

Human Pancreatic α -Cell Heterogeneity and Trajectory Inference Analyses Reveal *SMOC1* as a β -Cell Dedifferentiation Gene

Randy B. Kang¹, Miguel Varela¹, Eunjin Oh¹, Jungeun Lee¹, Tuo Zhang², Esra Karakose³, Andrew F. Stewart³, Donald K. Scott³, Debbie C. Thurmond¹, Adolfo Garcia-Ocana^{1*} and Geming Lu^{1*}

¹Department of Molecular and Cellular Endocrinology, Arthur Riggs Diabetes and Metabolic Research Institute, Beckman Research Institute of City of Hope, Duarte, CA, 91010.

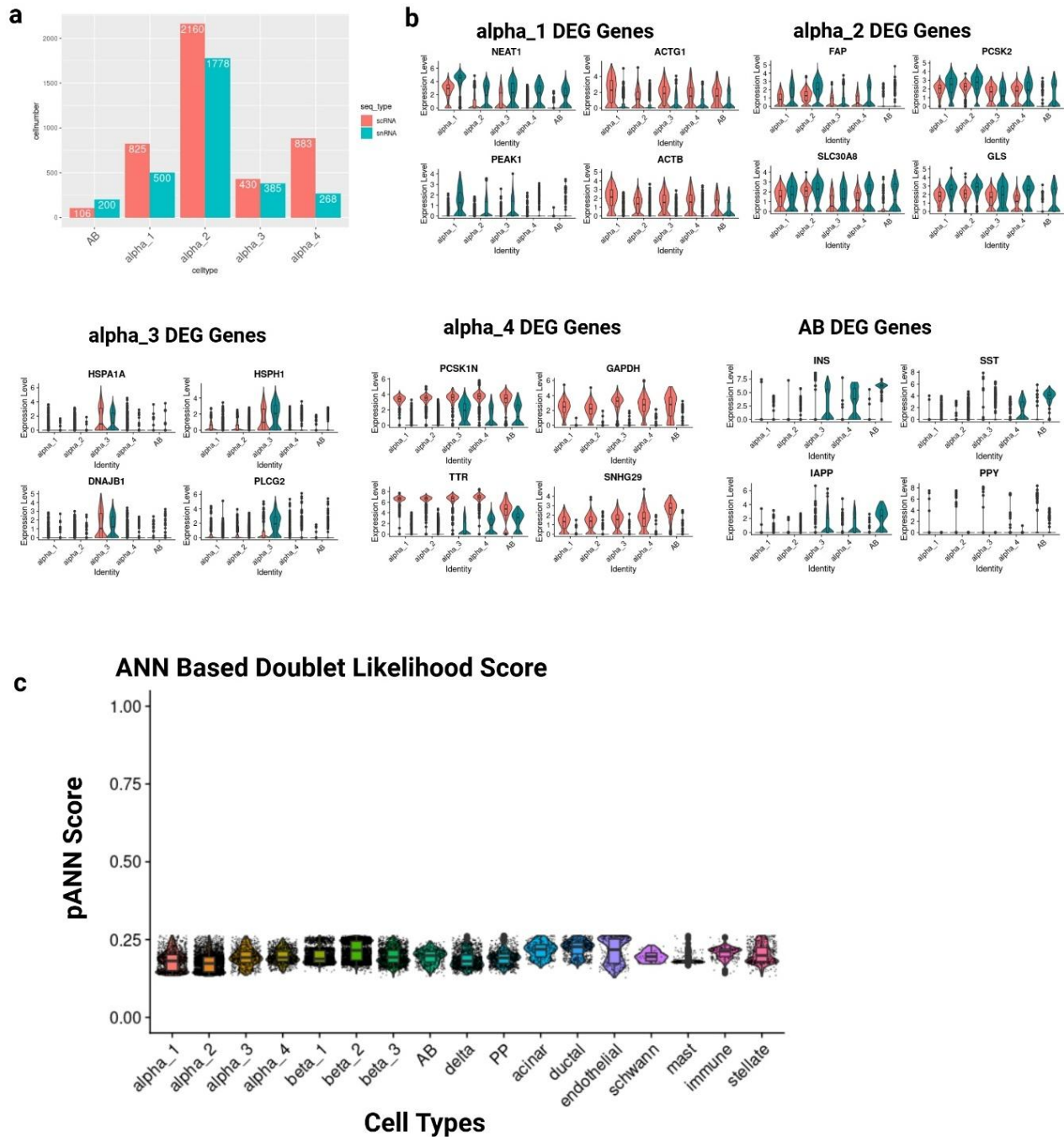
²Genomics Resources Core Facility, Weill Cornell Medicine, New York, NY, 10065.

³Diabetes Obesity and Metabolism Institute, and Division of Endocrinology, Diabetes and Bone Diseases, Icahn School of Medicine at Mount Sinai, New York, NY, 10029.

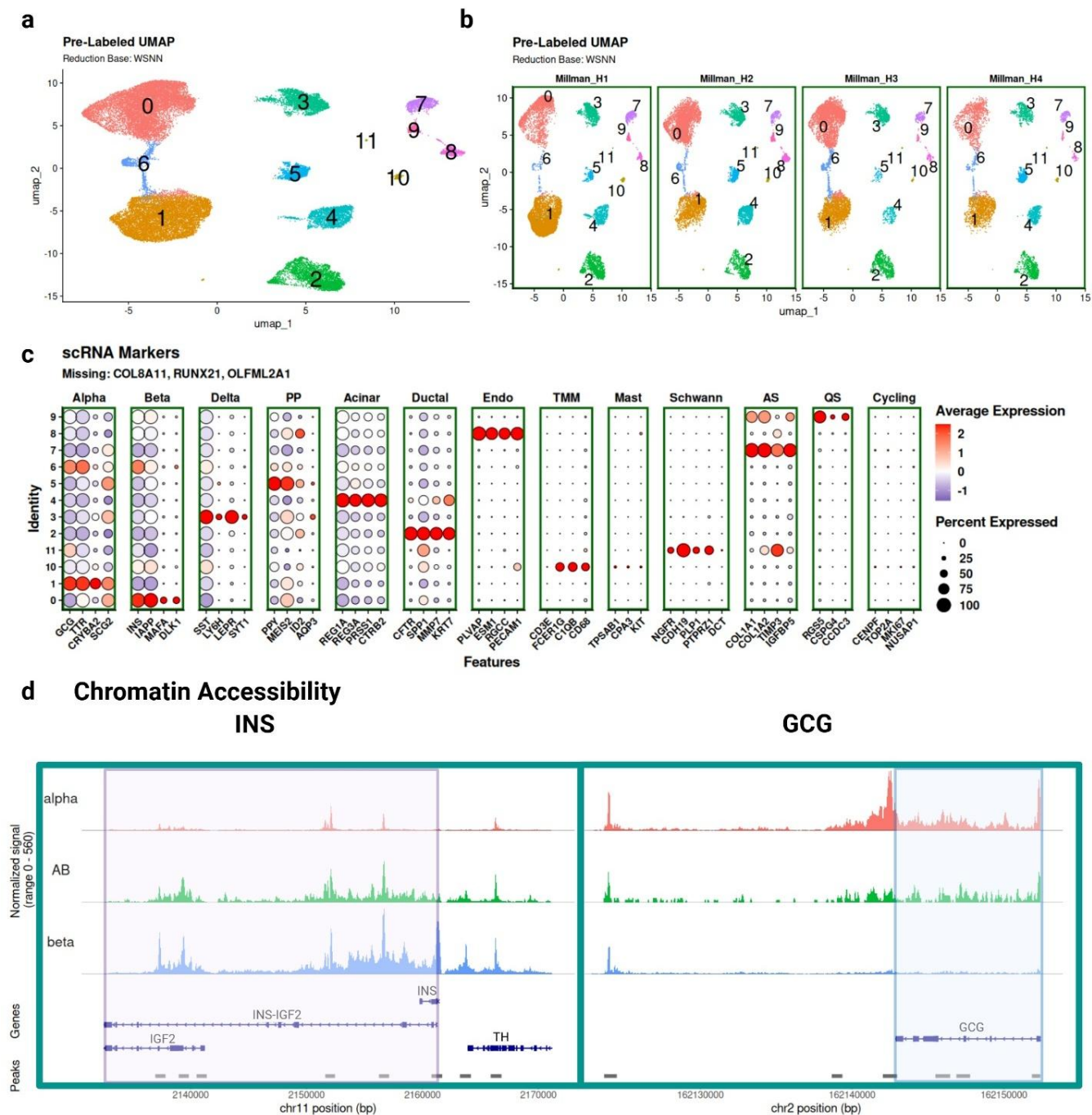
*Equally contributing authors.

