Supplementary Information

Low Expression of CCKBR in the Acinar Cells is Associated with Insufficient Starch Hydrolysis in Ruminants

Yan Cheng[1,2], Tianxi Zhang[1,2], Chao Yang[1,2], Kefyalew Gebeyew[1,2], Chengyu Ye[3], Xinxin Zhou[4], Tianqi Zhang[5], Ganyi Feng[1,2], Rui Li[1,2], Zhixiong He[1,2*], Oren Parnas[6*], Zhiliang Tan[1,2*]

Affiliations

[1]CAS Key Laboratory for Agro-Ecological Processes in Subtropical Region, National Engineering Laboratory for Pollution Control and Waste Utilization in Livestock and Poultry Production, Hunan Provincial Key Laboratory of Animal Nutritional Physiology and Metabolic Process, Institute of Subtropical Agriculture, The Chinese Academy of Sciences, Changsha, Hunan 410125, China

[2]University of Chinese Academy of Sciences, Beijing 100049, China

[3]The Department of Microbiology and Immunology, Emory University, 201 Dowman Dr, Atlanta, GA 30322

[4] LC-Bio Technology (Hanghzhou) co.ltd., Hanghzhou 310000, China

[5] College of Horticulture, Nanjing Agricultural University, Nanjing, Jiangsu 210095, China

[6] The Lautenberg Center for Immunology and Cancer Research, Faculty of Medicine, The Hebrew University of Jerusalem, Jerusalem, 91120, Israel

* Corresponding author:

Zhixiong He, Tianqi Zhang, Oren Parnas, Zhiliang Tan,

Address: Institute of Subtropical Agriculture, the Chinese Academy of Sciences, Changsha, Hunan 410125, P.R. China; College of Horticulture, Nanjing Agricultural University, Nanjing, Jiangsu 210095, China; The Lautenberg Center for Immunology and Cancer Research, Faculty of Medicine, The Hebrew University of Jerusalem, Jerusalem, 91120, Israel.

Email: zxhe@isa.ac.cn; oren.parnas@gmail.com; zltan@isa.ac.cn.

1. Materials and methods

1.1 *Sample preparation for single-cell RNA-seq*

One portion of tissue from twin female goats (n=2/group) was minced into 2-3 mm pieces and then incubated in a 37 °C water bath with multiple enzymes (1.0 mg/mL collagenase IV, 2.0 mg/mL dispase, and 50 U/mL DNase I) at 100 × rpm for 30 min. We added 10% FBS to stop the digestion process, followed by a filtration step using 70-μm and 30-μm Smart Strainer (Miltenyi Biotec, Bergisch Gladbach, Germany). Then the supernatant was discarded, and cells were collected (400 × g for 7 min at 12 °C). To remove red blood cells, dead cells, and cellular debris from the dissociated cells, we used Red Blood Cell Lysis Buffer and the MACS Dead Cell Removal Kit (Miltenyi Biotec, Bergisch Gladbach, Germany) as per the manufacturer's direction. After that, the total cell count and viability were checked using a hemocytometer (Thermo Fisher Scientific) and trypan blue (Thermo Fisher Scientific, Waltham, MA, USA), respectively. Finally, single-cell suspensions without cell debris and having greater than 90% viability were passed to the subsequent analysis. The appropriate volume was calculated for a target capture of 8000 cells, and then the cells were further diluted to a user guide concentration (700-1200 cells/μL) with 1× PBS with 0.04% BSA for 10x Genomics sequencing.

1.2 *Single-cell RNA-seq library construction and sequencing*

Briefly, to obtain single-cell gel bead-in-emulsions (GEM), we loaded the cellular suspensions onto the latest 10x Chromium™ Single Cell 3′ Solution system. Then, we used the 10x Genomics 3' Reagent Kits v3 and Agilent Bioanalyzer High Sensitivity chip to construct and check the quality of the scRNA-seq libraries, respectively. To reduce the batch effects of samples, we constructed the libraries using the exact versions of kits as per the manufacturer's protocols. The libraries were sequenced on the same Illumina NovaSeq 6000 sequencing system in a 150-bp paired-ended manner

by LC-Bio Technology Co., Ltd. (Hangzhou, China) at a minimum depth of 20,000 reads per cell.

