© 2022. Published by The Company of Biologists Ltd | Development (2022) 149, dev200432. doi:10.1242/dev.200432

HUMAN DEVELOPMENT

TECHNIQUES AND RESOURCES ARTICLE

A PDX1 cistrome and single-cell transcriptome resource of the developing pancreas

Xiaodun Yang[1,2,*], Jeffrey C. Raum[1,2,*], Junil Kim[3,*], Reynold Yu[1], Juxiang Yang[4], Gabriella Rice[5], Changhong Li[4], Kyoung-Jae Won[6], Diana E. Stanescu[4,7,‡] and Doris A. Stoffers[1,2,‡]

ABSTRACT

Pancreatic and duodenal homeobox 1 (PDX1) is crucial for pancreas organogenesis, yet the dynamic changes in PDX1 binding in human or mouse developing pancreas have not been examined. To address this knowledge gap, we performed PDX1 ChIP-seq and single-cell RNA-seq using fetal human pancreata. We integrated our datasets with published datasets and revealed the dynamics of PDX1 binding and potential cell lineage-specific PDX1-bound genes in the pancreas from fetal to adult stages. We identified a core set of developmentally conserved PDX1-bound genes that reveal the broad multifaceted role of PDX1 in pancreas development. Despite the well-known dramatic changes in PDX1 function and expression, we found that PDX1-bound genes are largely conserved from embryonic to adult stages. This points towards a dual role of PDX1 in regulating the expression of its targets at different ages, dependent on other functionally congruent or directly interacting partners. We also showed that PDX1 binding is largely conserved in mouse pancreas. Together, our study reveals PDX1 targets in the developing pancreas *in vivo* and provides an essential resource for future studies on pancreas development.

KEY WORDS: Pancreas development, ChIP-seq, Single-cell RNA-seq, Mouse, Human, PDX1 cistrome

INTRODUCTION

Environmental and genetic factors contribute to diabetes pathogenesis. Mutations of key transcription factors cause defects in endocrine pancreas development and impose susceptibility for diabetes. One of the most well-studied pancreatic transcription factors is the human diabetes gene pancreatic and duodenal homeobox 1 (*PDX1*), which plays essential roles in pancreas organogenesis, endocrine pancreas development, and the growth and function of insulin-secreting β-cells in both mouse and human (Pan and Wright, 2011; Jennings et al., 2020).

PDX1 is a hallmark transcription factor of pancreas development. One of the early signs of pancreas organogenesis is the induction of *Pdx1*/*PDX1* in the foregut endoderm at embryonic day 8.5 (E8.5) in mouse (Guz et al., 1995) and at ∼29 days post conception (dpc) in human (Jennings et al., 2013). PDX1 expression is gradually enriched in both mouse and human pancreatic β-cells at later stages (reviewed by Pan and Wright, 2011). Rare homozygous loss of function of *Pdx1*/*PDX1* causes pancreas agenesis in both mouse (Jonsson et al., 1994; Offield et al., 1996) and human (Stoffers et al., 1997b), whereas heterozygous mutations in *PDX1* are linked to human type 2 diabetes (Stoffers et al., 1997a; Hani et al., 1999).

PDX1 genome-wide target genes have been identified in human embryonic stem cell-derived pancreatic progenitors (hPSC-PP) (Teo et al., 2015; Wang et al., 2015, 2018), mouse β-cell lines (Keller et al., 2007; Perelis et al., 2015), and adult mouse and human islets (Khoo et al., 2012). PDX1 directly regulates important transcription factors in pancreas development (Wang et al., 2018) and genes that function in endocrine system and metabolic disorders in adult islet cells (Khoo et al., 2012). However, the PDX1 cistrome of the developing pancreas has not been examined in mouse or human, mainly due to the limited availability of mouse embryonic pancreas tissue and to limited access to human fetal pancreas samples, respectively. As a result, a significant gap of knowledge remains with regard to PDX1 targets during pancreas development, particularly for human pancreas development. Single-cell studies have examined cell type-specific gene expression during the differentiation and maturation of mouse and human pancreatic cells (reviewed by Yu and Xu, 2020), including two studies of human fetal pancreas (Ramond et al., 2018; Yu et al., 2021). Integration of PDX1 cistrome with single-cell RNA-seq data could reveal PDX1 cell lineage-specific targets in the developing pancreas.

To address this critical gap, we performed PDX1 chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq) using human fetal pancreata at 14 weeks gestation and mouse embryonic pancreata at E13.5 and E15.5, and single-cell RNA-seq using human fetal pancreata at 15-20 weeks gestation. By integrating our data with published datasets (Khoo et al., 2012; Bramswig et al., 2013; Krentz et al., 2018; Wang et al., 2018; Xin et al., 2018; Bastidas-Ponce et al., 2019), we characterized in parallel the transcriptomic signatures of human and mouse pancreas cells and integrated them with the PDX1 cistrome to identify the temporal pattern of PDX1 targets expressed in each lineage.