Correspondence To:

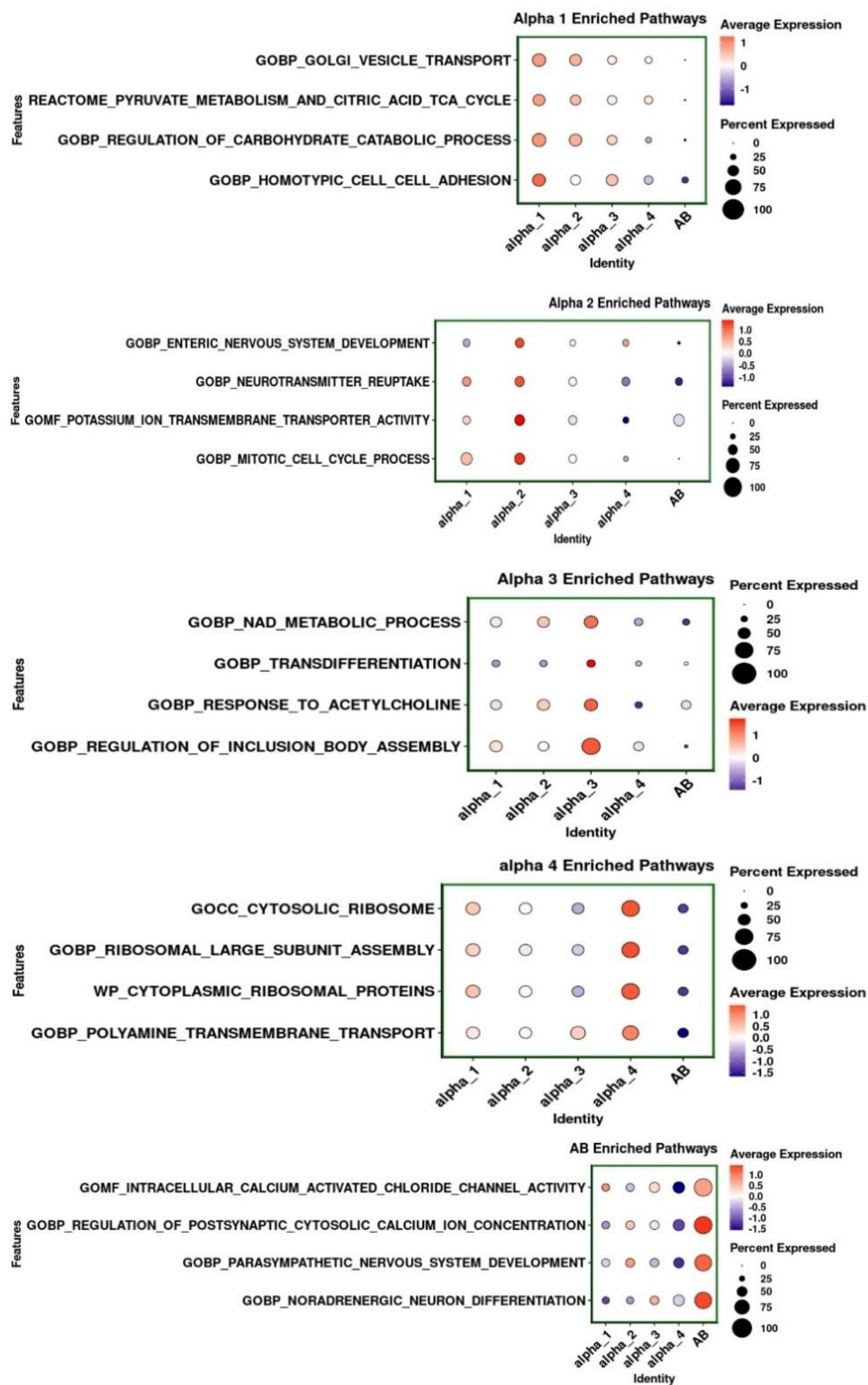
Adolfo Garcia-Ocana, PhD
Geming Lu, MD and
Department of Molecular and Cellular Endocrinology
Diabetes and Metabolism Research Institute
Beckman Research Institute
City of Hope
1500 East Duarte Road
Duarte, CA 91010
agarciaocana@coh.org
gelu@coh.org



Supplementary Figure 1. Cell numbers, gene expression, DEGs and pANN score of human α -cell subclusters by scRNA-seq and snRNA-seq of islets from adult human non-diabetic donors. **a. Number of cells in each α -cell subpopulation in scRNA-seq and snRNA-seq of islets from non-diabetic donors used in these experiments ($n=3$ pairs of human islet preparations). **b.** Differentially expressed genes for each α -cell subpopulation: α_1 with *NEAT1*, *ACTG1*, *PEAK1*, *ACTB*; α_2 with *FAP*, *PCSK2*, *SLC30A8*, *GLS*; α_3 with *HSPA1A*, *HSPH1*, *DNAJB1*, *PLCG2*; α_4 with *PCSK1N*, *GAPDH*, *TTR*, *SNHG29*; and AB with *INS*, *SST*, *IAPP*, *PPY* separated by scRNA-seq (red) and snRNA-seq (green). **c.** DoubletFinder Algorithm pANN (proportion of artificial nearest neighbors) score to distinguish between doublets and singlets in human islet cells.**