1.3 *Single-cell RNA-seq data analysis*

The Illumina bcl2fastq software (version 5.01) were applied to convert the sequencing results in to FASTQ files. Raw data were demultiplexed, barcode and mapped to the goat reference genome (https://www.ncbi.nlm.nih.gov/assembly/GCF_001704415.1/) using the Cell Ranger package (version 3.1.0). Primary alignment to the Cell Ranger showed 95% valid barcodes and 54.30% sequencing saturation. Feature-barcode raw digital gene expression matrix (UMI counts per gene per cell) was filtered and normalized by using the R package Seurat [1] (Version 3.1.1). Overall, genes detected in fewer than one cell were filtered out, and cells were excluded if they expressed fewer than 500 detected genes, UMI counts less than Inf, and the percent of mitochondrial-DNA derived gene expression below 25%. The Doublet Finder [2] package (version 2.0.3) was also employed to filter out doublets. scRNA-seq data that met quality control criteria were used for transcriptomic analysis.

1.4 *Uniform manifold approximation and projections (UMAP) analysis of single-cell RNA-seq datasets and identification of cell clusters*

Following the removal of the low-quality cells and doublets, the “LogNormalize” method in the Seurat package was used to normalize the expression of the data. Then the principal component analysis (PCA) was performed using the “RunPCA” function based on the normalized expression value. Using a Jackstraw substitution test algorithm, we select the top ten principal components (PC) from the PCA analysis results for subsequent unsupervised clustering and cluster analysis. The process of clustering cells was performed using the “FindClusters” function with an appropriate resolution Two-dimensional visualization was obtained with a UMAP[3].

Differentially expressed genes (DEGs) or marker genes for each significant cluster were found using the Wilcoxon rank-sum test (default parameter is "bimod": likelihood-ratio test) as implemented in the "FindAllMarkers" function ($|\log2\ FC| \geq 1$ and adjusted P value ≤ 0.01). Only genes representing or expressing greater than 10% of the cells in a given cluster were considered. Manual annotation (a combination of marker genes identified from the literature, marker gene list, and gene ontology for cell types) was applied to assigned cell cluster identity. The Seurat function of FeaturePlot, VlnPlot, and DotPlot were employed to plot the expression of selected genes. We used heatmap.2 function from the gplots v3.6.1 R package with the default complete-linkage clustering algorithm to construct heat maps.

1.5 *Inference of differentiation trajectories*

Single-cell trajectories were constructed using the R package monocle [4] Version 2.4 with the default parameter settings and the partition-based graph abstraction (PAGA) [5] with an edge significance threshold of 0.6.

1.6 *Estimation of transcriptional noise*

Estimation of transcriptional noise was performed mainly based on previous work [6]. Here, we presented two measurements of transcriptional noise. Specifically, we first ensure that each cell cluster has at least 10 cells. In order to avoid the impact brought by different UMI counts, each cell was down-sampled to 100,000 UMIs so that all cells had an equal number of total UMI counts. Similarly, to account for differences in cell-type frequency, we subsampled cell numbers to maintain equal numbers of cells for the two groups. We then divided genes into 10 equally sized bins by mean expression, with the top and bottom bins excluded. The 10% of genes with the lowest coefficient of variation (CV) from each bin were kept. These genes were then used to calculate the Euclidean

distance between each cell and its corresponding cell type mean. This Euclidean distance was used as one measure of transcriptional noise for each cell. The ratios of transcriptional noise between neonatal and adult goats for each cell type were subsequently calculated. We also calculated the average Euclidean distances for each goat. Alternatively, we used the 1–Spearman correlation coefficient as the second measure. Spearman's correlation coefficients were calculated using the subsampled gene expression matrices between each cell type and age group (neonatal and adult stages). We then used Wilcoxon’s rank sum test to statistically evaluate the relationship between transcriptional noise and age for each cell type. The Bonferroni–Hochberg method was used to correct the p values.

2. Supplementary Figures

Figure. S1. Transcriptomic characteristics of the goat Pancreas. (A) Split views showed the biological duplicate effect on NG and AG groups. (B) UAMP clustering of 35,449 cells isolated from pancreas of neonatal (NG group) and adult goats (AG group). Cells are colored according to the cluster. Each dot represents a single cell. (C) Expression level of marker of acinar cell in different subsets. (D) Expression level of marker of ductal cell in different subsets.