RESULTS

Temporal analysis of PDX1-bound genes in human developing pancreas and adult islets

To examine PDX1 genome-wide targets in human developing pancreas, we performed PDX1 ChIP-seq using human fetal

[1]Institute of Diabetes, Obesity and Metabolism, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA. [2]Department of Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA. [3]School of Systems Biomedical Science, Soongsil University, 369 Sangdo-Ro, Dongjak-Gu, Seoul 06978, Republic of Korea. [4]Division of Endocrinology and Diabetes, The Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA. [5]Department of Cell and Developmental Biology, Institute for Regenerative Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA. [6]Biotech Research & Innovation Centre, University of Copenhagen, Copenhagen 2200, Denmark. [7]Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA.
*These authors contributed equally to this work

‡Authors for correspondence (stanescud@chop.edu, stoffers@pennmedicine.upenn.edu)

R.Y., 0000-0002-8093-254X; D.A.S., 0000-0003-2626-2124

Handling Editor: James Wells
Received 8 December 2021; Accepted 31 May 2022

Development (2022) 149, dev200432. doi:10.1242/dev.200432

pancreata at 14 weeks gestational age. All fetal pancreatic epithelial cells expressed PDX1 at this stage (Fig. S1A). HOMER *de novo* motif analysis showed that the consensus PDX1-binding motif was the most enriched (Fig. S1B) and histogram distribution of PDX1 peaks shows the majority of PDX1-binding sites were located within 2 kb of the transcriptional starting site (Fig. S1C). We compared these newly generated PDX1 cistromic data with published PDX1 ChIP-seq data from hPSC-PPs and adult islets (Tables S1-3) (Khoo et al., 2012; Wang et al., 2018). PDX1-binding events detected were a combination of different cell types due to the heterogeneity of cell types expressing PDX1 in the hPSC-PPs and fetal pancreata. There were 3934 genes bound by PDX1 in hPSC-PP cells only (Fig. 1A, Table S3). Canonical pathways with enrichment of these genes included DNA methylation and transcriptional repression signaling, which may help to establish proper epigenetic regulation (Wu and Sun, 2006), and retinoic acid-mediated apoptosis signaling, which may facilitate pancreatic lineage specification (Lorberbaum et al., 2020) (Fig. S2A, Table S4). There were 1092 genes bound by PDX1 in fetal pancreas only (Fig. 1A, Table S3). Canonical pathways with enrichment of these genes included the sirtuin signaling pathway and fatty acid β-oxidation I, which may help to maintain energy homeostasis for pancreas development *in vivo* (Yaney and Corkey, 2003; Wu et al., 2012) (Fig. S2A, Table S4). This set of PDX1-bound genes in fetal pancreas may provide clues to identify novel signals that maintain energy homeostasis to improve *in vitro* differentiation process using stem cells. A total of 3131 genes were bound by PDX1 in all three datasets, representing ∼50% of bound genes in the human fetal pancreas (Fig. 1A). Significantly, PDX1-bound genes in hPSC-PP, fetal and adult cells with important roles in specification and differentiation of pancreatic endocrine progenitors (e.g. *NEUROG3*, *NKX6-1*, *NKX2-2*, *RFX3* and *HNF1B*) and/or in the function of adult β-cells (such as *JUN*, *GCK* and *KCNJ11*) (Fig. 1B). We analyzed changes in PDX1 binding in hPSC-PP, fetal and adult cells (Fig. 1B). We identified genes bound only in pancreas progenitors (*NEUROD1*, *JUND* and *SIX3*), during development (*HNF4A*, *ONECUT1*, *CLDN4* and *PTF1A*) or only in adult islet (*SYT4*, *NPY* and *CHGA*) (Fig. 1B). In human adult islets, genes involved in β-cell maturation were bound by PDX1 (Table S3) (Khoo et al., 2012), including *MAFA*, a β-cell maturation driver, *UCN3*, a β-cell maturation marker, and *GCK*, a low-affinity glucokinase (reviewed by Liu and Hebrok, 2017), confirming the involvement of PDX1 in β-cell maturation. Moreover, a δ-cell differentiation regulator *Hhex* (Zhang et al., 2014) is bound by PDX1 in both developing pancreas and adult islet (Table S3). The δ-cell hormone gene *SST* is bound by PDX1 in adult human islet (Table S3). The data highlighted the role of PDX1 in δ-cell differentiation and function. Gene-ontology (GO) and upstream regulator analysis of each category for PDX1-bound genes is presented in Fig. S2A,B and Tables S4, S5. These findings expand the list of transcription factors and co-factors that are potentially either functionally congruent or directly interacting PDX1 partners.

