnData object for downstream analysis with Scanpy. All analysis related to RNA velocity such as estimating velocities and obtaining the RNA velocity force field were done using scVelo. Here, we followed the recommended procedures described in scVelo, consisting of four main steps: preprocessing, computing first- and second-order moments, estimating velocities, and constructing a velocity graph. That is, we filtered genes with less than 10 spliced and 10 unspliced counts and normalized unspliced and spliced reads by their initial counts and log transformed the data. Then, for each stage separately, we recomputed the neighborhood graph on the represented cell subset using 15 neighbors and the 50 first PCs that sufficiently explain the variation in the data, and the UMAP embedding. Next, first-order and second-order moments for each cell were computed across its nearest neighbors. Velocities were estimated using a stochastic model of transcriptional dynamics, whereby we used a 95% quantile fit for more conservative velocity estimates. Finally, the velocity graph was obtained by computing the correlations between potential cell transitions and the predicted cell state change given by the velocity vector. Only genes with an $r^2 > 0.1$ were considered. The graph was then used to project the velocities into the low dimensional UMAP embedding.

Software specifications

All analyses from UMI count matrices were run with python 3 with the Scanpy API v.1.3.2 to v.1.3.4 and anndata v.0.6.11 (<https://github.com/theislab/scanpy>). Versions of packages required by Scanpy that might influence numerical results are indicated in the custom scripts. Unspliced and spliced reads were extracted using velocity v0.17.7 and loompy v2.0.12 (<http://velocityto.org>). Velocities were calculated with scVelo v0.1.16 (<https://github.com/theislab/scvelo>). All figures were plotted with matplotlib and seaborn and data was exported to excel-sheets with xlswriter. Custom scripts of code for all analyses will be made available upon publication.

Supplementary references

- Bastidas-Ponce, A., Roscioni, S. S., Burtscher, I., Bader, E., Sterr, M., Bakhti, M. and Lickert, H.** (2017). Foxa2 and Pdx1 cooperatively regulate postnatal maturation of pancreatic β -cells. *Mol. Metab.* **6**, 524–534.
- Blondel, V. D., Guillaume, J. L., Lambiotte, R. and Lefebvre, E.** (2008). Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* **10**, P10008.
- Carrasco, M., Delgado, I., Soria, B., Martín, F. and Rojas, A.** (2012). GATA4 and GATA6 control mouse pancreas organogenesis. *J. Clin. Invest.* **122**, 3504–3515.
- Haghverdi, L., Büttner, M., Wolf, F. A., Buettner, F. and Theis, F. J.** (2016). Diffusion pseudotime robustly reconstructs lineage branching. *Nat. Methods* **13**, 845–8.
- La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastrioti, M. E., Lönnerberg, P., Furlan, A., et al.** (2018). RNA velocity of single cells. *Nature* **560**, 494–498.
- McInnes, L. and Healy, J.** (2018). UMAP : Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv:1802.03426* 1–18.
- Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. and Regev, A.** (2015). Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502.
- Soneson, C. and Robinson, M. D.** (2018). Bias, robustness and scalability in single-cell differential expression analysis. *Nat. Methods* **15**, 255–261.
- Tirosh, I., Izar, B., Prakadan, S. M., Wadsworth, M. H., Treacy, D., Trombetta, J. J., Rotem, A., Rodman, C., Lian, C., Murphy, G., et al.** (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science (80-.)*. **352**, 189–96.
- Weinreb, C., Wolock, S. and Klein, A. M.** (2018). SPRING: A kinetic interface for visualizing high dimensional single-cell expression data. *Bioinformatics* **34**, 1246–1248.
- Wolf, F. A., Angerer, P. and Theis, F. J.** (2018). SCANPY: Large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15.
- Wolf, F. A., Hamey, F., Plass, M., Solana, J., Dahlin, J. S., Gottgens, B., Rajewsky, N., Simon, L. and Theis, F. J.** (2019). PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biol.* **20**, 59.
- Wolock, S. L., Lopez, R. and Klein, A. M.** (2019). Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. *Cell Syst.* **8**, 281–291.e9.
- Zheng, G. X. Y., Terry, J. M., Belgrader, P., Ryvkin, P., Bent, Z. W., Wilson, R., Ziraldo, S. B., Wheeler, T. D., McDermott, G. P., Zhu, J., et al.** (2017). Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* **8**, 14049.