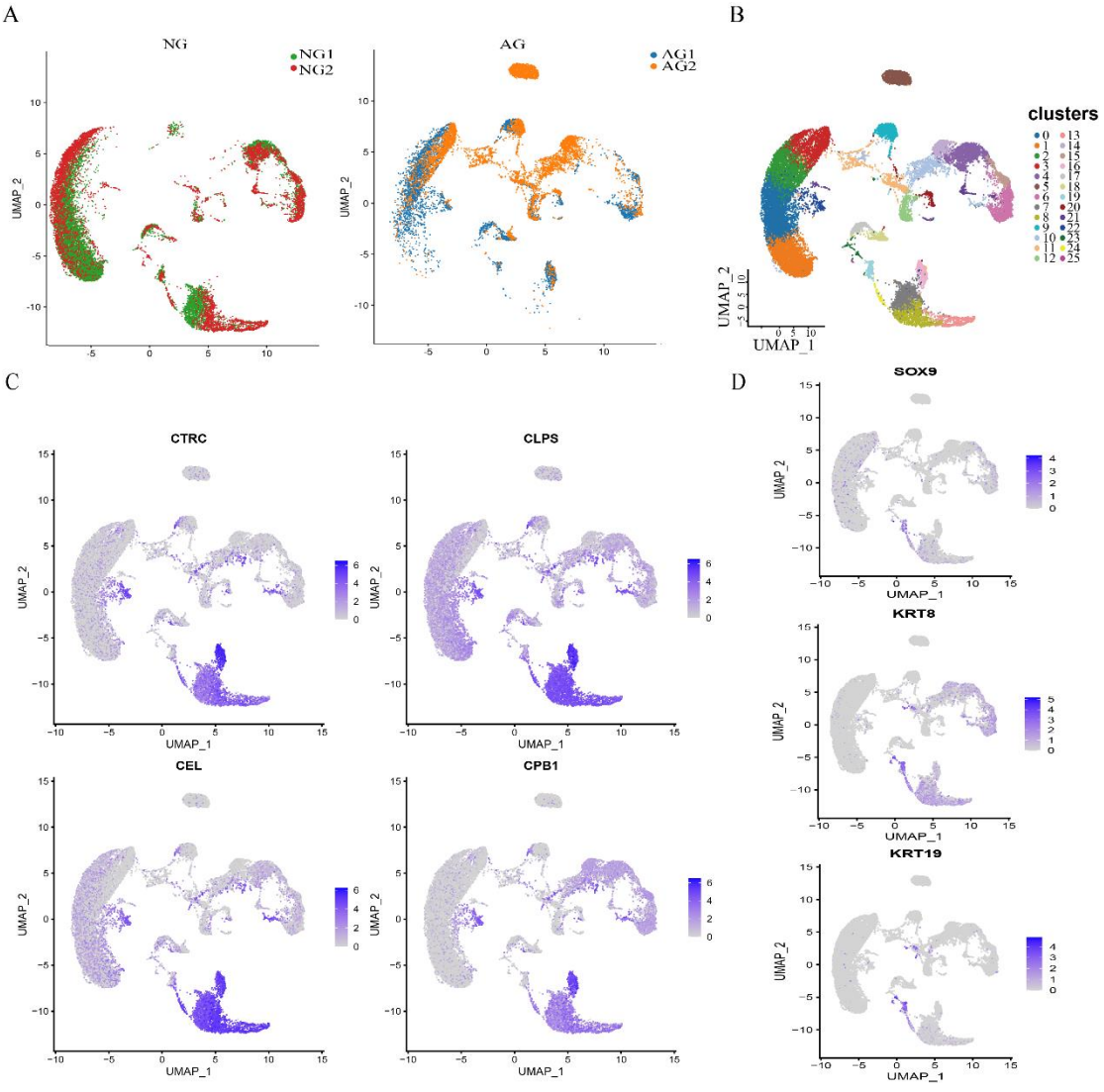


Figure. S2. Heterogeneity of AC cells reveal the differential metabolic responses. UAMP clustering of 3483 pancreatic AC cells isolated from NG and AG samples. Cells are colored based on the clusters (A) and samples (B), respectively. (C) Expression level of markers of acinar cell subtype in different subsets. (D) List of the significant GSEA terms. (E) Expression of representative candidate genes of acinar cell subtype along pseudotime.