Comparative single-cell transcriptome analysis of human fetal pancreas and adult islet cells reveals cell lineage-specific and temporally regulated genes

We used single-cell RNA-seq analysis to address the role of PDX1 in the various cell types found in the human fetal pancreas at 15-20 weeks gestational age, and we compared these newly generated fetal-cell transcriptomes with adult pancreas cell transcriptomes (Xin et al., 2018). Fetal and adult pancreatic cells largely overlapped in each cluster (Fig. 2A,B, Figs S3, S4, S6) by anchor-based method implemented in Seurat. The Seurat integration method identifies matching cell states of different datasets by considering mutual nearest neighbors as 'anchors'. By using this anchor-based method, Seurat can exclude unique populations for the matching cell states. The differentially expressed genes among different cell lineages at fetal or adult stage were curated into transcriptomic signatures for each cell lineage (Fig. 2C-D, Fig. S5, Tables S6, S7).

Interestingly, we identified two clusters of putative acinar cells, acinar 1 and acinar 2, with acinar 1 containing almost exclusively fetal cells, while acinar 2 contained both fetal and adult cells. Despite separate sub-clustering, several transcripts had much higher expression in fetal compared with the adult cells (Fig. 3C,D). The expression of NKX6.1 and SOX9 in the acinar 1 cluster suggest that this cluster contains acinar cells in an early progenitor state. Overall, within the same cell lineage, a significant number of genes showed differential expression between fetal and adult cells (Fig. 3A-F). It is possible that contamination of highly expressed endocrine transcripts during single-cell encapsulation may affect differential expression analysis. Nevertheless, these gene profile variations over age constitute a resource for further investigation of maturational changes occurring during pancreas organogenesis and tissue homeostasis.

Integrating PDX1 ChIP-seq with single-cell RNA-seq identifies putative functional roles of PDX1 in target regulation in fetal pancreas and adult islets

We imputed plausible cell lineage-specific PDX1 targets in human pancreas by integrating single-cell transcriptomes with PDX1

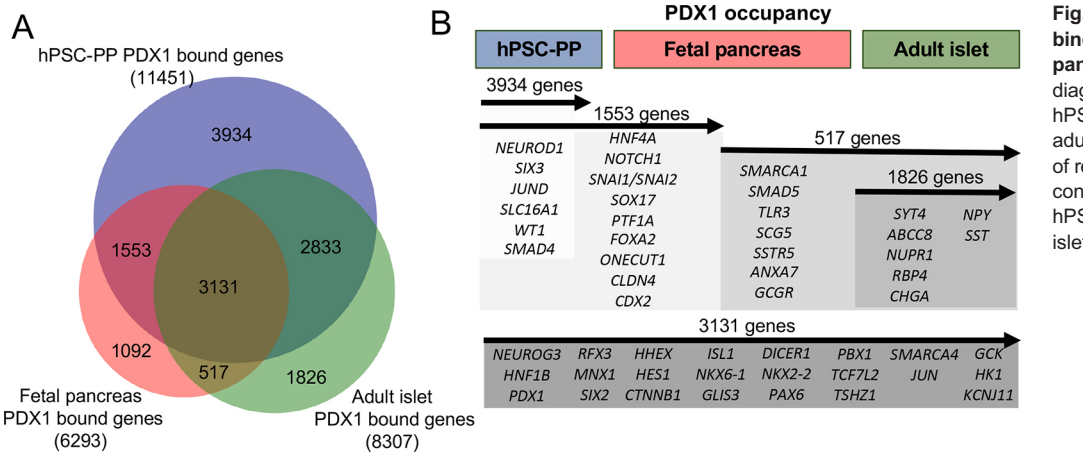

Fig. 1. Temporal analysis of PDX1 binding in human developing pancreas and adult islets. (A) Venn diagram of PDX1-bound genes in hPSC-PP cells, fetal pancreas and adult islet. (B) Temporal arrangement of representative stage-specific and conserved genes bound by PDX1 in hPSC-PP, fetal pancreas and adult islet.