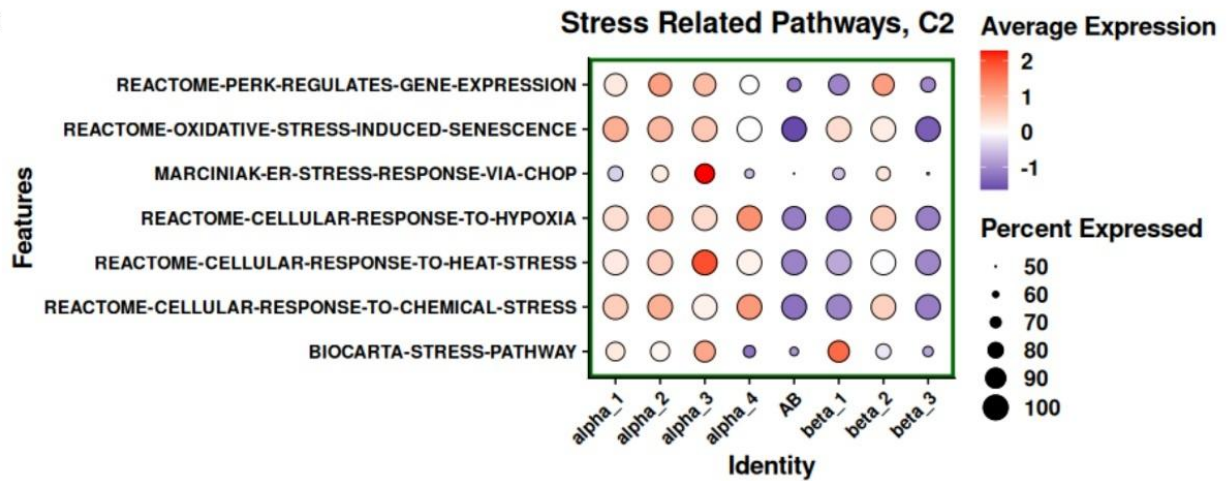


Supplementary Figure 2. External snRNA/snATAC-seq multiome data analysis. **a.** UMAP visualization of the WNN-integrated multi-omics dataset (34). **b.** UMAP divided by individual donor **c.** Gene expression profiles using canonical scRNA-seq cell type markers **d.** Chromatin opening of the *INS* and *GCG* genes in α -, AB and β -cell clusters.

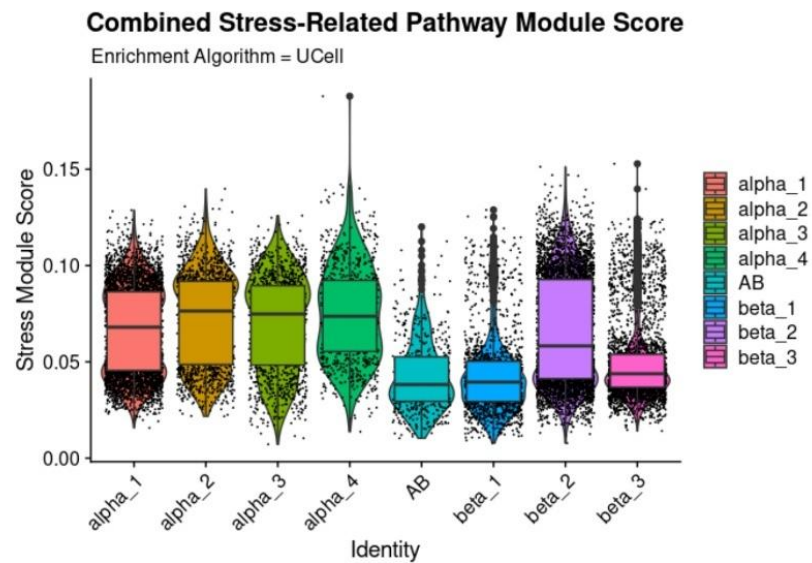


Supplementary Figure 3. Pathway enrichment analysis of human α -cell subclusters in islets from adult human non-diabetic donors. The top four pathways according to the p-value and FDR of each α -cell subcluster by non-targeted GSEA.