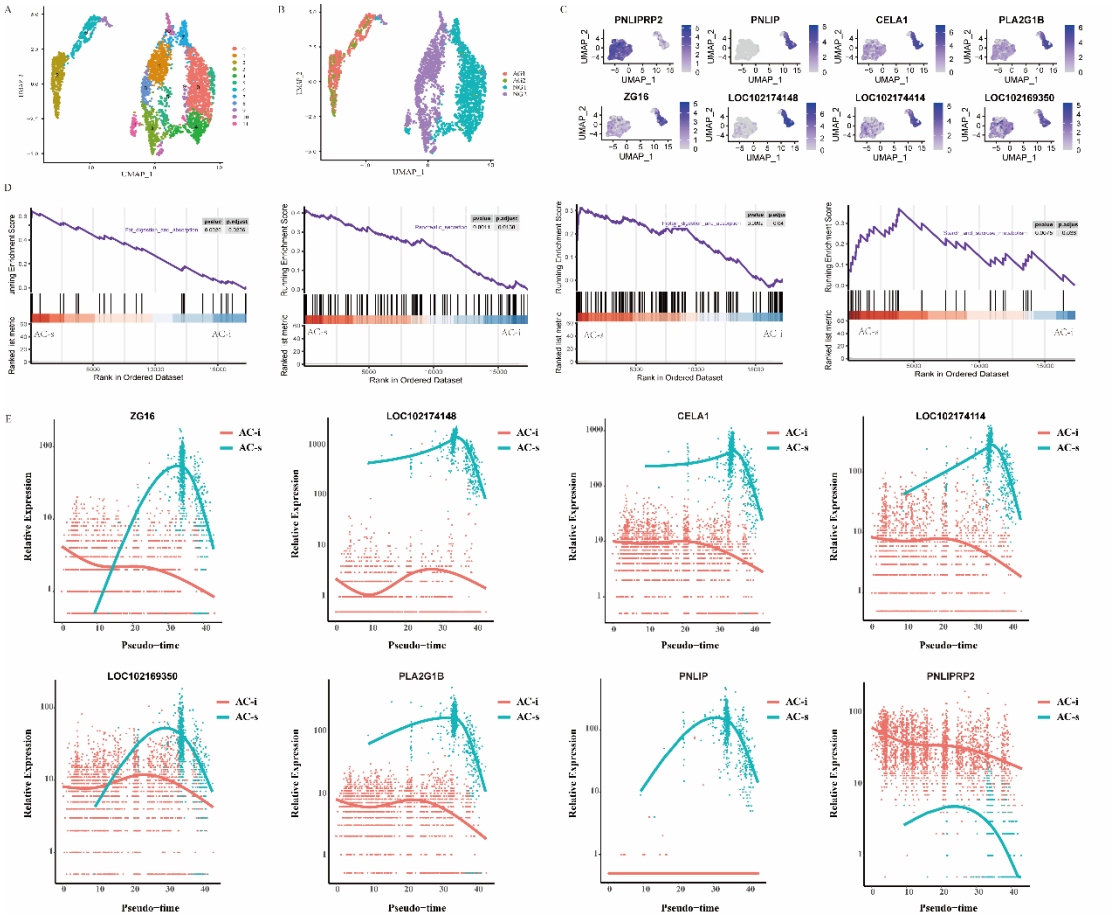


Figure. S3. Cell-to-cell signaling networks and gene expression alterations of pancreatic endocrine cells of adult goats. Cell-to-cell signaling networks in neonatal goats (A) and adult goats (B). Networks depicting cell types as nodes and interactions as edges. Edge thickness is proportional to the number of interactions between the connecting types. (B) Heat map depicting the number of all possible interactions between the clusters analyzed in neonatal goats (C) and adult goats (D). Volcano map showing the number of DE genes between cells from AG and NG samples in β (E) and δ (F), with red and blue colors corresponding to up-and-downregulated genes in AG samples, respectively. Non-differential genes are shown in grey color.

