Single Cell RNA-Seq Analysis of Pancreatic Cells

Yu Zhong, ZhiYu Zhang, JianFeng Ke, HuiSiYu Yu

Introduction

The human pancreas is a composite organ with both exocrine and endocrine functions[1]. The exocrine compartment consists of two cell types: acinar cells that produce digestive enzymes, and ductal cells that secrete bicarbonate and channels these enzymes into the duodenum. The endocrine portion helps maintain glucose homeostasis and comprises five major cell types found in islets of Langerhans: alpha cells, beta cells, delta cells, gamma (PP) cells, and epsilon cells[2]. In their study *A Single-Cell Transcriptomic Map of the Human and Mouse Pancreas Reveals Inter- and Intra-cell Population Structure*, Baron et al performed single cell RNA sequencing using the inDrop platform in a set of human donor pancreatic cells from four subjects and two strains of mice in order to describe pancreatic cell types at a high resolution[3].

The goal of this project is to replicate their primary findings using current analytical methodology and software packages by analyzing sequencing data from one of the four human donors (donor 2). We utilized a graph-based clustering method different from the original authors and identified six major pancreatic cell types characterized by different transcriptional expression profiles, four of which confirmed by gene set enrichment analysis results.

Methods

Datasets

Single Cell RNA-Seq datasets comprising 3 replicates sequenced from the 51-year-old female donor (donor 2, GSM2230758) were obtained from the GEO database (GEO Accession viewer (nih.gov)). The barcoding scheme of the inDrop sequencing library was designed to be 19 nucleotides long, followed by an unique molecular identifier (UMI) with six bases. The barcode for each cell and UMIs for each transcript had been processed beforehand, such that read1 files only contain barcoding sequences (barcodes and UMIs).

Barcode preprocessing

The number of reads by barcodes (bc1 and bc2) was calculated for each run, and statistical visualization was performed in R. Infrequent reads were eliminated by filtering out barcodes with a number of reads less than 10,000 across all replicates. The remaining barcodes were used as a whitelist for the purpose of transcript quantification.

Transcript quantification

Using the latest transcriptome reference and annotations (GRCh38.p13), transcriptome indices were built with comprehensive decoy sequences of the genome by using the mapping-based mode in Salmon[4]. A cell-by-gene count matrix was generated using the Alevin module which worked under the same indexing scheme.

Data preprocessing

The UMI count matrix was filtered with two criteria: 1) having at least 200 non-zero genes per cell, 2) having a minimum of three non-zero count cells per gene before entering downstream analysis. The Seurat package in R was used for exploring the quality control metrics, for which we further filtered out cells with unique features counts over 4000 or less than 200, as well as cells having more than 10% mitochondrial counts. We normalized the feature expression measurements for each cell by the total expression, multiplied by a scale factor of 10,000, and log-transformed the result. The 2,000 most highly variable features were selected, and clusters of cell type subpopulations were identified using a graph-based clustering approach.

Marker genes analysis

Differential expression analysis was performed to examine maker genes for each cluster. Only positive maker genes with at least 25% coverage in either group of cells and passed the threshold of log2FC greater than 0.25 were reported as valid cluster biomarkers. Cell type annotations and literature reference were identified using CellMarker [ref]. UMAP (Uniform Manifold Approximation and Projection), was used as a dimension reduction technique to visualize the clustering results.

Gene set enrichment analysis (GSEA)

Gene set enrichment analysis was performed using the enrichR (version 3.0)[5] R package, an R interface of the Enrichr database (R version 4.0.3). Four annotation databases used for the analysis were: GO Molecular Function 2018, GO Cellular Component 2018, GO Biological Process 2018, and KEGG 2019 Human. Enrichment results were filtered using adjusted P-value < 0.1, and the top 5 enriched terms ranked by combined score were selected for each annotation source. The combined score is a composite evaluation metric computed by multiplying the log of the p-value computed with the Fisher exact test by a corrected z-score[5].

Results

Barcode preprocessing is critical for the single cell sequencing datasets due to the presence of unknown technical bias derived from library preparation or sequencing errors. There were in total 1,267,907, 1,300,473 and 1,197,807 unique UMIs for the three replicates SRR3879604, SRR3879605, and SRR3879606, respectively. The cumulative distributions of log barcode counts demonstrated that reads were distributed similarly from each other across different replicates (Figure 1). The majority of barcodes had extremely low expression levels (Figure 2A) as we expected. After removing barcodes with low quality (See Methods), 4809 unique UMIs (Figure 2B) were left to be served as a whitelist for cell detection and cellular barcode sequence correction.

To evaluate expression levels across different cells, the UMIs count matrix was estimated and quantified based on the barcodes derived from above. Table 1 highlights that a noticeable proportion of reads were detected as noise, and mapping rate was low (about 30%).

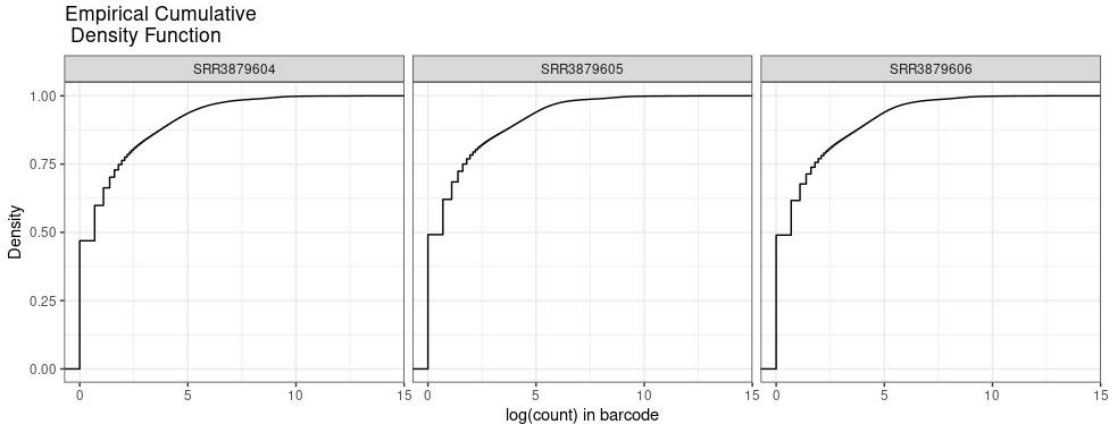

Figure 1. Cumulative distribution of log(barcode) over three replicates.