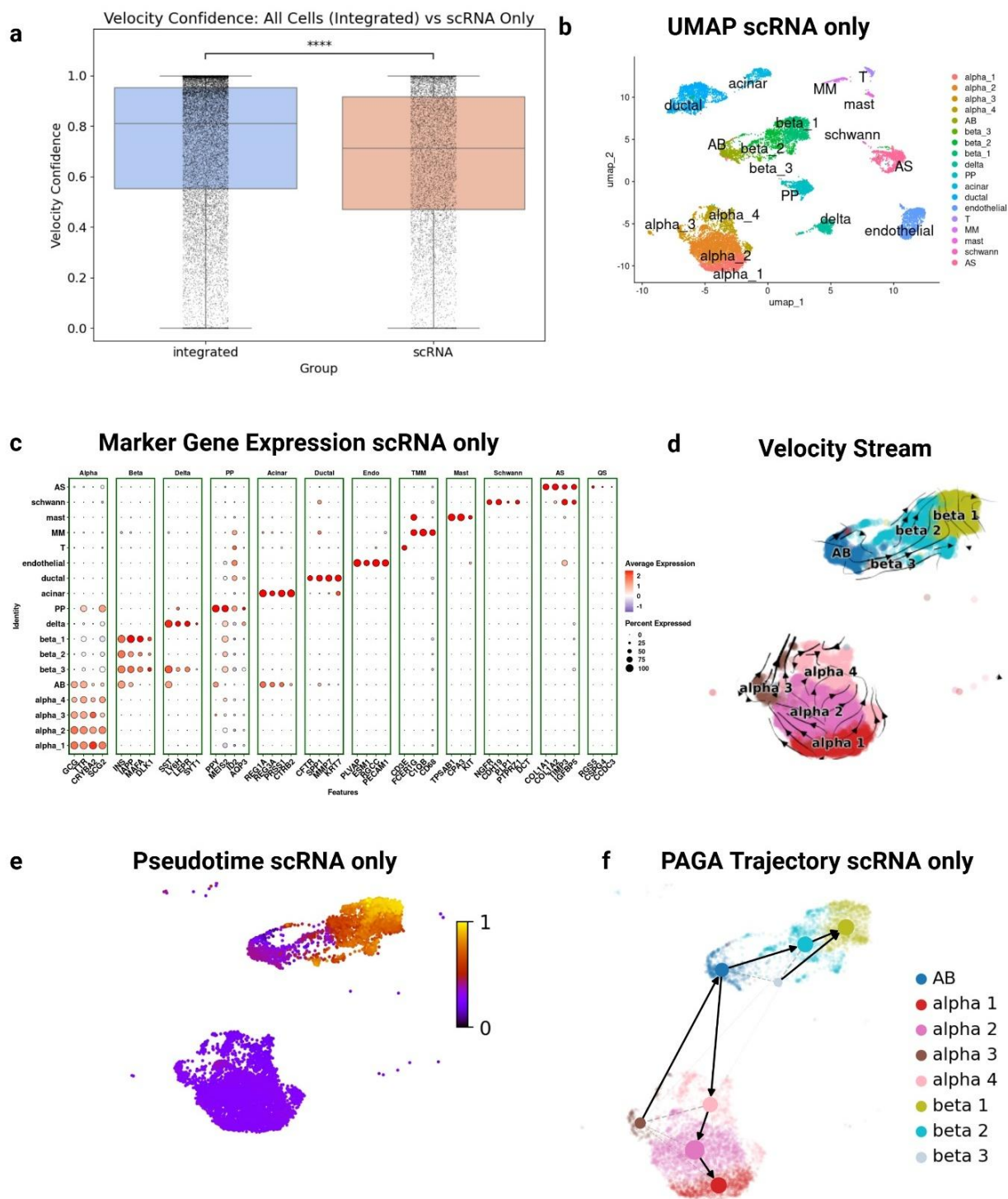
a



b

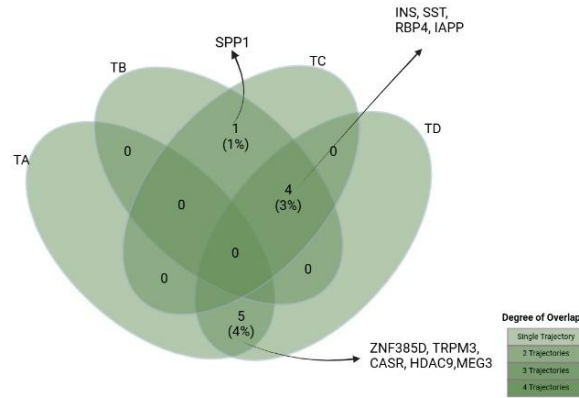


Supplementary Figure 4. GSEA of stress pathways in α - and β -cell subclusters. a. Stress related pathways in C2 and **b.** combined stress-related pathways scores for α - and β -cell subclusters.

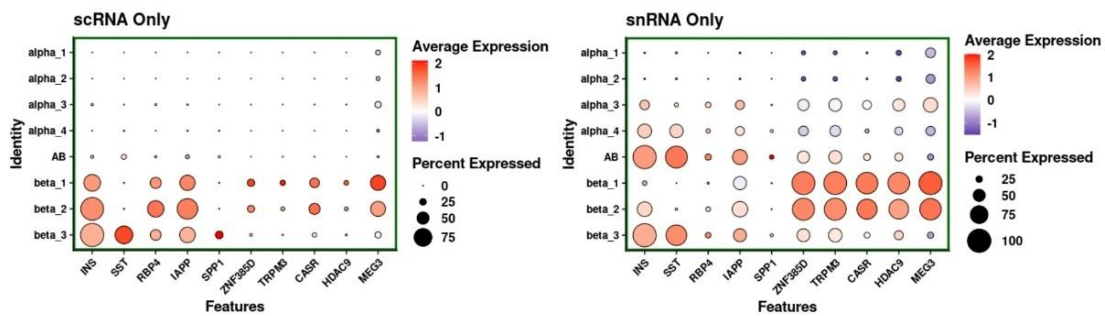


Supplementary Figure 5. RNA velocity and trajectories only in scRNA-seq analysis of islets from adult human non-diabetic donors, and RNA velocity confidence scores. **a.** Comparison of RNA velocity confidence scores between only scRNA-seq and integrated scRNA-seq samples. Statistical significance was assessed using a two-sided Mann–Whitney–Wilcoxon test. **** $p < 0.0001$. **b.** UMAP depicting human islet cell type annotation of the scRNA-seq dataset. **c.** Expression level and percentage of cells expressing canonical genes of annotated human islet cell types in the scRNA-seq data. **d.** RNA velocity streamline among α - and β -cell subclusters with scVelo. **e.** Pseudotime in the subset of α - and β -cell subclusters. **f.** PAGA trajectory analysis between α - and β -cells showing 4 main trajectories: 1) from AB to β_1 ; 2) from β_3 to β_1 ; 3) from AB to α_1 ; 4) from α_3 to AB.