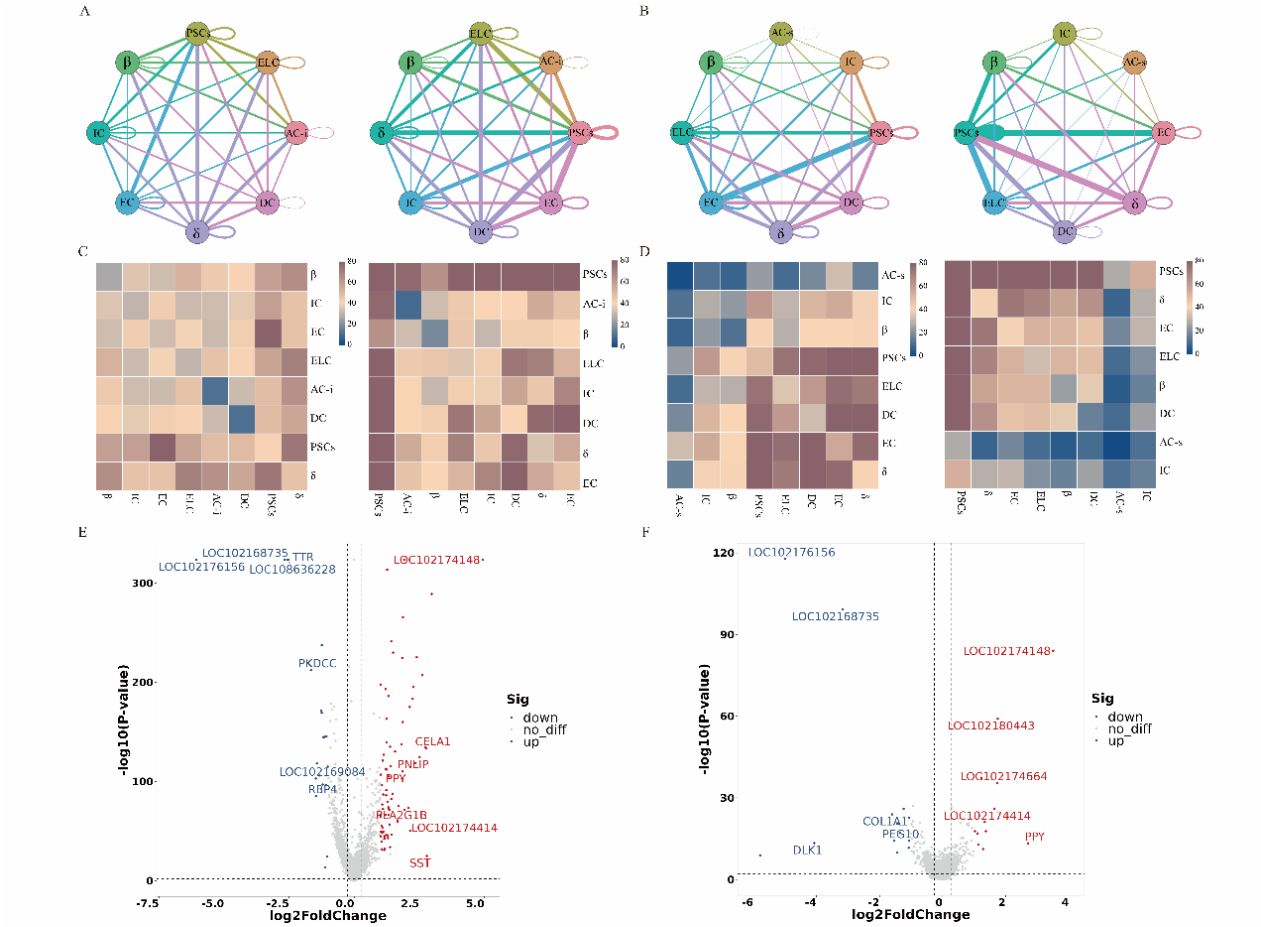
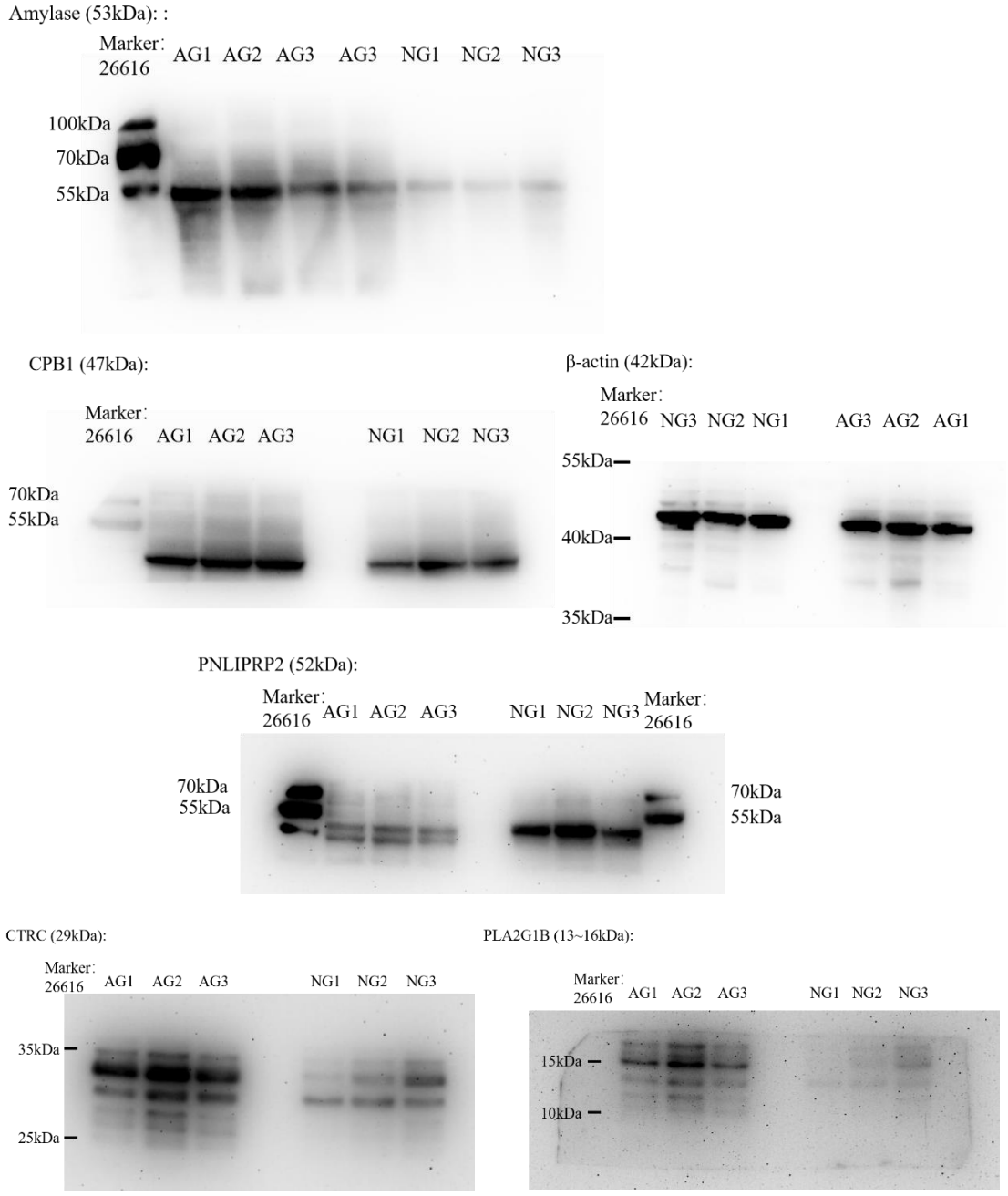