Supplementary figure legends

Supplementary figure 1. Characterization of endocrine cell formation and function in NVF mice.

(A) Immunohistochemical analysis shows similar spatio-temporal expression pattern of NVF and endogenous Ngn3 expression at E16.5 and E18.5 heterozygous NVF pancreatic section. Scale bar, 20 μm . (B) Plot profile of Venus staining together with two different antibodies against endogenous Ngn3 in WT and homozygous NVF pancreatic sections. No Venus signal was detected in the WT sample. Scale bar, 20 μm . (C) qPCR analysis of isolated cells from E15.5 WT and homozygous NVF pancreata for expression of Ngn3 target genes (N= 1 experiment derived from 8-10 embryos). (D) Staining of PP-cells in E18.5 pancreatic sections from WT and homozygous NVF mice show no significant difference in the formation of these cells. Scale bar, 50 μm . (E) qPCR analysis of key pancreatic hormones from isolated islets from 3-month-old WT and homozygous NVF mice (N = 3). (F) Normoglycemic levels in 3-month-old homozygous NVF mice compared to WT. (G) Glucose-stimulated insulin secretion test shows no differences in insulin secretion from isolated islets from 3-month-old homozygous NVF and WT mice.

Supplementary figure 2. Cell cycle classification and *Fev*^{high} cell-enriched genes.

(A) UMAP plots of proliferative status of all pancreatic epithelial populations. Classification into cycling and non-cycling cells (left) based on cell score levels for cell cycle phases S and G2/M calculated based on the expression of a set of genes involved in cell cycle progression (right). (B) Heatmap showing the mean expression in *Ngn3*⁺ and *Ngn3*⁻ cells of the 58 EP-signature genes. Expression values are normalized and scaled to unit variance. (C) qPCR analysis of E15.5 isolated pancreatic epithelial cells indicates no significant changes in the expression levels of several EP-signature genes in WT and homozygous NVF mice (N= 1 experiment derived from 8-10 embryos). (D) Gene expression along pseudotime from *Ngn3*^{low} progenitors to endocrine cells. The top 200 genes enriched in *Fev*^{high} cells are shown (Table S3) and selected genes highlighted. Expression is normalized and approximated by polynomial regression fits along pseudotime. Fitted values of each gene are then scaled to the range between 0 and 1. Cluster membership of the cells is indicated at the bottom. (E) Gene expression along pseudotime of selected genes starting expression in *Fev*^{high} cells and continuously expressed in differentiated endocrine cells. Lines are polynomial regression fits of normalized data. Cluster membership of the cells is indicated at the bottom.

Supplementary figure 3. Pathway analysis of EPs and *Fev*^{high} cells.

(A) Heatmap showing the top 20 genes of subtype-specific expression profiles of ductal, *Ngn3*^{low} progenitor, *Ngn3*^{high} precursor and *Fev*^{high} cells. Normalized expression values are shown. (B) GO-term enrichment analysis of gene sets specific to ductal, *Ngn3*^{low} progenitors, *Ngn3*^{high} precursors and *Fev*^{high} populations.

Supplementary figure 4. Early and late α -cells are transcriptionally heterogeneous. (A) Expression of selected differentially expressed genes between primary and secondary α -cells indicates stage-dependent α -cell heterogeneity. Normalized expression values are shown. (B) Proportion of primary and secondary α -cells from E12.5 to E15.5 developmental stages. (C) Dotplot of the expression levels of EP-signature genes in $Ngn3^{\text{low}}$ progenitors derived from different developmental stages. (D) The expression of *Pax4* and *Arx* in different pancreatic clusters overlaid onto the PAGA graph. The expression of *Pax4* starts in late $Ngn3^{\text{high}}$ precursors, while the expression of *Arx* is mainly evident in Fev^{high} alpha cells. Mean of normalized expression per cluster is shown.