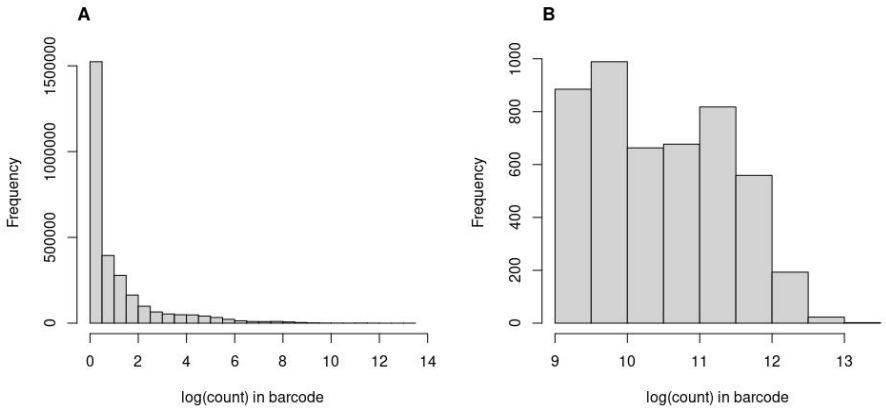

Figure 2. The histograms of barcode counts. A. The frequency distribution of log(count) before filtering; B. The frequency distribution of log(count) after filtering.

Table 1. A summary of the mapping statistics from the Alevin

Class	Count/Rate
total_reads	1,324,837,961
reads_with_N	67,930
noisy_cb_reads	554,924,208
noisy_umi_reads	35,254
used_reads	769,775,315
mapping_rate	30.35%
total_cbs	4,251,176
used_cbs	86,516
mean_umis_per_cell	2,572
mean_genes_per_cell	1,239

Exploratory Data Analysis can reveal biological patterns within the massive datasets. There were in total 60232 genes and 4798 cells in the unfiltered count matrix. After filtering out low-quality cells as well as features, 25950 genes and 1678 cells remained. Eight clusters were identified based on the top 2000 features of cells (**Figure 3**), suggesting that each cluster may represent a specific set of cellular activities and functions.

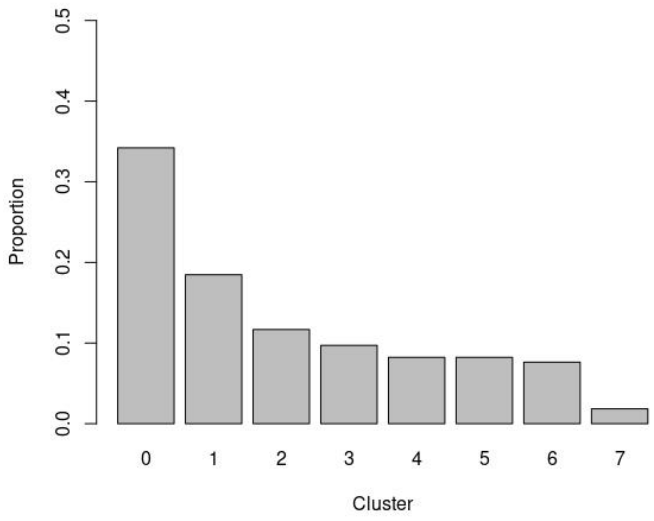

Figure 3. The proportion of cells across different clusters.

To further understand biological patterns embedded in different clusters, we identified marker genes derived from differential expression analysis. Provided with annotations of marker genes, it is evident that the cells clustered together by cell types (**Figure 4**), despite that a small proportion of cells mixed up with each other. Expression profiles of the top ten maker genes across different clusters was plotted in **Figure 5**, which showed significant differences in gene expression across different cell types.

It was possible that there were other differentially expressed genes in each cluster which can be used to discriminate cell types besides the existing marker genes, therefore we selected the single most differentially expressed gene from each cluster (**Figure 6**). Given that both GCG and PPY were already identified as maker genes according to the original study[ref], the rest of genes were considered as novel maker genes capable of identifying cell types.