Development (2022) 149, dev200432. doi:10.1242/dev.200432

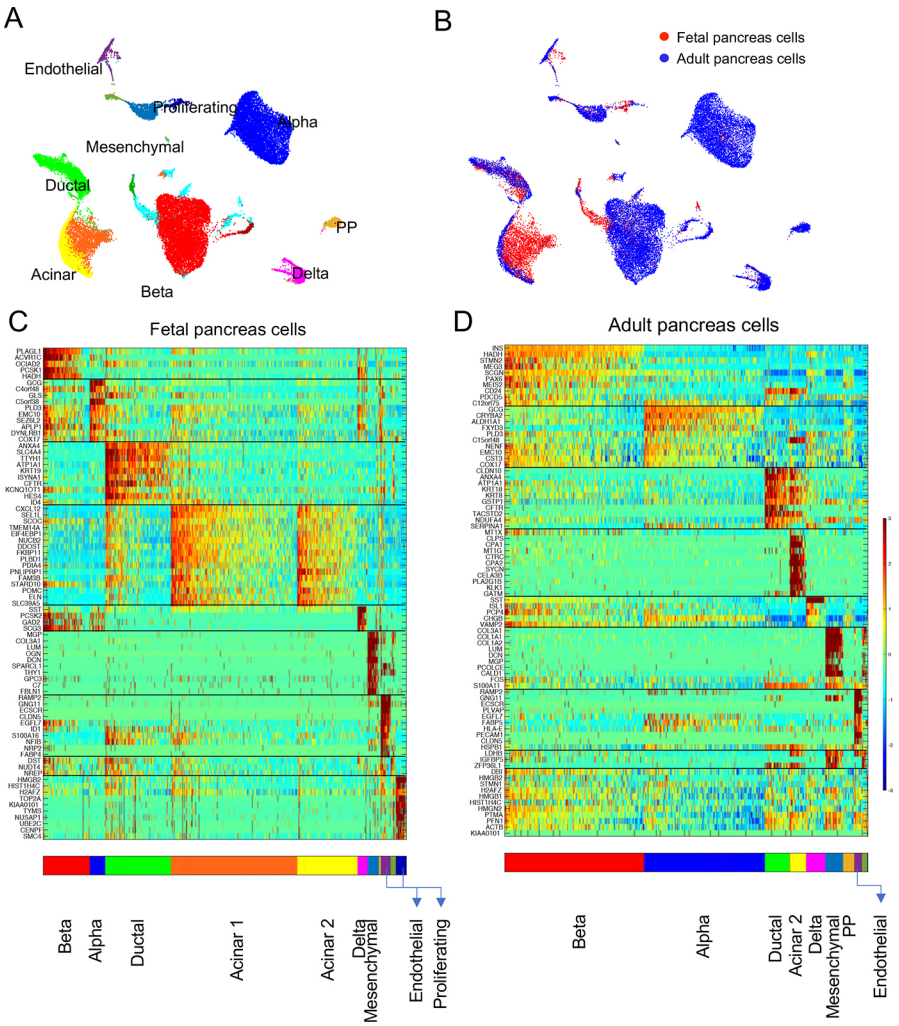

Fig. 2. Comparative single-cell transcriptome analysis of human fetal pancreas and adult islet cells reveals cell-lineage-specific and temporally regulated genes. (A,B) UMAP plots showing cell clusters (A) identified in single-cell RNA-seq from fetal and adult cells (B). (C) Heatmap showing differentially expressed genes in each cluster among human fetal pancreas cells. (D) Heatmap showing differentially expressed genes in each cluster among human adult pancreas cells. For C,D, the DEGs in each cluster were picked up by FDR <0.01 and log ratio >0.5. The DEGs commonly found in multiple clusters were excluded. If there were more than 10 DEGs in each cluster, only the top 10 DEGs by log ratio were included in the heatmap. The color codes of different clusters in A are related to different cell types in C and D. For example, the red cluster in A contains β-cells and the red bars in C,D show differentially expressed genes in β-cells. The acinar 1 cluster is in orange and the acinar 2 cluster is in yellow. Some clusters contain unidentified cell types, such as those in cyan.

cistrome data. By these criteria (binding and associated transcriptional expression), we called out 'putative functional PDX1 targets' upregulated in fetal or adult β-cells, ductal cells and acinar cells (Fig. 3G, Tables S8). PDX1 surprisingly bound important early acinar and ductal specific genes, not only during early development but also in the adult human islet. For example, fetal and adult-specific ductal genes, such as *HHEX*, *ANXA2*, *KRT8* and *KRT18* were bound by PDX1 from fetal pancreas to adult islet (Fig. 3G). Similarly, genes enriched in acinar cells, such as *CPA1*, *RBPJ*, *SPINK1* and *CELA2A*, were expressed in fetal and/or adult cells, and were bound by PDX1 in fetal pancreas and in adult islets. By integrating PDX1-bound ductal/acinar genes in adult islets with histone mark ChIP-seq data of human adult β-cells (Bramswig et al., 2013), we found that the majority of PDX1 bound ductal/acinar genes in adult islets were marked by the active histone mark histone H3 lysine 4 trimethylation (H3K4me3) (Fig. S7), indicating potential active transcriptional status.