Figure. S4. Western blot analysis of PNLIPRP2, CTRC, CPB1, PLA2G1B, Amylase and β-Actin in pancreatic tissue. Note: 1) the PVDF membrane was cut according to the size of the target protein. Then, the membranes were incubated with primary antibodies in TBST (including 5% BSA). The complete picture is provided here without cropping as it is. 2) The protein bands (40~10kDa) did not show color, so we labeled the protein size on the PVDF membrane based on the color on the membrane.



3. Supplementary Tables

Table. S1. Overview of key genes from scRNA-seq data, divided by each cell compartment, that exhibit relatively high or specific expression in each cell types used for characterization annotation, related to Figure 1.

Compartment	Cell type	Key genes	Reference
Islet endocrine cell types	β	*INS*	Lawlor et al., 2017 [7], Li et al., 2016a [8]
	δ	*GCG*	
	ELC	*SST, PPY*	
exocrine cell types	AC	*CTRC, CLPS, CEL, CPB1*	Whitcomb & Lowe, 2007 [9]
	DC	*KRT8, KRT19, SOX9*	Lawlor et al., 2017 [7], Li et al., 2016a [8], Prévot et al., 2012 [10]
Nonpancreatic cell types	EC	*PECAM1, VWF*	Segerstolpe et al., 2016 [11]
	activated PSCs	*PDGFRA*, *COL1A2*	Tang et al., 2022 [12], Baron et al., 2016 [13]
	B	*MS4A1, CD19*, *CD79A*	Adams et al., 2009 [14]
	T	*CD3E*, *CD3D*	Tang et al., 2022 [12], de Saint et al., 2004 [15]
	macrophages	*C1QA, C1QB*, *C1QC*	Tang et al., 2022 [12], Li et al., 2023[16]
	granulocytes	*S100A9*, *S100A8*	Scott et al., 2020 [17]

149 **Table. S2.** List of specific and highly expressed pancreatic digestive enzymes-related genes identified in acinar cells of goats.

Item	Description	Gene transcript	Proportion of AC transcripts
Digestive zymogen			
LOC102176156	chymotrypsinogen B	1652	2010/10000
LOC102174148	chymotrypsinogen A	217	
LOC102173875	chymotrypsinogen B	128	
ZG16	zymogen granule protein 16	10	
CYM	prochymosin	3	
Protease			
PRSS2	protease, serine 2	365	957/10000
LOC1021783207	carboxypeptidase A1	243	
CTRC	Chymotrypsin C	90	
CELA1	chymotrypsin like elastase family member 1	75	
CPB1	carboxypeptidase B1	68	

LOC102179184	cationic trypsin	55	
LOC102174414	chymotrypsin-like elastase family member 2A	50	
LOC102178719	cationic trypsin	8	
PRSS23	protease, serine 23	1	
CPM	carboxypeptidase M	1	
Lipase			
CLPS	colipase	327	526/10000
CEL	carboxyl ester lipase	114	
PLA2G1B	phospholipase A2 group 1B	33	
PNLIPRP2	pancreatic lipase related protein 2	30	
PNLIP	pancreatic lipase	23	
Amylase			
LOC102169350	alpha-amylase 2B	15	15/10000

Table. S3. Nutrient contents of milk in the NG group.