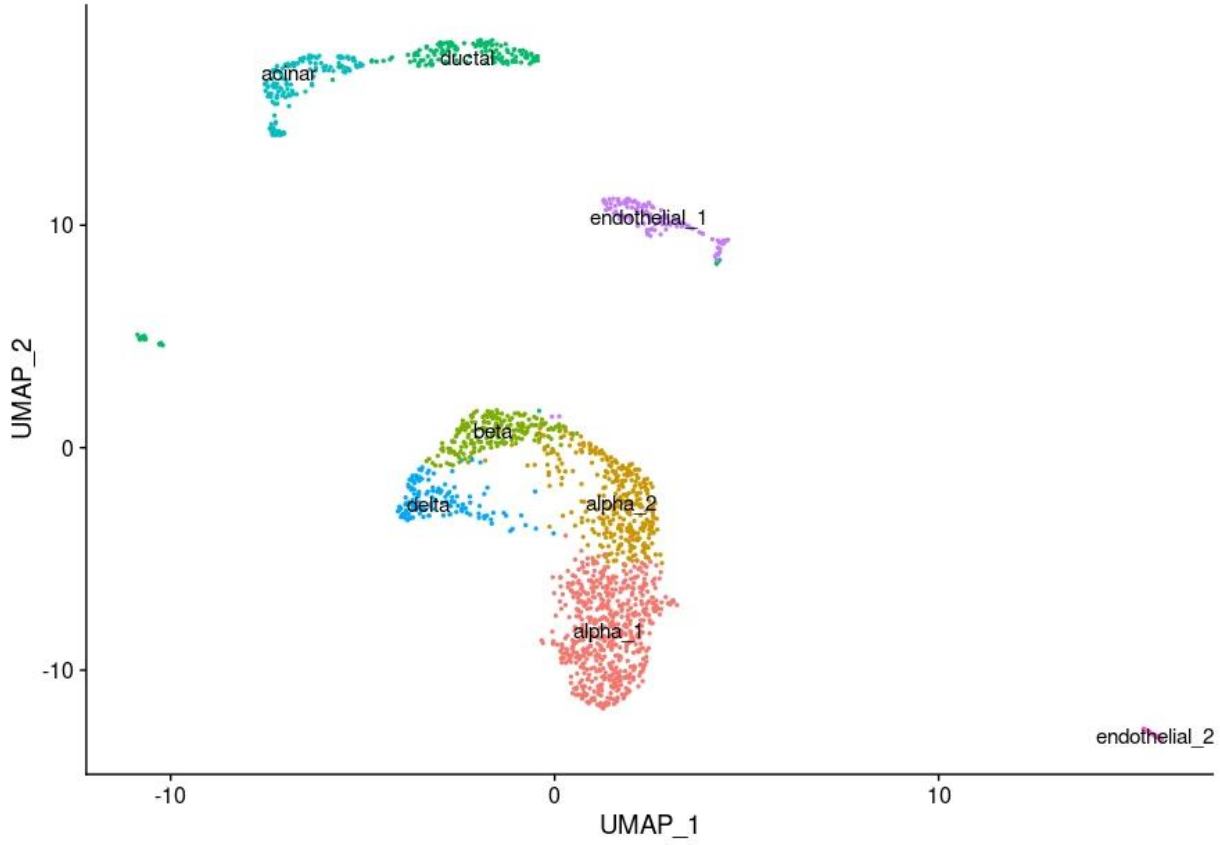

Figure 4. Unsupervised clustering results by cell types

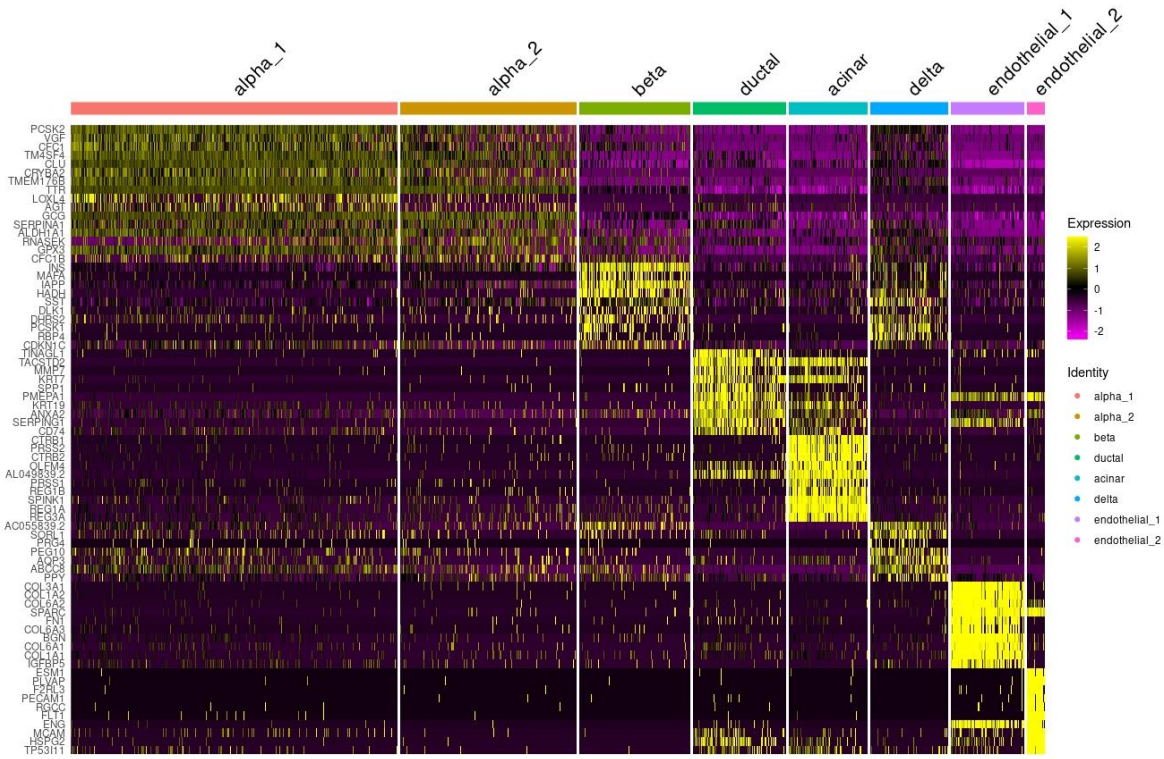

Figure 5. Unsupervised clustering of the top ten marker genes.