Temporal analysis of PDX1 binding and cell lineage-specific target expression from developing pancreas to adult islet in mouse

To examine species-conserved and -specific PDX1-binding in mouse pancreas, we performed ChIP-seq of whole mouse embryonic pancreata at E13.5 and E15.5 (Fig. S8A,B, Tables S9, S10). We compared these newly generated data with our previously published PDX1 cistrome of mouse adult islets (Khoo et al., 2012) (Table S11). Similar to the human cistromic analysis, the majority of PDX1 target genes (2040) were consistently bound by PDX1 in mouse pancreas from embryonic to adult stages (Fig. 4A). Some of the known genes required for pancreas development and adult β-cell function were continuously bound from E13.5-E15.5 into the adult islet, including *Neurog3*, *Onecut1*, *Pax6*, *Foxa2*, *Mnx1*, *Nkx6-1* and *Tshz1* (Pan and Wright, 2011; Raum et al., 2015). Pathway analysis and upstream regulator analysis revealed similar results to the analysis from human pancreas (Fig. S8C,D, Tables S12, S13). To characterize the functional roles of PDX1 in different cell types in mouse pancreas, we integrated the identified PDX1 targets with published single-cell RNA-seq data of mouse developing pancreas (Krentz et al., 2018; Bastidas-Ponce et al., 2019). We identified PDX1-bound genes that were highly expressed in multipotent progenitors (including *Dlk1*), endocrine progenitors (including *Neurog3* and *Neurod2*), embryonic endocrine cells (including *Pax6* and *Mafb*), and other cell types (Fig. 4B, Tables S14-S15).

PDX1 directs a developmentally and evolutionarily conserved ductal and endocrine program

We examined species-conserved and specific PDX1 targets in mouse and human pancreas at developmental and adult stages. We found that 762 and 1509 common mouse/human gene targets were bound by PDX1 in developing pancreas and adult islets,

Development (2022) 149, dev200432. doi:10.1242/dev.200432

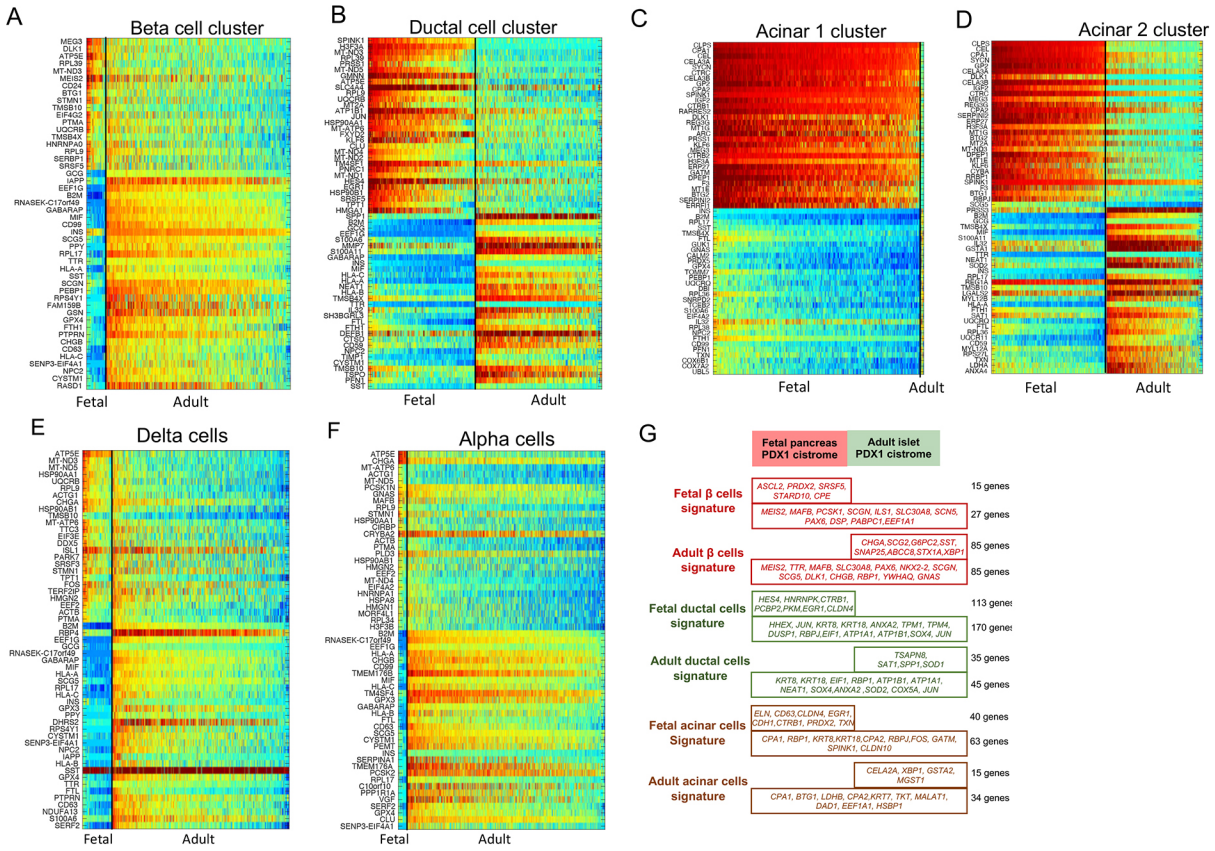

Fig. 3. Heatmap showing differentially expressed genes between human fetal and adult pancreas, and potential PDX1 cell type-specific targets. (A) Genes differentially expressed between human fetal and adult β-cells. (B) Genes differentially expressed between human fetal and adult ductal cells. (C,D) Genes differentially expressed between human fetal and adult acinar cells (the acinar 1 cluster is mainly composed of fetal acinar cells; the acinar 2 cluster is composed of both fetal and adult acinar cells). (E) Genes differentially expressed between human fetal and adult δ-cells. (F) Genes differentially expressed between human fetal and adult α-cells. (G) Integrating PDX1 ChIP-seq with single-cell RNA-seq identifies stage-specific PDX1-bound genes differentially expressed in β, ductal and acinar cells.