Figure S1

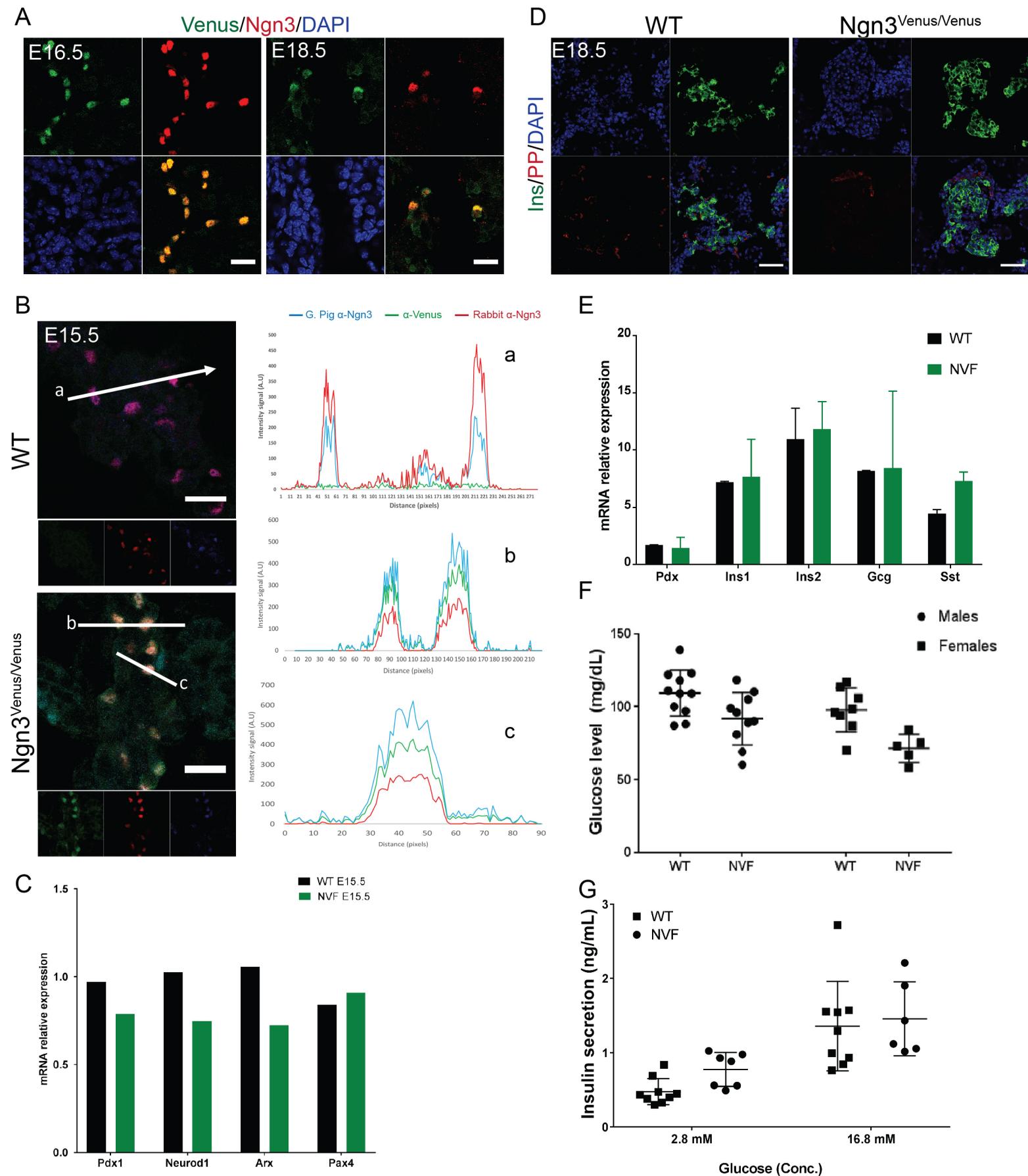
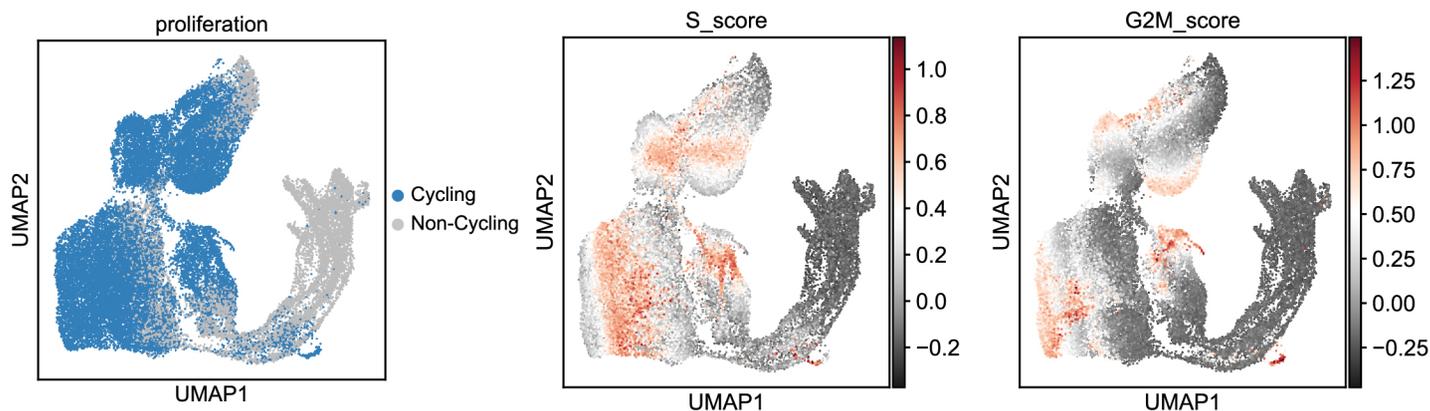
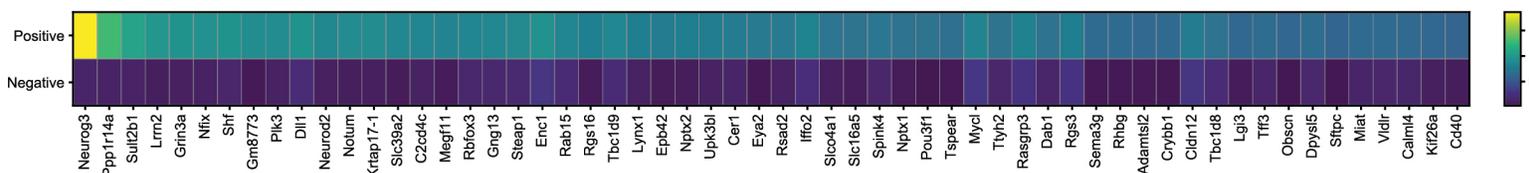