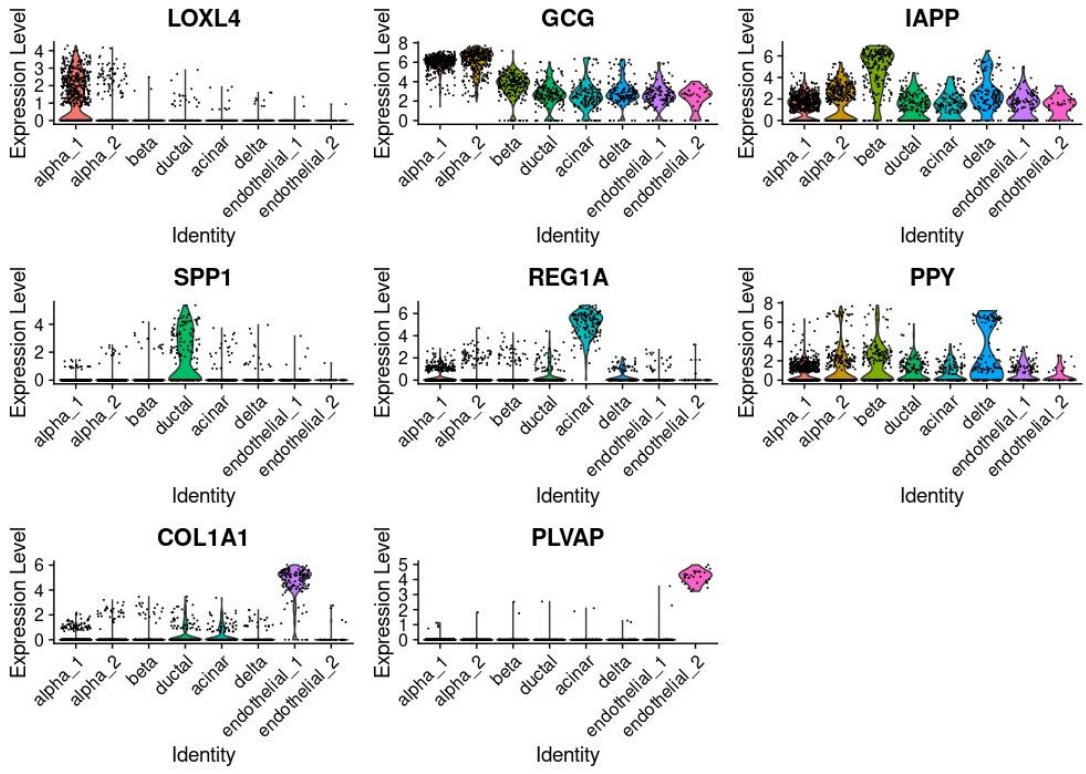

Figure 6. The top one maker genes across different cell types.

A summary of enrichment analysis results are listed in Table 2. Complete GSEA results are reported in Appendix Table S1.

Table 2. GSEA summary

cluster	cell type	# marker genes	identity confirmed?	brief description
0	alpha	634	No	intracellular membrane trafficking; cell signaling; lysosomal catabolism
1	alpha	24	No	mixed metabolic enzymes; blood-clotting-related processes; neural signaling and CNS related processes
2	beta	67	Ambiguous	translational activities; protein targeting; insulin and diabetes related processes; viral contamination
3	ductal	676	Yes	cell-cell/cell-matrix junctions; secretory granules and exocytosis
4	acinar	726	Yes	cell-cell/cell-matrix junctions; protein modifications and folding; unfolded protein response (UPR); secretion
5	delta	301	Yes	neurotransmitting activities; regulation of peptide hormone secretion; potassium ion channels and calcium ion dependent activities
6	endothelial	790	Ambiguous	binding of multiple growth factor receptors; cytoskeleton and extracellular matrix organization
7	endothelial	565	Yes	angiogenesis, cell adhesion, TGF-β and VEGF receptor binding, leukocyte recruitment

Discussion

Single cell sequencing libraries sometimes contain unwanted variations and contaminants that lead to technical biases. As we expected, the majority of barcode sequences were potentially results of PCR amplification errors. The low mapping rate (about 30%) during transcript quantification also demonstrated that our library quality was less confident for the downstream analysis.

Unsupervised clustering methods were used to classify cells types. We identified in total eight cell clusters in pancreatic tissue, whereas some of them were classified into the same group, suggesting that the graph-based method implemented in this project is less robust for performing clustering analysis than the iterative hierarchical clustering method used by Baron et al. In addition, we observed that the clustering results were sensitive to the way we filtered cells in the UMI count matrix, indicating that our results are somewhat biased compared with the original study.

Comparing with cell types identified by Baron et al, the relatively rare cell types of epsilon-cells, T cells, vascular cells, Schwann cells, quiescent and activated stellate cells, and four types of immune cells were missing from our clustering results, while most of the major endocrine and exocrine cell types were successfully identified. In order to confirm cell type classification results, gene set enrichment analysis was performed for each cluster, and enriched terms were closely examined for each by comparing with existing literatures on the identified cell types.

Cluster 0-2

Alpha cells secrete glucagon, a hormone that raises blood glucose levels by stimulating hepatic glucose synthesis and mobilization[6]; beta cells secretes insulin, which lowers blood glucose levels by stimulating cells to take up glucose from the bloodstream[[7]. While these two hormones have opposing effects on blood sugar levels and on the metabolism of nutrients, alpha and beta cells share similar glucose sensing mechanisms and are both regulated by insulin and glucagon simultaneously, thereby creating difficulty for confirming cell type classification through gene set enrichment analysis (GSEA). Based on GSEA results, cluster 0 is possibly a misclassification of beta cell as alpha cell because of the term insulin secretion and several lysosomal catabolism and glycolysis related terms, which are some of the GO terms with highest combined score.

Classification of cluster 1 is also unreliable, as only 24 marker genes have been identified and the most significant terms consist of a mix of several metabolic enzymes, blood clotting, and neural signaling related processes. Cluster 2 is classified as beta cells. The most prominent enrichment terms with extremely high combined scores (~400 to ~40000) relate to translational activities and protein targeting. These characteristics coincide with the signature function of pancreatic acinar cells, which are known to have high rates of protein synthesis and export[8]. Insulin secretion and diabetes related

processes are also captured in KEGG human annotation terms, and the presence of the term viral gene expression (combined score=26439.8) suggests viral contamination. Therefore, cluster 2 is considered a hybrid and contaminated cluster of beta cells and acinar cells.

Cluster 3&4

Acinar and ductal cells of the exocrine pancreas form a close functional unit. Acinar cell synthesizes, stores, and secretes digestive enzymes, while ductal cells provide a structural framework for delivering these enzymes to the duodenum[8,9], and secrete bicarbonate that neutralizes stomach acidity[10]. Production of digestive enzymes in acinar cells requires protein synthesis, trafficking, and processing[8]. Cluster 3 is classified as ductal cell, which is supported by enrichment terms involving structural components such as cell-cell/cell-matrix junctions as well as terms that relate to intercellular transport and communication. Cluster 4 is classified as acinar cells, supported by the enrichment terms that relate to protein modifications and folding, unfolded protein response (UPR) and secretion. However, protein synthesis related processes are absent from this cluster.