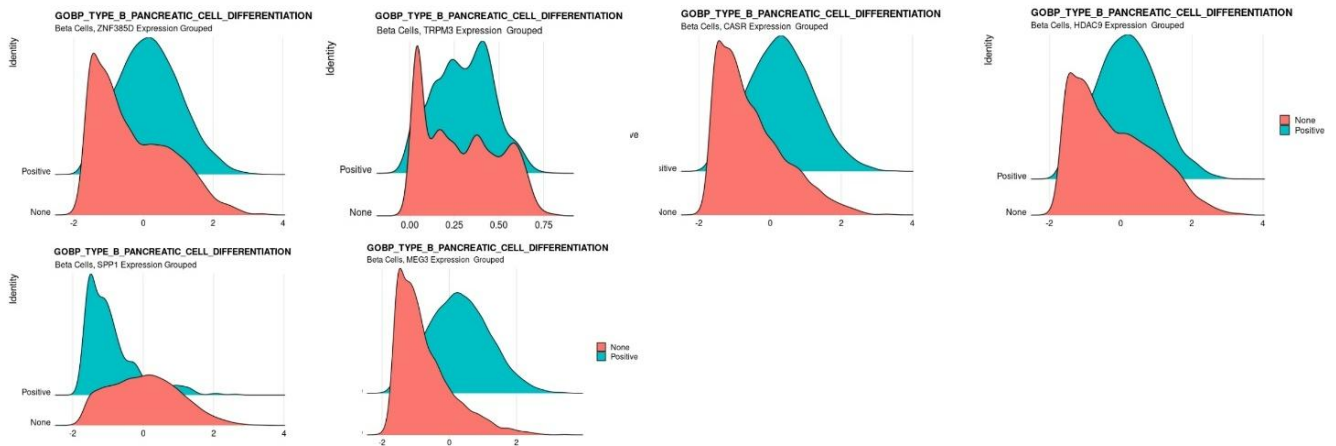
a Commonality to Beta Cell Trajectory



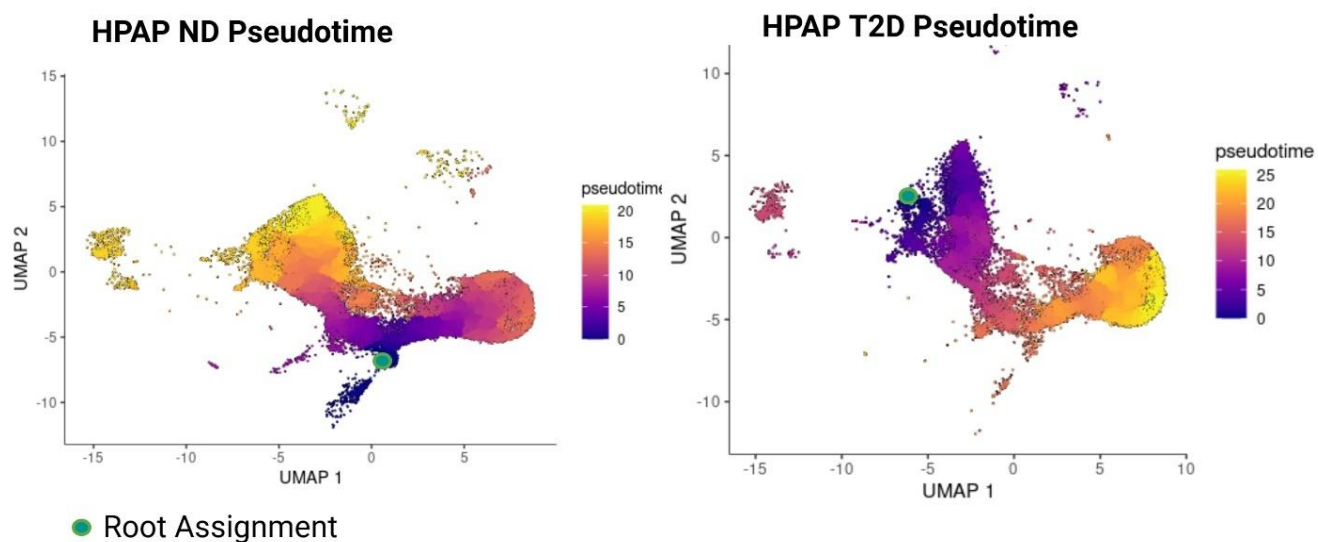
b



c



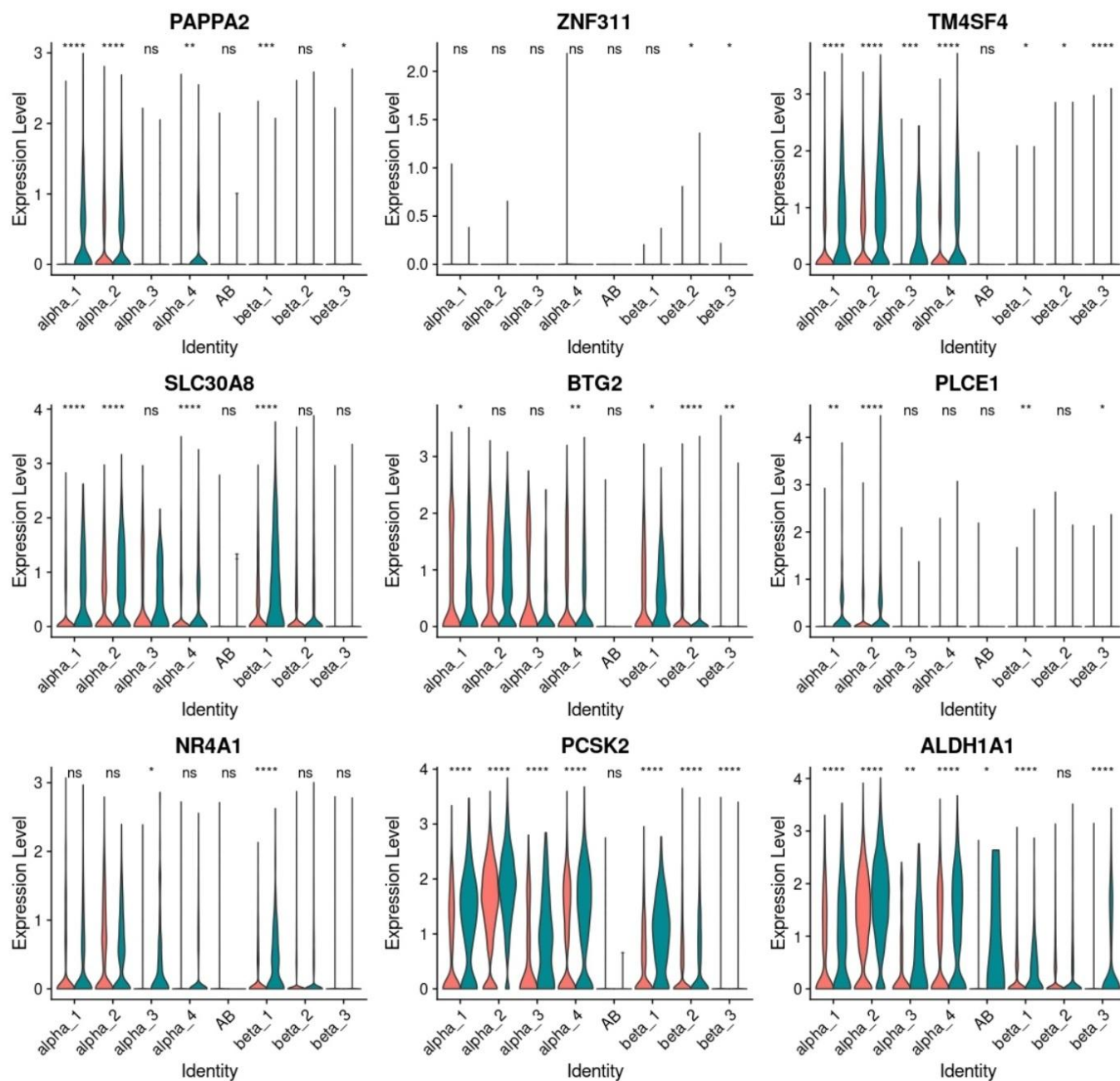
Supplementary Figure 6. Gene commonality analysis towards β -cell transition in islets from adult human non-diabetic donors. **a.** Genes in increasing pattern in Trajectory A and decreasing patterns in Trajectories B, C, and D were intersected leading to common genes among the different trajectories. **b.** Dot plots representing gene expression of the commonality genes between Trajectories A, B, C and D in scRNA-seq and snRNA-seq analysis of human islets. **c.** Enrichment ridge plot of GOBP's pancreatic β -cell differentiation pathway for *ZNF385D*, *CASR*, *HDAC9*, *MEG3*, *SPP1* and *TRPM3* genes. We grouped β -cells in two groups by expression level for each gene – negative (0% normalized expression), positive (>0%).