Milk feed composition, %	Content
Total Solids	23.50
Protein	15.67
Fat	4.15
Lactose	3.48
Urea, mg/dL	28.70

Table. S4. Nutrient contents of solid feed in the AG group.

TMR feed composition, %	Content
Ricestraw	70
Soybeans	15
Corn	8.2
CaCO3	2.9
CaH2PO4	0.1
Fat	0.3
Nacl	1
Premix	0.5
TMR nutritional level	
DM (%)	96.2
ASH (%)	12.71
CP (%)	10.8
NDF (%)	49.8
ADF (%)	28.4
Starch (%)	11.5
Energy (MJ/Kg)	16.8

Note: Premix (per kilogram) contains 6.9 g Fe, 4.4 g Cu, 1.1 g Co, 11.2 g I, 11.0 g Mn, 4.6 g Zn, 0.3 g Se, 104.2 g Mg, 10,000,000 IU vitamin A, 16,000,000 IU vitamin D, 12,000 IU vitamin E, 400 g NaHCO3, and 400.9 g carrier.

Table. S5. Overview of antibody information in this study.

Application	Antibody name	Company	Batch	Dilution
IHC	Amylase	Santa	sc-46657	1: 1000
	CPB1	abcam	ab153716	1: 1000
	PNLIPRP2	Proteintech	26218-1-AP	1: 400
	CCKAR	Proteintech	16550-1-AP	1: 200
	CCKBR	Proteintech	16549-1-AP	1: 50
	VIPR1	Proteintech	14878-1-AP	1: 200
	VIPR2	Bioss	bs-0197R	1: 400
	SCTR	Proteintech	14172-1-AP	1: 400
	Secondary antibody	Sparkjade	EF0002	1: 200
WB	PNLIPRP2	Proteintech	26218-1-AP	1: 1000
	CTRC	Bioss	bs-13948R	1: 1000
	CPB1	Abcam	ab153716	1: 1000
	PLA2G1B	Proteintech	66397-1-IG	1: 500
	Amylase	Santa	sc-46657	1: 1000
	β-actin	Bioss	bs-0061R	1: 1000
IF	CPA1	Proteintech	15836-1-AP	1: 100
	CCKBR	Proteintech	16549-1-AP	1: 100
	CCK	Proteintech	13074-2-AP	1: 400
	Secondary antibody	YEASEN	33107ES60	1: 200

References

[1] Y.H. Hao, S. Hao, E. Andersen-Nissen, W.M. Mauck, S.W. Zheng, A. Butler, M.J. Lee, A.J. Wilk, C. Darby, M. Zager, P. Hoffman, M. Stoeckius, E. Papalexi, E.P. Mimitou, J. Jain, A. Srivastava, T. Stuart, L.M. Fleming, B. Yeung, A.J. Rogers, J.M. McElrath, C.A. Blish, R. Gottardo, P. Smibert, R. Satija, Integrated analysis of multimodal single-cell data, Cell 184(13) (2021) 3573-+.

[2] C.S. McGinnis, L.M. Murrow, Z.J. Gartner, DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors, Cell Systems 8(4) (2019) 329-+.

[3] E. Becht, L. McInnes, J. Healy, C.A. Dutertre, I.W.H. Kwok, L.G. Ng, F. Ginhoux, E.W. Newell, Dimensionality reduction for visualizing single-cell data using UMAP, Nature Biotechnology 37(1) (2019) 38-+.

[4] J.Y. Cao, M. Spielmann, X.J. Qiu, X.F. Huang, D.M. Ibrahim, A.J. Hill, F. Zhang, S. Mundlos, L. Christiansen, F.J. Steemers, C. Trapnell, J. Shendure, The single-cell transcriptional landscape of mammalian organogenesis, Nature 566(7745) (2019) 496-+.

[5] F.A. Wolf, F.K. Hamey, M. Plass, J. Solana, J.S. Dahlin, B. Göttgens, N. Rajewsky, L. Simon, F.J. Theis, PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells, Genome Biology 20 (2019).

[6] M. Enge, E. Arda, M. Mignardi, J. Beausang, R. Bottino, S.K. Kim, S.R. Quake, Single-Cell Analysis of Human Pancreas Reveals Transcriptional Signatures of Aging and Somatic Mutation Patterns, Cell 171(2) (2017) 321-+.