Cluster 5

Delta cells are known to secrete somatostatin, a hormone that regulates secretion of many other hormones including insulin and glucagon, and is involved in the exocrine, endocrine, and CNS systems[11]. Secretion of somatostatin is triggered by changes in blood glucose levels detected via K_{ATP} channels[12] as well as action potential firing and subsequent Ca^{2+} influx [13]. GSEA results support the classification of cluster 5 as delta cells. Although the main function of “somatostatin secretion” is absent from top ranked enrichment terms, physiological mechanisms that initiate or dependent upon somatostatin secretion are indeed reflected in terms involving neurotransmitting activities, regulation of peptide hormone secretion, and potassium ion channels and calcium ion dependent activities.

Cluster 6&7

The endocrine pancreas is highly vascularized[14], where blood vessels are lined by endothelial cells that regulate exchanges between the bloodstream and the surrounding tissues[15]. Previous studies have demonstrated that pancreatic endothelial cells play an important role in beta cell differentiation and proliferation[14], and signals from endothelial cells organize the growth and development of connective tissue cells that form the surrounding layers of the blood-vessel wall[15]. Cluster 6 and 7 are classified as endothelial cells, while major enriched terms involve binding of several growth factors such as PDGF (platelet-derived growth factor), TGF-β (transforming growth factor beta), and VEGF (vascular endothelial growth factor), as well as terms that relate

to cytoskeleton and ECM structures and organizations. Classification of this cluster is partially supported by the aforementioned growth factors, which are known to interact with endothelial cells and participate in angiogenesis[16]. However, the presence of a large amount of actin cytoskeleton and ECM related terms could also suggest identity of stellate cells, a type of pancreatic cells that support the formation and maturation of pancreatic endothelial cells[16] and are shown to promote islet fibrosis and regulate ECM turnover[17].
Cluster 7 is enriched with angiogenesis related processes, TGF-β and VEGF receptor binding as well as leukocyte recruitment, all of which are typical functions of pancreatic endothelial cells. Therefore, cluster 6 remains unconfirmed and cluster 7 is considered correct.

Conclusion

In summary, we implemented a single-cell RNA-seq method to characterize the transcriptomes of thousands of pancreatic cells from one human donor. Transcript quantification was performed against the latest transcriptome reference. We customized the UMI count matrix by filtering out unwanted variations, as well as low quality records. Cells were clustered into six different types based on the expression profiles: alpha cells, beta cells, delta cells, acinar cells, ductal cells, and endothelial cells. Based on GSEA results, cluster 0 is likely a misclassification of beta cell as alpha cell; cluster 1 has ambiguous identity due to mixed enrichment terms; cluster 2 is considered a hybrid and contaminated cluster of beta cells and acinar cells; cluster 6 is likely a hybrid of endothelial and stellate cells; cluster 3,4,5 and 7 were correctly classified.

References

[1] El Sayed SA, Mukherjee S. Physiology, Pancreas. [Updated 2020 Jul 10]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2021 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK459261/

[2] Muraro, M. J., Dharmadhikari, G., Grün, D., Groen, N., Dielen, T., Jansen, E., van Gurp, L., Engelse, M. A., Carlotti, F., de Koning, E. J., & van Oudenaarden, A. (2016). A Single-Cell Transcriptome Atlas of the Human Pancreas. *Cell systems*, *3*(4), 385–394.e3. https://doi.org/10.1016/j.cels.2016.09.002

[3] Baron, M., Veres, A., Wolock, S. L., Faust, A. L., Gaujoux, R., Vetere, A., Ryu, J. H., Wagner, B. K., Shen-Orr, S. S., Klein, A. M., Melton, D. A., & Yanai, I. (2016). A Single-Cell Transcriptomic Map of the Human and Mouse Pancreas Reveals Inter- and Intra-cell Population Structure. *Cell systems*, *3*(4), 346–360.e4. https://doi.org/10.1016/j.cels.2016.08.011

[4] Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., & Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nature methods, 14(4), 417–419. https://doi.org/10.1038/nmeth.4197

[5] Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, Clark NR, Ma'ayan A. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics. 2013; 128(14).

[6] Quesada, I., Tudurí, E., Ripoll, C., & Nadal, A. (2008). Physiology of the pancreatic alpha-cell and glucagon secretion: role in glucose homeostasis and diabetes. *The Journal of endocrinology*, *199*(1), 5–19. https://doi.org/10.1677/JOE-08-0290

[7] Types of Cells in the Pancreas. (2020, August 14). Retrieved May 4, 2021, from https://med.libretexts.org/@go/page/7777

[8] Cooley, Michele M, Jones, Elaina K, Gorelick, Fred S, and Groblewski, Guy E. (2020). Pancreatic Acinar Cell Protein Synthesis, Intracellular Transport, and Export. *Pancreapedia: Exocrine Pancreas Knowledge Base*, DOI: 10.3998/panc.2020.15

[9] Hegyi, P., Maléth, J., Venglovecz, V., & Rakonczay, Z., Jr (2011). Pancreatic ductal bicarbonate secretion: challenge of the acinar Acid load. *Frontiers in physiology*, *2*, 36. https://doi.org/10.3389/fphys.2011.00036

[10] Grapin-Botton A. (2005). Ductal cells of the pancreas. *The international journal of biochemistry & cell biology*, *37*(3), 504–510. https://doi.org/10.1016/j.biocel.2004.07.010

[11] O'Toole TJ, Sharma S. Physiology, Somatostatin. [Updated 2020 Aug 11]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2021 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK538327/

[12] Rorsman, P., & Huising, M. O. (2018). The somatostatin-secreting pancreatic δ-cell in health and disease. *Nature reviews. Endocrinology*, *14*(7), 404–414. https://doi.org/10.1038/s41574-018-0020-6