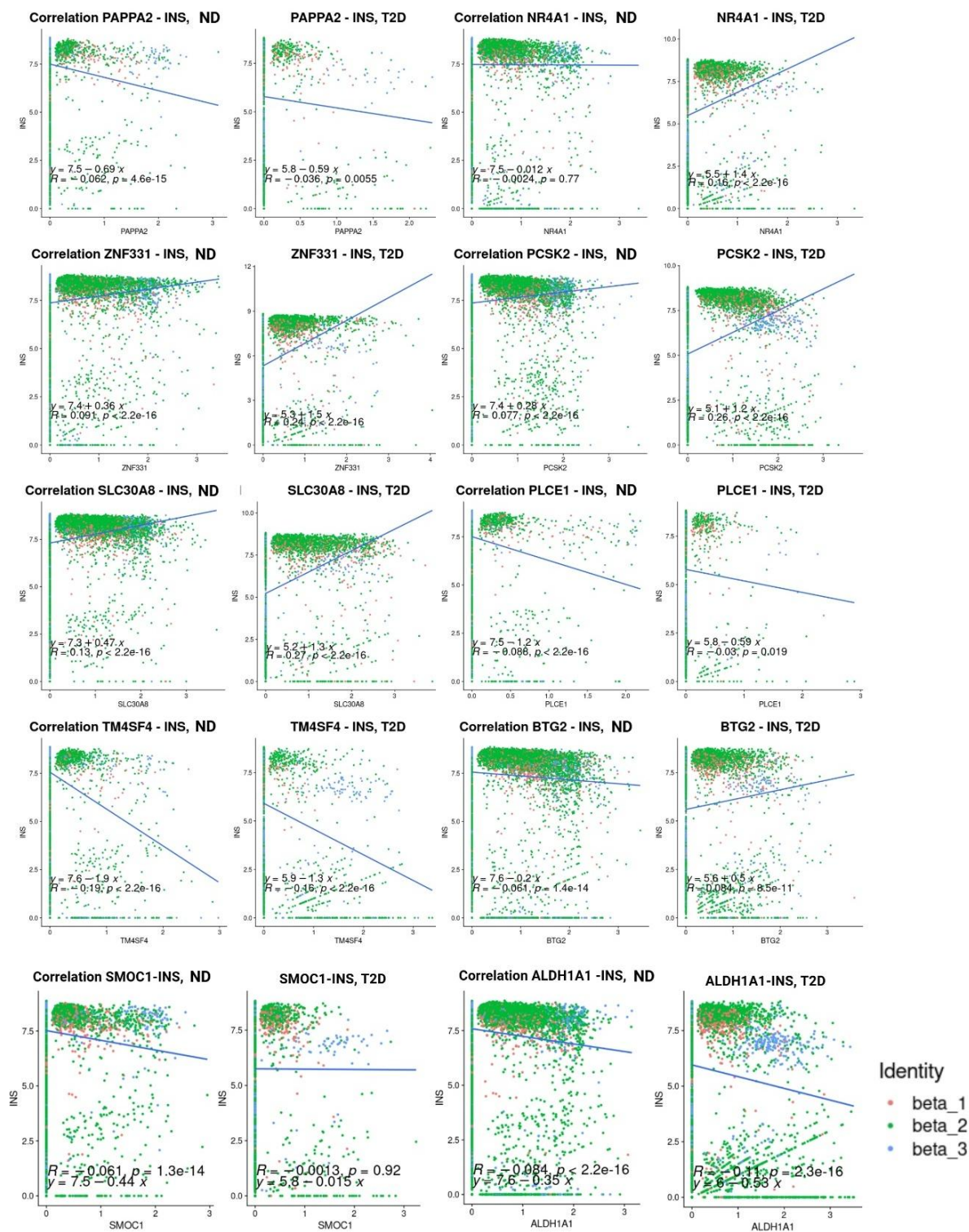
Supplementary Figure 7. Pseudotime trajectory with Monocle 3, root assigned by PAGA analysis.
a. HIRN-HPAP non-diabetic human islet cell data, alpha and beta cell clusters. **b.** HIRN-HPAP type 2 diabetic human islet cell data, alpha and beta cell clusters.