[7] N. Lawlor, J. George, M. Bolisetty, R. Kursawe, L.L. Sun, V. Sivakamasundari, I. Kycia, P. Robson, M.L. Stitzel, Single-cell transcriptomes identify human islet cell signatures and reveal cell-type specific expression changes in type 2 diabetes, Genome Research 27(2) (2017) 208-222.

[8] J. Li, J. Klughammer, M. Farlik, T. Penz, A. Spittler, C. Barbieux, E. Berishvili, C. Bock, S. Kubicek, Single-cell transcriptomes reveal characteristic features of human pancreatic islet cell types, Embo Reports 17(2) (2016) 178-187.

[9] D.C. Whitcomb, M.E. Lowe, Human pancreatic digestive enzymes, Digestive Diseases and Sciences 52(1) (2007) 1-17.

[10] P.P. Prévot, A. Simion, A. Grimont, M. Colletti, A. Khalaileh, G. Van den Steen, C. Sempoux, X.B. Xu, V. Roelants, J. Hald, L. Bertrand, H. Heimberg, S.F. Konieczny, Y. Dor, F.P. Lemaigre, P. Jacquemin, Role of the ductal transcription factors HNF6 and Sox9 in pancreatic acinar-to-ductal metaplasia, Gut 61(12) (2012) 1723-1732.

[11] Å. Segerstolpe, A. Palasantza, P. Eliasson, E.M. Andersson, A.C. Andréasson, X.Y. Sun, S. Picelli, A. Sabirsh, M. Clausen, M.K. Bjursell, D.M. Smith, M. Kasper, C. Ämmälä, R. Sandberg, Single-Cell Transcriptome Profiling of Human Pancreatic Islets in Health and Type 2 Diabetes, Cell Metabolism 24(4) (2016) 593-607.

[12] W.J. Tang, Y.F. Zhong, Y.S. Wei, Z.X. Deng, J.D. Mao, J.L. Liu, T.G. Valencak, J.X. Liu, H.P. Xu, H.F. Wang, Ileum tissue single-cell mRNA sequencing elucidates the cellular architecture of pathophysiological changes associated with weaning in piglets, Bmc Biology 20(1) (2022).

[13] M. Baron, A. Veres, S.L. Wolock, A.L. Faust, R. Gaujoux, A. Vetere, J.H. Ryu, B.K. Wagner, S.S. Shen-Orr, A.M. Klein, D.A. Melton, I. Yanai, A Single-Cell Transcriptomic Map of the Human and Mouse Pancreas Reveals Inter- and Intra-cell Population Structure, Cell Systems 3(4) (2016) 346-+.

[14] H. Adams, P. Liebisch, P. Schmid, S. Dirnhofer, A. Tzankov, Diagnostic Utility of the B-cell

Lineage Markers CD20, CD79a, PAX5. and CD19 in Paraffin-embedded Tissues From Lymphoid Neoplasms, Applied Immunohistochemistry & Molecular Morphology 17(2) (2009) 96-101.
[15] G. de Saint Basile, F. Geissmann, E. Flori, B. Uring-Lambert, C. Soudais, M. Cavazzana-Calvo, A. Durandy, N. Jabado, A. Fischer, F. Le Deist, Severe combined immunodeficiency caused by deficiency in either the δ or the ε subunit of CD3, Journal of Clinical Investigation 114(10) (2004) 1512-1517.
[16] H.M. Li, Y. Miao, L.Q. Zhong, S.J. Feng, Y. Xu, L. Tang, C. Wu, X.Z. Zhang, L. Gu, H.Y. Diao, H.Y. Wang, Z.S. Wen, M.L. Yang, Identification of TREM2-positive tumor-associated macrophages in esophageal squamous cell carcinoma: implication for poor prognosis and immunotherapy modulation, Frontiers in Immunology 14 (2023).
[17] N.R. Scott, R.V. Swanson, N. Al-Hammadi, R. Domingo-Gonzalez, J. Rangel-Moreno, B.A. Kriel, A.N. Bucsan, S. Das, M. Ahmed, S. Mehra, P. Treerat, A. Cruz-Lagunas, L. Jimenez-Alvarez, M. Muñoz-Torrico, K. Bobadilla-Lozoya, T. Vogl, G. Walzl, N. du Plessis, D. Kaushal, T.J. Scriba, J. Zú ñiga, S.A. Khader, S100A8/A9 regulates CD11b expression and neutrophil recruitment during chronic tuberculosis, Journal of Clinical Investigation 130(6) (2020) 3098-3112.