[13] Arrojo E Drigo, R., Jacob, S., García-Prieto, C. F., Zheng, X., Fukuda, M., Nhu, H., Stelmashenko, O., Peçanha, F., Rodriguez-Diaz, R., Bushong, E., Deerinck, T., Phan, S., Ali, Y., Leibiger, I., Chua, M., Boudier, T., Song, S. H., Graf, M., Augustine, G. J., Ellisman, M. H., … Berggren, P. O. (2019). Structural basis for delta cell paracrine regulation in pancreatic islets. *Nature communications*, *10*(1), 3700. https://doi.org/10.1038/s41467-019-11517-x
[14] Jonsson, A., Hedin, A., Müller, M. *et al.* Transcriptional profiles of human islet and exocrine endothelial cells in subjects with or without impaired glucose metabolism. *Sci Rep* 10, 22315 (2020). https://doi.org/10.1038/s41598-020-79313-y
[15] Alberts B, Johnson A, Lewis J, et al. Molecular Biology of the Cell. 4th edition. New York: Garland Science; 2002. Blood Vessels and Endothelial Cells. Available from: https://www.ncbi.nlm.nih.gov/books/NBK26848/
[16] Ranjan, A. K., Joglekar, M. V., & Hardikar, A. A. (2009). Endothelial cells in pancreatic islet development and function. *Islets*, *1*(1), 2–9. https://doi.org/10.4161/isl.1.1.9054
[17] Zhou, Y., Sun, B., Li, W., Zhou, J., Gao, F., Wang, X., Cai, M., & Sun, Z. (2019). Pancreatic Stellate Cells: A Rising Translational Physiology Star as a Potential Stem Cell Type for Beta Cell Neogenesis. *Frontiers in physiology*, *10*, 218. https://doi.org/10.3389/fphys.2019.00218

Appendix

Table S1. Gene set enrichment analysis result

cluster 0 alpha_1			
Term	P_adj	Combined.Score	AnnotationDB
amyloid-beta binding (GO:0001540)	7.35E-02	53.5	GO_Molecular_Function_2018
clathrin coat of trans-Golgi network vesicle (GO:0030130)	1.52E-03	297.1	GO_Cellular_Component_2018
invadopodium (GO:0071437)	7.49E-03	120	GO_Cellular_Component_2018
lysosomal lumen (GO:0043202)	5.00E-05	87.8	GO_Cellular_Component_2018
integral component of endoplasmic reticulum membrane (GO:0030176)	5.00E-05	68.3	GO_Cellular_Component_2018
AP-1 adaptor complex (GO:0030121)	8.88E-02	65.6	GO_Cellular_Component_2018
protein retention in ER lumen (GO:0006621)	9.94E-02	228.2	GO_Biological_Process_2018
maintenance of protein localization in endoplasmic reticulum (GO:0035437)	4.01E-02	223.9	GO_Biological_Process_2018
insulin secretion (GO:0030073)	6.71E-03	173	GO_Biological_Process_2018
regulation of long-term neuronal synaptic plasticity (GO:0048169)	8.44E-02	143.9	GO_Biological_Process_2018

regulation of insulin secretion (GO:0050796)	1.85E-04	100	GO_Biological_Process_2018
Lysosome	4.57E-07	122.5	KEGG_2019_Human
Circadian entrainment	2.73E-04	68.3	KEGG_2019_Human
Sphingolipid metabolism	5.06E-03	58.1	KEGG_2019_Human
Glycosaminoglycan degradation	4.83E-02	48.6	KEGG_2019_Human
Biosynthesis of unsaturated fatty acids	3.19E-02	45.8	KEGG_2019_Human

cluster 1 alpha_2			
Term	P_adj	Combined.Score	AnnotationDB
ubiquinone binding (GO:0048039)	7.01E-02	857.3	GO_Molecular_Function_2018
retinal dehydrogenase activity (GO:0001758)	7.01E-02	692.1	GO_Molecular_Function_2018
3-chloroallyl aldehyde dehydrogenase activity (GO:0004028)	7.01E-02	576.7	GO_Molecular_Function_2018
tau protein binding (GO:0048156)	7.01E-02	576.7	GO_Molecular_Function_2018
activin receptor binding (GO:0070697)	7.01E-02	576.7	GO_Molecular_Function_2018
apical dendrite (GO:0097440)	3.71E-02	576.7	GO_Cellular_Component_2018
spherical high-density lipoprotein particle (GO:0034366)	3.71E-02	576.7	GO_Cellular_Component_2018
secretory granule lumen (GO:0034774)	8.04E-04	171.6	GO_Cellular_Component_2018
platelet alpha granule lumen (GO:0031093)	3.24E-02	162.6	GO_Cellular_Component_2018

platelet alpha granule (GO:0031091)	3.24E-02	108.1	GO_Cellular_Component_2018
positive regulation of cardiac muscle contraction (GO:0060452)	8.20E-02	857.3	GO_Biological_Process_2018
negative regulation of catecholamine secretion (GO:0033604)	8.20E-02	857.3	GO_Biological_Process_2018
regulation of relaxation of cardiac muscle (GO:1901897)	8.20E-02	857.3	GO_Biological_Process_2018
nucleobase metabolic process (GO:0009112)	8.20E-02	857.3	GO_Biological_Process_2018
negative regulation of erythrocyte differentiation (GO:0045647)	8.20E-02	857.3	GO_Biological_Process_2018
Thyroid hormone synthesis	4.02E-02	141.9	KEGG_2019_Human
Complement and coagulation cascades	4.02E-02	129.6	KEGG_2019_Human

cluster 2 beta			
Term	P_adj	Combined.Score	AnnotationDB
small ribosomal subunit rRNA binding (GO:0070181)	4.04E-03	1027.5	GO_Molecular_Function_2018
mRNA 5'-UTR binding (GO:0048027)	3.54E-05	985.8	GO_Molecular_Function_2018
insulin-like growth factor receptor binding (GO:0005159)	4.66E-04	859.9	GO_Molecular_Function_2018
RNA binding (GO:0003723)	4.91E-19	621.1	GO_Molecular_Function_2018