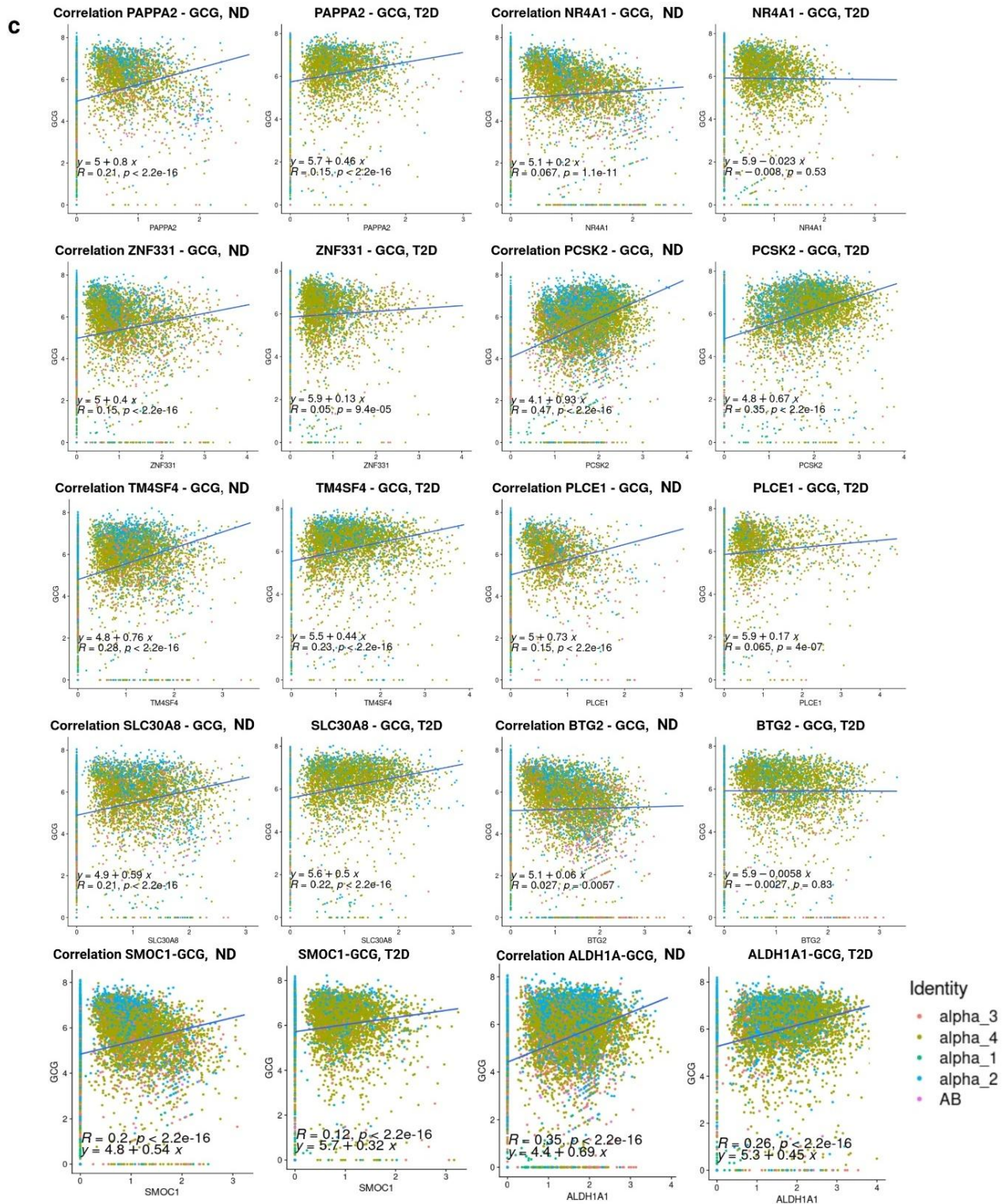
a

ND T2D

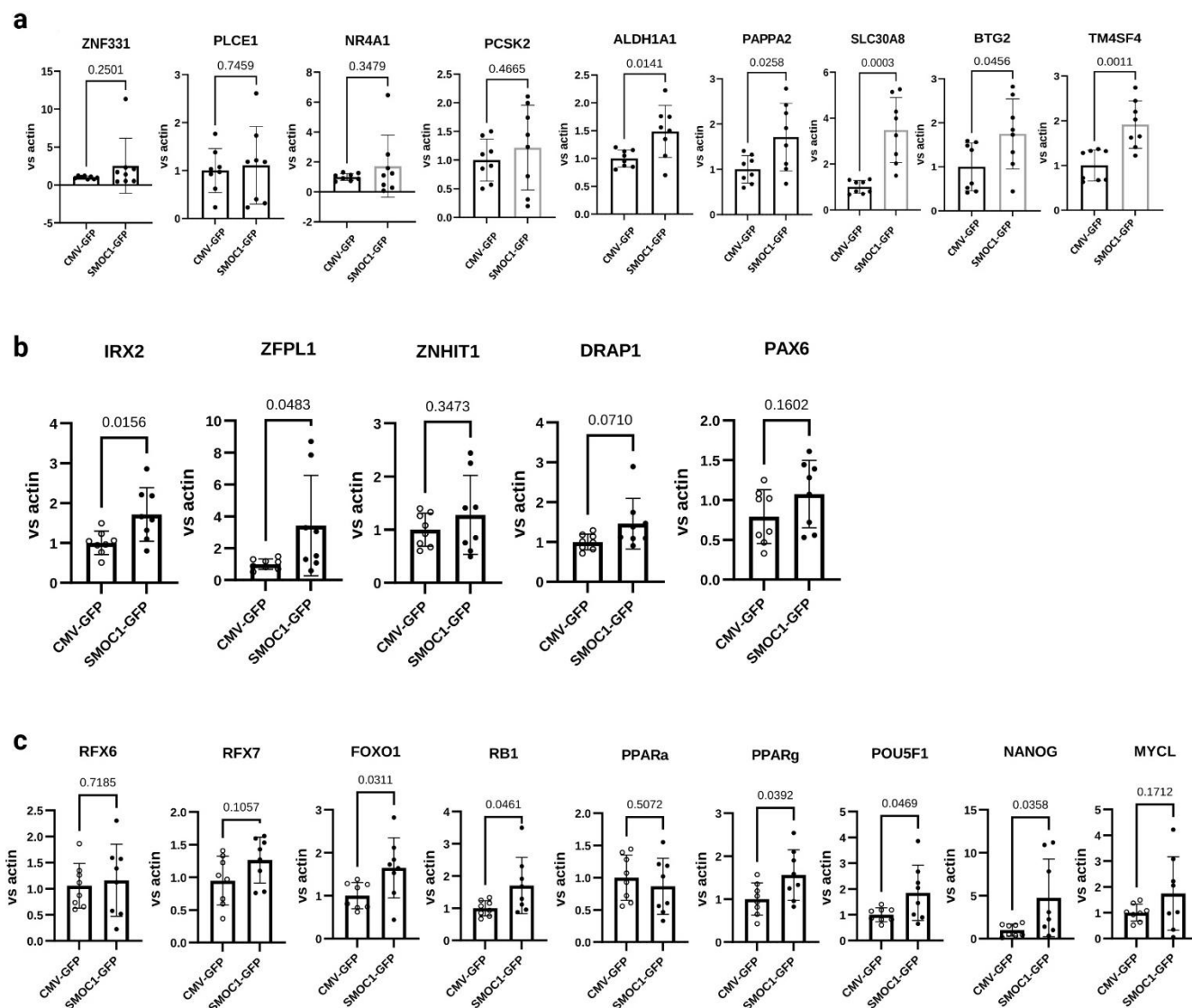


b

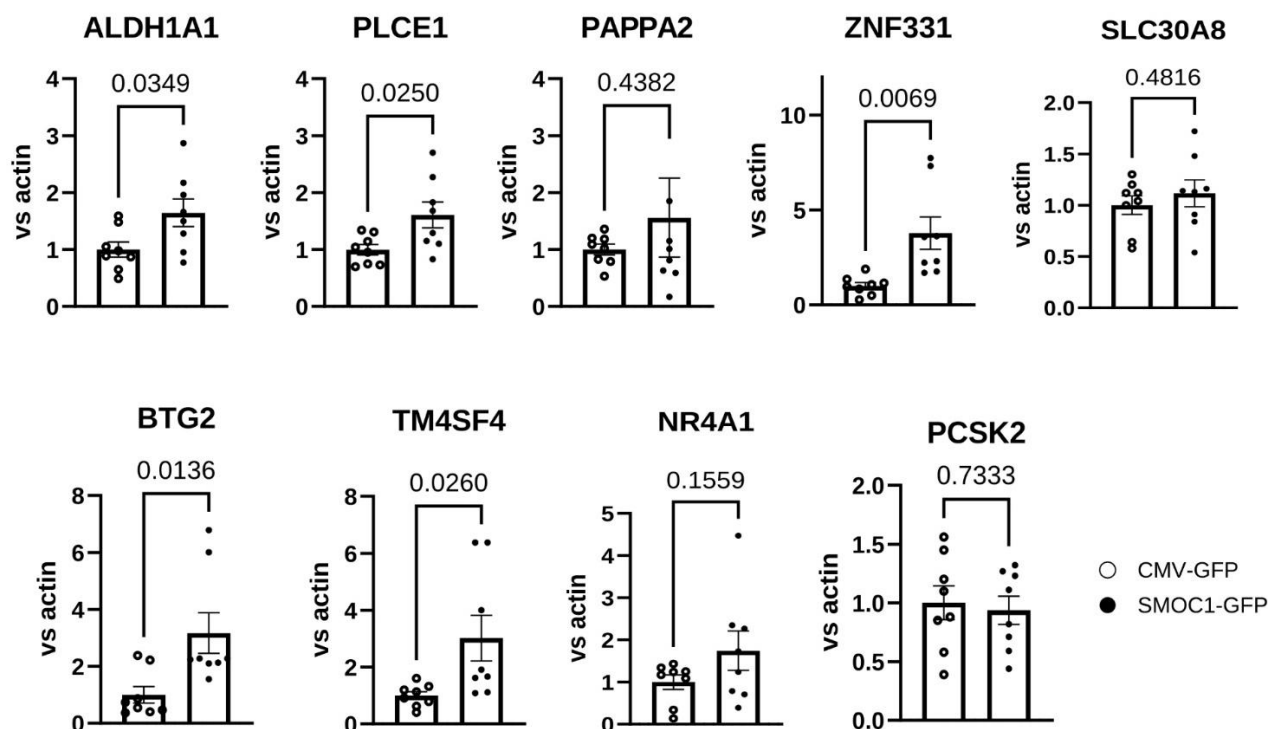




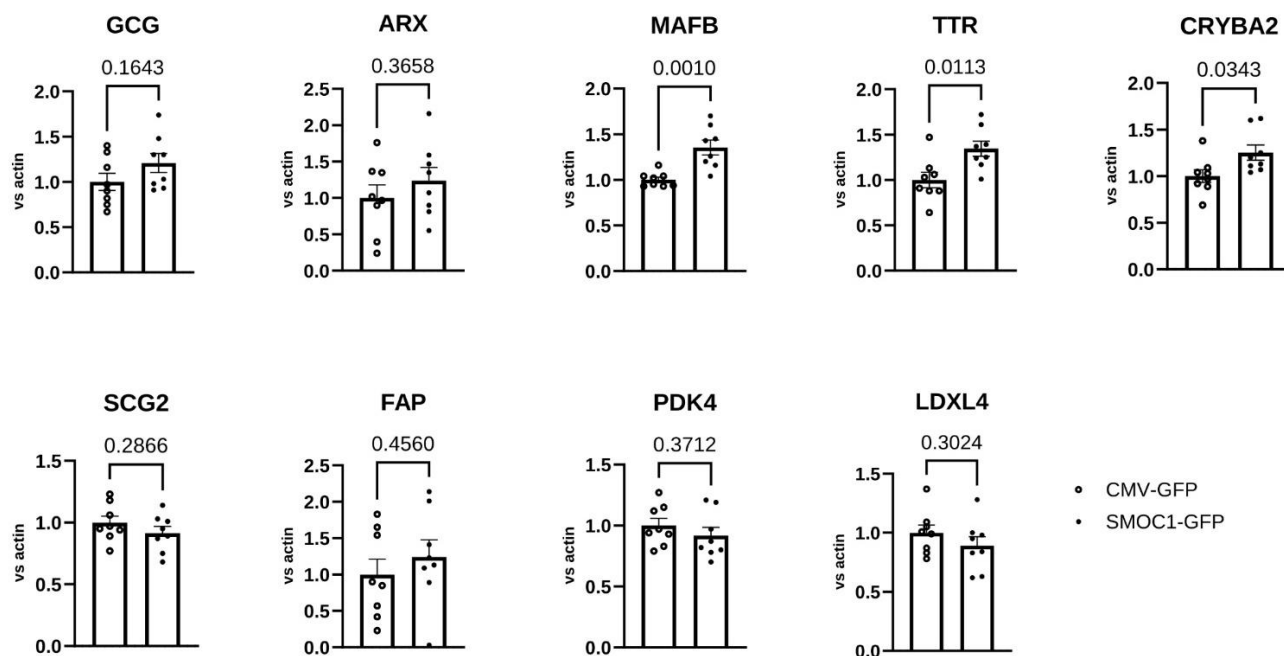
Supplementary Figure 8. Correlation between expression of common genes and *INS* and *GCG* in islets from adult human non-diabetic and T2D donors from HPAP-HIRN. a. Expression of common genes in the different β -cell to α -cell subpopulations between non-diabetic (red) and T2D (green) human islet data and analyzed statistically by Wilcoxon rank-sum test. ns: $0.05 < p < 1.0$; *: $0.01 < p < 0.05$; **: $0.001 < p < 0.01$; ***: $0.0001 < p < 0.001$; ****: $p < 0.0001$. b. Correlation between the expression of common genes in the α - to β -cell trajectories with *INS* expression in β -cell subclusters in human islets from non-diabetic (ND) or T2D donors from HPAP-HIRN dataset. c. Correlation between the expression of common genes in the α - to β -cell trajectories with *GCG* expression in α -cell subclusters in human islets from non-diabetic (ND) or T2D donors from HPAP-HIRN dataset.



Supplementary Figure 9. Effect of SMOC1 upregulation in human islet cells from non-diabetic donors. **a.** Impact of SMOC1 upregulation in the expression of common genes in the trajectory from β - to α -cells in non-diabetic human islets transduced with adenovirus-CMV-SMOC1-GFP. Impact of SMOC1 upregulation on gene signatures associated with T2D in **b.