Figure S2

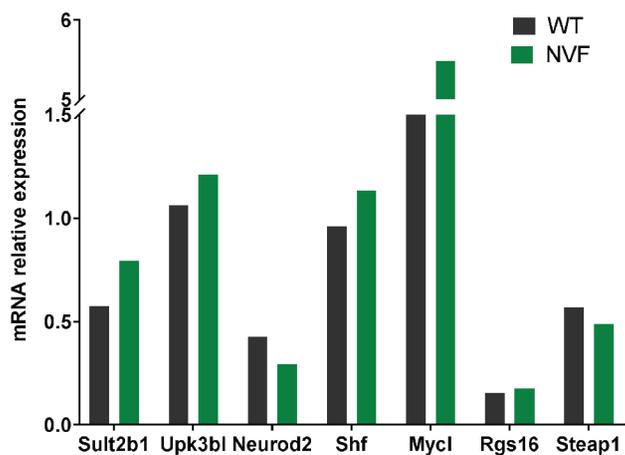
A



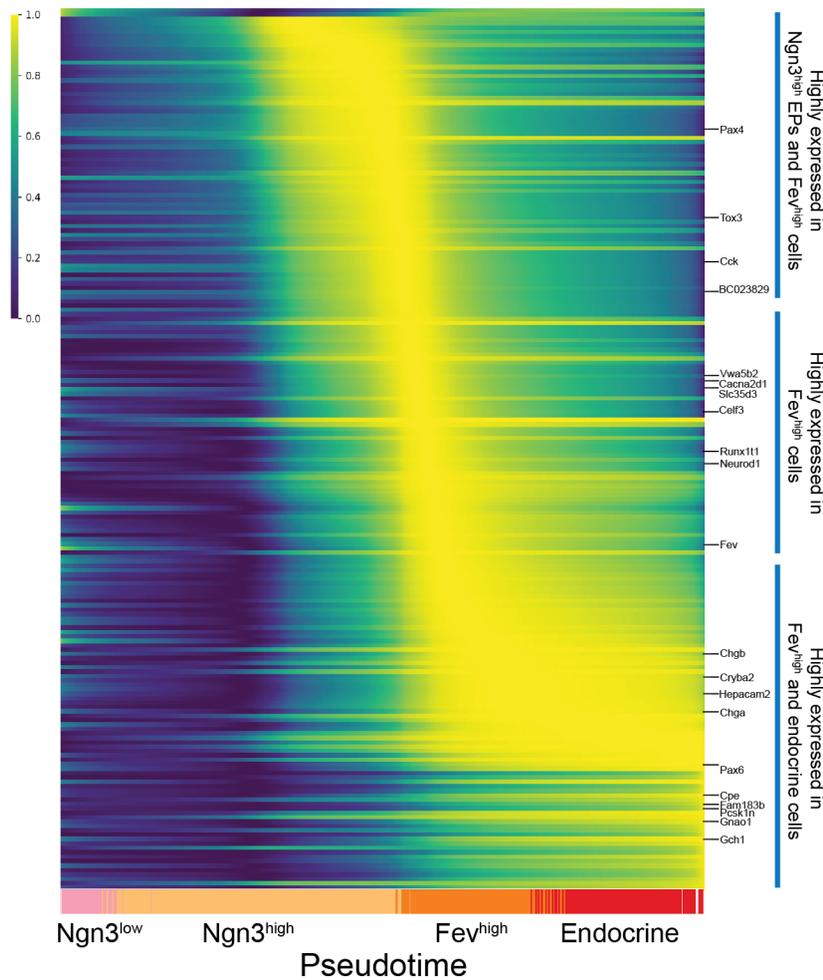