respectively (Fig. S9A-D, Tables S17-S19). Of these, 518 were conserved both developmentally and evolutionarily (Fig. 5A, Table S16). Integration of this data subset with our single-cell RNASeq-derived transcriptomic signatures suggested a stable occupancy by PDX1 of embryonic and adult ductal, as well as embryonic β cell, signature genes in adult islets (Fig. 5A). We next integrated the embryonic mouse and human PDX1-bound genes with the single-cell transcriptomic signatures of ductal, acinar and β-cells. None of the genes enriched in human fetal acinar cells or mouse embryonic acinar cells was part of the embryonic PDX1 cistrome. Ten genes expressed in embryonic ductal cells were PDX1 targets in pancreas development in both mouse and human (Fig. 5B). Six of these were also bound by PDX1 in adult islets in both mouse and human (bold in Fig. 5B). These ductal-specific genes are involved in mitochondrial function (*ATP1B1* and *ATP1A1*), or related to pancreatic cancer progression (*ACTN4*, *TPM1*, *CYR61* and *CLDN3*) (Westmoreland et al., 2012; Abou-Kheir et al., 2020). We found that some co-factors, potentially either functionally congruent or directly interacting PDX1 partners, showed stage-specific interactions with PDX1 (Table S19), suggesting that these co-factors contribute to temporally divergent regulatory functions of PDX1. In parallel, only four genes were expressed in embryonic β-cells and bound by PDX1 in embryonic mouse and fetal human samples: *MEIS2*, *ISL1*, *PAX6*, and *SCGN* – all of which play known and important roles in pancreas development (Swift et al., 1998; Ashery-Padan et al., 2004; Du et al., 2009; Malenczyk et al., 2018). *MEIS2*, *ISL1* and *PAX6* were also bound in the adult mouse and human islet.

DISCUSSION

Identification of genome-wide PDX1 targets in mouse and human developing pancreas *in vivo*

PDX1 has been studied extensively in the pancreas due to its roles in pancreas development and β-cell function, and its identification as a human diabetes gene. Previous studies examined PDX1 genome-wide targets in hPSC-PPs (Teo et al., 2015;Wang et al., 2015, 2018), which revealed the transcriptional networks regulated by PDX1 in the induced differentiation model system. By performing PDX1 ChIP-seq using human fetal pancreata at 14 weeks gestation and mouse embryonic pancreata at E13.5 and E15.5, we characterized the PDX1 cistrome of the developing pancreas. There are 11451 PDX1 targets identified in hPSC-PPs *in vitro* and 6293 PDX1 targets identified in fetal human pancreas *in vivo*. The much lower number of PDX1 targets in fetal pancreas than those in hPSC-PPs maybe due to a lower sensitivity of the ChIP-seq assay and/or to a lower ratio of endocrine lineage cells in human fetal pancreas than in hPSC-PPs. A higher degree of heterogeneity of cell types in fetal human pancreas may also contribute to this difference. PDX1 directly regulated key transcription factors and signaling pathways important in pancreatic lineage establishment and expansion. By integrating PDX1 ChIP-seq with single-cell RNA-seq data, we revealed potential cell lineage-specific PDX1 targets. When