** α -cells (*IRX2*, *ZNHIT1*, *ZFPL1*, *PAX6*, and *DRAP1*) and **c.** β -cells (*RFX6*, *RFX7*, *FOXO1*, *PPAR α* , *PPAR γ* , *RB1*, *POU5F1*, *NANOG*, and *MYCL*) based on well-known processes of metabolic inflexibility, stress response and endocrine progenitors/stemness, as previously reported (41). Bars indicate mean $\pm$ SEM; n=8 different human islet preparations. P values indicated in the figure calculated using two-tailed Student's t-test



Supplementary Figure 10. Expression of common genes in the trajectory from β - to α -cells in EndoC- β H1 cells expressing SMOC1. Analysis of ZNF331, PLCE1, PAPP2, TM4SF4, NR4A1, BTG2, SLC30A8, PCSK2, and ALDH1A1 expression by qPCR in EndoC- β H1 cells transduced with adenovirus CMV-GFP or adenovirus CMV-SMOC1-GFP. Bars indicate mean $\pm$ SEM; n=8 different human islet preparations. P values indicated in the figure calculated using two-tailed Student's t-test.



Supplementary Figure 11. Expression of α -cell gene markers in EndoC- β H1 cells expressing SMOC1. Analysis of GCG, ARX, MAFB, TTR, CRYBA2, SCG2, FAP, PDK4 and LOXL4 expression by qPCR in EndoC- β H1 cells transduced with adenovirus CMV-GFP or adenovirus CMV-SMOC1-GFP. Bars indicate mean $\pm$ SEM; n=8 different human islet preparations. P values indicated in the figure calculated using two-tailed Student's t-test.