B



C



D



E

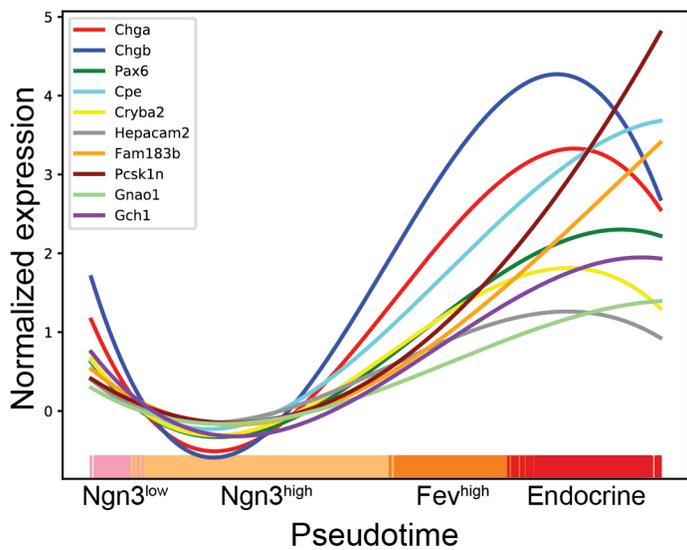
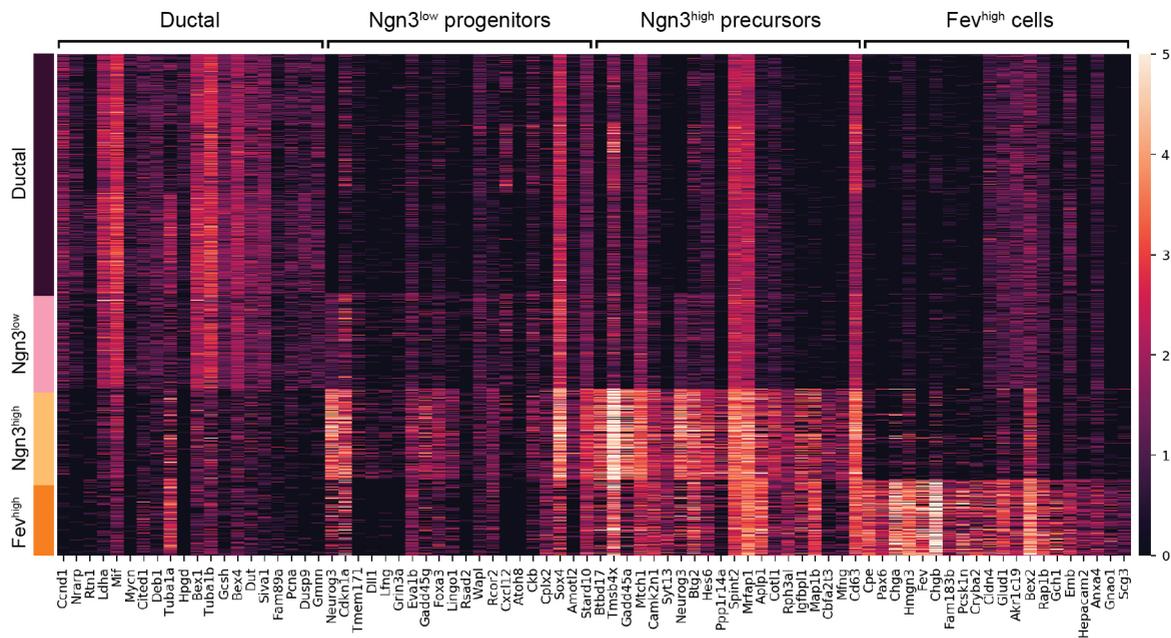