Development (2022) 149, dev200432. doi:10.1242/dev.200432

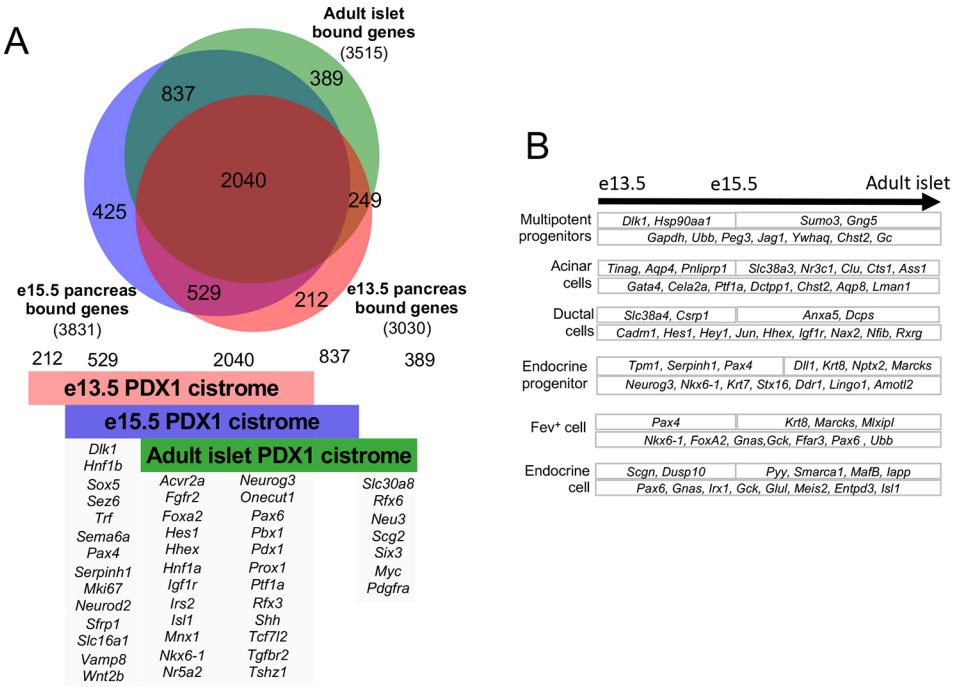

Fig. 4. Temporal analysis of PDX1 binding from developing pancreas to adult islet in mouse. (A) Venn diagram of PDX1-bound genes in E13.5, E15.5 and adult islets, and representative genes among the three datasets. (B) Temporal arrangement of representative stage-specific and conserved genes bound by PDX1 in mouse developing pancreas at E13.5 and E15.5, and in adult islets.

comparing the PDX1 cistrome of mouse and human pancreas, we identified an evolutionarily conserved gene program regulated by PDX1, which functions in key pathways in pancreas development and function, including Wnt/β-catenin signaling at the developmental stage (Sharon et al., 2019), Pi3K/AKT signaling at adult stage (Kaneko et al., 2010), and protein kinase A signaling at both developmental and adult stages (Kaihara et al., 2013) (Fig. S9C).

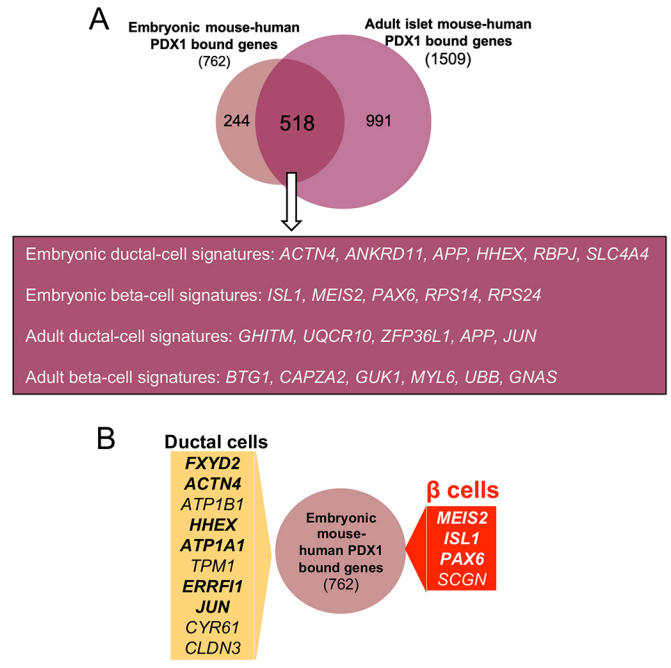

Fig. 5. PDX1 may direct a developmentally and evolutionarily conserved ductal and endocrine program. (A) Venn diagram and representative genes bound by PDX1 in developing pancreas and in adult islet in both mouse and human. (B) Genes bound by PDX1 in both mouse and human developmental pancreas, and expressed at high levels in embryonic ductal cells or β-cells. Genes in bold are also bound by PDX1 in both mouse and human adult islets.

PDX1 may maintain cellular plasticity in adult β-cells by binding the same set of targets from developing pancreas to adult islet

Perhaps the most surprising and intriguing finding was the observation of binding of PDX1 on important acinar and ductal genes from early development to adult islet. During pancreas development, PDX1 is initially expressed in pancreatic progenitors to establish the pancreatic lineage and then gradually enriched in pancreatic β-cells to maintain β-cell function and identity (Pan and Wright, 2011; Jennings et al., 2020). The dramatic changes in PDX1 function could be caused by changes in PDX1 binding. However, although some PDX1 targets were stage specific, as also showed by Wang and colleagues (Wang et al., 2018), we found that PDX1 binding was largely stable from embryonic pancreas to adult islet in both mouse and human. These data indicate that a subset of PDX1-binding events in embryonic progenitors are maintained in the pancreatic islet, which may contribute to adult β-cell plasticity (Remedi and Emfinger, 2016), enabling β-cells to dedifferentiate and acquire a progenitor-like state under stress conditions.