translation elongation factor activity (GO:0003746)	9.26E-03	394.5	GO_Molecular_Function_2018
cytosolic ribosome (GO:0022626)	3.37E-50	21659.1	GO_Cellular_Component_2018
cytosolic large ribosomal subunit (GO:0022625)	7.01E-38	18294.1	GO_Cellular_Component_2018
large ribosomal subunit (GO:0015934)	1.59E-37	16978.8	GO_Cellular_Component_2018
cytosolic part (GO:0044445)	9.31E-47	14560.9	GO_Cellular_Component_2018
polysomal ribosome (GO:0042788)	1.54E-15	5879.5	GO_Cellular_Component_2018
SRP-dependent cotranslational protein targeting to membrane (GO:0006614)	2.18E-54	38352.3	GO_Biological_Process_2018
cotranslational protein targeting to membrane (GO:0006613)	5.67E-54	35415.8	GO_Biological_Process_2018
protein targeting to ER (GO:0045047)	1.81E-53	32857.4	GO_Biological_Process_2018
viral gene expression (GO:0019080)	1.33E-51	26439.8	GO_Biological_Process_2018
nuclear-transcribed mRNA catabolic process, nonsense-mediated decay (GO:0000184)	2.03E-51	25647.5	GO_Biological_Process_2018
Ribosome	6.29E-49	17301.4	KEGG_2019_Human
Maturity onset diabetes of the young	1.11E-06	1346.3	KEGG_2019_Human

Type II diabetes mellitus	1.21E-02	165.2	KEGG_2019_Human
Insulin secretion	6.49E-03	131	KEGG_2019_Human
Butanoate metabolism	7.74E-02	130.4	KEGG_2019_Human

cluster 3 ductal			
Term	P_adj	Combined.Score	AnnotationDB
cadherin binding (GO:0045296)	1.41E-25	467	GO_Molecular_Function_2018
cadherin binding involved in cell-cell adhesion (GO:0098641)	1.32E-05	341.4	GO_Molecular_Function_2018
protein binding involved in cell-cell adhesion (GO:0098632)	2.67E-05	272.3	GO_Molecular_Function_2018
MAP kinase tyrosine/serine/threonine phosphatase activity (GO:0017017)	4.50E-02	144.5	GO_Molecular_Function_2018
protein binding involved in heterotypic cell-cell adhesion (GO:0086080)	4.50E-02	144.5	GO_Molecular_Function_2018
focal adhesion (GO:0005925)	1.07E-38	830.2	GO_Cellular_Component_2018
cortical cytoskeleton (GO:0030863)	3.22E-10	331.7	GO_Cellular_Component_2018
specific granule (GO:0042581)	2.57E-14	265.8	GO_Cellular_Component_2018
tertiary granule (GO:0070820)	4.09E-14	252.2	GO_Cellular_Component_2018
pseudopodium (GO:0031143)	9.68E-04	203.7	GO_Cellular_Component_2018
establishment of endothelial intestinal barrier (GO:0090557)	8.29E-04	627.1	GO_Biological_Process_2018

positive regulation of myeloid leukocyte cytokine production involved in immune response (GO:0061081)	8.29E-04	627.1	GO_Biological_Process_2018
regulation of extracellular exosome assembly (GO:1903551)	8.29E-04	627.1	GO_Biological_Process_2018
hemidesmosome assembly (GO:0031581)	5.61E-05	496.7	GO_Biological_Process_2018
cell-substrate junction assembly (GO:0007044)	5.61E-05	496.7	GO_Biological_Process_2018
ECM-receptor interaction	2.13E-12	339	KEGG_2019_Human
Focal adhesion	5.26E-15	263.6	KEGG_2019_Human
Leukocyte transendothelial migration	2.77E-12	261.8	KEGG_2019_Human
Tight junction	3.87E-14	257.3	KEGG_2019_Human
Arrhythmogenic right ventricular cardiomyopathy (ARVC)	6.83E-10	240.5	KEGG_2019_Human

cluster 4 acinar			
Term	P_adj	Combined.Score	AnnotationDB
cadherin binding involved in cell-cell adhesion (GO:0098641)	1.00E-06	454.3	GO_Molecular_Function_2018
protein binding involved in cell-cell adhesion (GO:0098632)	2.39E-06	356.5	GO_Molecular_Function_2018
cadherin binding (GO:0045296)	1.83E-21	342.5	GO_Molecular_Function_2018

bile acid binding (GO:0032052)	4.13E-03	244	GO_Molecular_Function_2018
protein-disulfide reductase activity (GO:0047134)	1.70E-02	187.6	GO_Molecular_Function_2018
focal adhesion (GO:0005925)	5.10E-22	321.7	GO_Cellular_Component_2018
zonula adherens (GO:0005915)	5.83E-03	187.6	GO_Cellular_Component_2018
pseudopodium (GO:0031143)	1.87E-03	183.2	GO_Cellular_Component_2018
tertiary granule lumen (GO:1904724)	2.41E-06	138.7	GO_Cellular_Component_2018
integral component of lumenal side of endoplasmic reticulum membrane (GO:0071556)	1.51E-04	122.3	GO_Cellular_Component_2018
positive regulation of myeloid leukocyte cytokine production involved in immune response (GO:0061081)	1.81E-03	567.1	GO_Biological_Process_2018
PERK-mediated unfolded protein response (GO:0036499)	2.69E-04	355.2	GO_Biological_Process_2018
positive regulation of apoptotic cell clearance (GO:2000427)	3.35E-03	348.9	GO_Biological_Process_2018
positive regulation of viral entry into host cell (GO:0046598)	4.50E-03	244	GO_Biological_Process_2018

regulation of apoptotic cell clearance (GO:2000425)	4.50E-03	244	GO_Biological_Process_2018
Pancreatic secretion	6.17E-08	153.1	KEGG_2019_Human
Fluid shear stress and atherosclerosis	6.17E-08	111.4	KEGG_2019_Human
Adherens junction	4.81E-06	98.8	KEGG_2019_Human
Focal adhesion	6.17E-08	93.6	KEGG_2019_Human
Tight junction	1.24E-07	92.9	KEGG_2019_Human