Figure S3

A



B

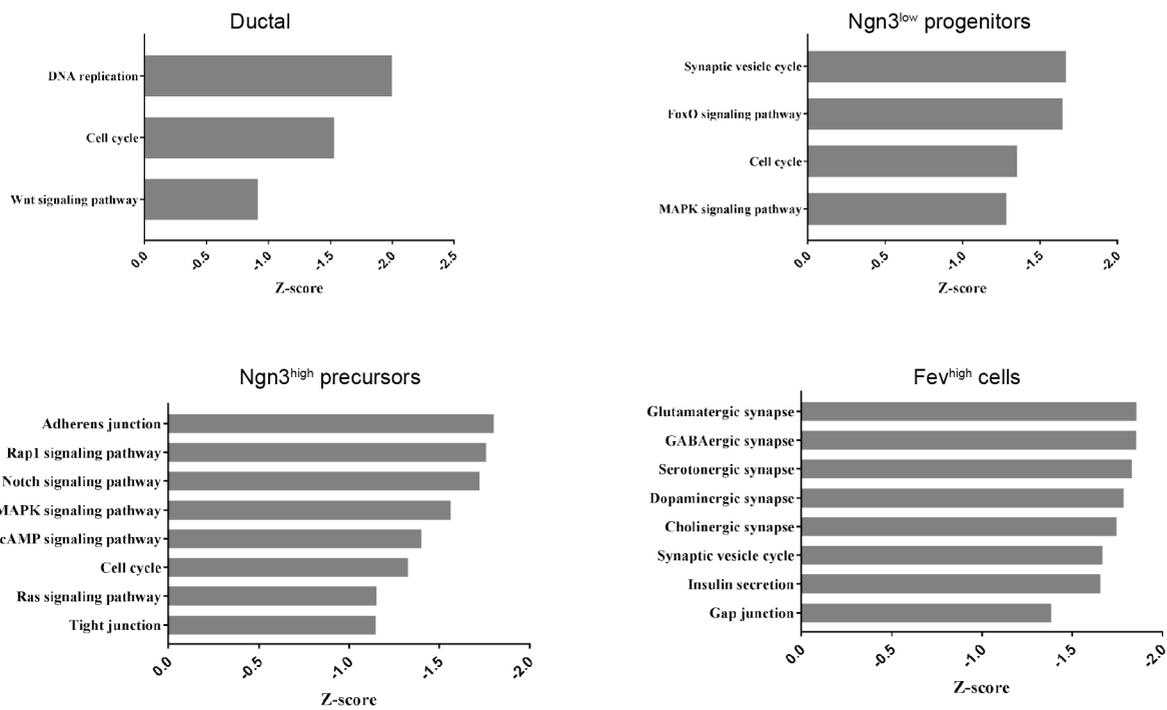


Figure S4