cluster 5 delta			
Term	P_adj	Combined.Score	AnnotationDB
syntaxin-1 binding (GO:0017075)	6.79E-02	109.2	GO_Molecular_Function_2018
syntaxin binding (GO:0019905)	7.26E-03	83.8	GO_Molecular_Function_2018
ankyrin binding (GO:0030506)	9.92E-02	80.1	GO_Molecular_Function_2018
hydrogen-exporting ATPase activity (GO:0036442)	9.92E-02	61.7	GO_Molecular_Function_2018
inward rectifier potassium channel activity (GO:0005242)	9.92E-02	61.7	GO_Molecular_Function_2018
adrenal gland development (GO:0030325)	9.89E-02	223.4	GO_Biological_Process_2018
regulation of insulin secretion (GO:0050796)	4.15E-04	129.6	GO_Biological_Process_2018
regulation of peptide hormone secretion (GO:0090276)	1.16E-03	127	GO_Biological_Process_2018
negative regulation of neurogenesis (GO:0050768)	4.44E-02	106.6	GO_Biological_Process_2018

regulation of calcium ion-dependent exocytosis (GO:0017158)	4.44E-02	75.5	GO_Biological_Process_2018
Insulin secretion	1.06E-07	259.6	KEGG_2019_Human
Maturity onset diabetes of the young	2.62E-03	161.1	KEGG_2019_Human
Synaptic vesicle cycle	2.34E-03	81.9	KEGG_2019_Human
Vasopressin-regulated water reabsorption	1.24E-02	64.7	KEGG_2019_Human
Type II diabetes mellitus	1.24E-02	59.8	KEGG_2019_Human

cluster 6 endothelial_1			
Term	P_adj	Combined.Score	AnnotationDB
platelet-derived growth factor binding (GO:0048407)	2.21E-09	2788	GO_Molecular_Function_2018
collagen binding (GO:0005518)	5.11E-15	719.2	GO_Molecular_Function_2018
vascular endothelial growth factor receptor 2 binding (GO:0043184)	1.41E-03	502.9	GO_Molecular_Function_2018
transforming growth factor beta binding (GO:0050431)	1.20E-06	399.4	GO_Molecular_Function_2018
protein binding involved in cell-matrix adhesion (GO:0098634)	2.62E-03	308.7	GO_Molecular_Function_2018
focal adhesion (GO:0005925)	8.67E-52	1260.3	GO_Cellular_Component_2018
endoplasmic reticulum lumen (GO:0005788)	3.19E-21	335.4	GO_Cellular_Component_2018

actomyosin (GO:0042641)	1.05E-07	200.5	GO_Cellular_Component_2018
contractile actin filament bundle (GO:0097517)	6.39E-07	184.2	GO_Cellular_Component_2018
stress fiber (GO:0001725)	6.39E-07	184.2	GO_Cellular_Component_2018
bleb assembly (GO:0032060)	6.80E-05	807	GO_Biological_Process_2018
negative regulation of platelet-derived growth factor receptor-beta signaling pathway (GO:2000587)	6.80E-05	807	GO_Biological_Process_2018
extracellular matrix organization (GO:0030198)	4.59E-32	797.6	GO_Biological_Process_2018
collagen fibril organization (GO:0030199)	5.75E-10	640.4	GO_Biological_Process_2018
epithelial to mesenchymal transition involved in endocardial cushion formation (GO:0003198)	7.51E-04	502.9	GO_Biological_Process_2018
Focal adhesion	2.98E-21	418.9	KEGG_2019_Human
ECM-receptor interaction	4.63E-11	258.3	KEGG_2019_Human
Regulation of actin cytoskeleton	3.94E-12	164.4	KEGG_2019_Human
Bacterial invasion of epithelial cells	1.67E-07	144.4	KEGG_2019_Human
Protein digestion and absorption	3.38E-07	113.1	KEGG_2019_Human

cluster 7 endothelial_2			
Term	P_adj	Combined.Score	AnnotationDB

type I transforming growth factor beta receptor binding (GO:0034713)	7.70E-05	803.7	GO_Molecular_Function_2018
ephrin receptor binding (GO:0046875)	3.87E-07	659.7	GO_Molecular_Function_2018
vascular endothelial growth factor-activated receptor activity (GO:0005021)	8.96E-04	498.3	GO_Molecular_Function_2018
vascular endothelial growth factor receptor 2 binding (GO:0043184)	7.81E-03	268.8	GO_Molecular_Function_2018
vascular endothelial growth factor receptor binding (GO:0005172)	2.22E-03	264.7	GO_Molecular_Function_2018
focal adhesion (GO:0005925)	6.36E-18	265.9	GO_Cellular_Component_2018
platelet alpha granule (GO:0031091)	3.53E-08	176.1	GO_Cellular_Component_2018
membrane raft (GO:0045121)	1.43E-08	162.8	GO_Cellular_Component_2018
platelet alpha granule lumen (GO:0031093)	1.39E-06	144	GO_Cellular_Component_2018
platelet alpha granule membrane (GO:0031092)	1.33E-03	135.9	GO_Cellular_Component_2018
mitral valve morphogenesis (GO:0003183)	2.95E-06	1186.6	GO_Biological_Process_2018
vascular smooth muscle cell development (GO:0097084)	2.33E-04	804.6	GO_Biological_Process_2018
vasculogenesis (GO:0001570)	6.19E-10	657.7	GO_Biological_Process_2018

branching involved in blood vessel morphogenesis (GO:0001569)	2.96E-06	593.1	GO_Biological_Process_2018
negative regulation of endothelial cell migration (GO:0010596)	1.21E-09	574.8	GO_Biological_Process_2018
ECM-receptor interaction	7.25E-12	344.3	KEGG_2019_Human
Focal adhesion	4.65E-14	260.4	KEGG_2019_Human
Rap1 signaling pathway	9.80E-12	180.6	KEGG_2019_Human
Proteoglycans in cancer	2.55E-11	171.2	KEGG_2019_Human
Leukocyte transendothelial migration	8.88E-09	157.9	KEGG_2019_Human