PDX1 interacts with other transcription factors and cofactors to regulate pancreas development

PDX1 has been found to act as both an activator and a repressor in some cell types. In hESC-pancreatic progenitors, PDX1 activates pancreatic genes and represses hepatic genes to establish pancreatic lineages (Teo et al., 2015). For example, two canonical hepatic genes, *FOXA3* (Kaestner et al., 1994) and *HNF4A* (Watt et al., 2003), were bound by PDX1 but not expressed in hESC-PPs and/or human fetal pancreas (Tables S3 and S8). In adult β-cells, PDX1 activates key β-cell functional genes and represses α-cell genes to maintain normal β-cell function and identity (Gao et al., 2014). PDX1 may interact with different transcription factors and/or recruit different co-factors to activate or repress the expression of its targets (Spaeth et al., 2016). Indeed, our upstream regulator analysis

Development (2022) 149, dev200432. doi:10.1242/dev.200432

suggested that a large number of these PDX1 targets were co-regulated by other transcription factors, such as P53 and SOX2, and co-factors, such as LDB1 and EP300, at different stages (Figs S2B, S8D). We have previously identified PDX1 in complex with LDB1 in murine β-cells, providing validation for these bioinformatic predictions (Ediger et al., 2017). This pattern of interaction is remarkably well conserved between mouse and human.

In summary, our study comprehensively examines PDX1 genome-wide targets in mouse and human developing pancreas and reveals a core evolutionarily conserved gene program in the pancreatic islet. PDX1 surprisingly bound endocrine progenitor-, acinar- and duct-specific genes, from early development to adult islet in human pancreas, suggesting that PDX1 maintains a subset of its embryonic binding events in adult β-cells. Our data suggest a protective role of PDX1 in maintaining adult β-cell plasticity to allow adaptation and dedifferentiation under pathological conditions, and also set a valuable foundation for evaluating the efficiency and improving the protocols for human stem cell differentiation to the β-cell fate.

MATERIALS AND METHODS

Animals

Mouse experiments were performed at University of Pennsylvania, with IACUC approval. Timed pregnant CD1 mice were purchased from Charles River Laboratories. Dams were housed in AAALC approved rodent colony. Embryonic pancreata were dissected at E13.5 and E15.5.

Human fetal pancreas samples

A total of seven samples were used for these experiments (three pancreas samples at 14 weeks gestation, and one sample each at 15, 18, 19 and 20 weeks gestation). Gestational age was determined by ultrasound, usually using a crown-rump length. No other information on maternal history was available. Samples were received deidentified, in RPMI/5% FBS/antibiotics media on ice. All samples were processed for ChIP-seq or single cell RNA-seq on the day they were received, as detailed below. Fetal human pancreas tissues were obtained in the period of time between 2011-2017 from StemExpress (Los Angeles, CA, USA) and from Advance Bioscience Resources (Alameda, CA, USA) after elective termination of pregnancy. An informed consent was obtained at the time of the elective termination of pregnancy and is held by supplying company. A sample of the Advance Bioscience Resource consent was obtained and complies with NIH requirements (specifically, donation of human fetal tissues was obtained by someone other than the person who obtained the informed consent for abortion, occurred after the informed consent for abortion and does not affect the method of abortion; no enticements, benefits or financial incentives were used at any level of the process to incentivize abortion or the donation of human fetal tissues; and to be signed by both the woman and the person who obtains the informed consent). Experiments performed on these samples were exempt from IRB evaluation by the University of Pennsylvania and The Children's Hospital of Philadelphia.

ChIP-seq

We performed ChIP-seq for PDX1 in mouse embryonic pancreata (at E13.5 and E15.5) and in human fetal pancreata (at 14 weeks gestational age). We used the goat anti-PDX1 antibody (a gift from Chris Wright, Vanderbilt University, TN, USA; BCBC AB2027). Fifty to 60 individual mouse pancreata were pooled for one chromatin immunoprecipitation replicate. All analysis was performed in three replicates. Chromatin was prepared for individual replicates from three human fetal pancreata. ChIP-seq was performed as previously described (Khoo et al., 2012). All samples were sequenced in the Functional Genomic Core at University of Pennsylvania. The ChIP-seq reads were mapped to the mouse genome (mm10) using Bowtie.

PDX1 binding in hPSC-derived pancreas progenitor cells was obtained from previously published data by the Lickert laboratory (Wang et al., 2018). The PDX1 binding in adult human and mouse islets were previously published by our group (Khoo et al., 2012). After peak calling by HOMER, we continued the analysis with the ∼25% top peaks. For the mouse data, at E13.5 there were 4107 peaks (lowest HOMER score 5.28), at E15.5 were 5343 peaks (lowest HOMER score 5.03) and in mouse islet were 4559 peaks (lowest HOMER score 5.9) (Tables S9, S10). For the human data, in fetal pancreas there were 8183 peaks, representing the top 25% of the peaks (lowest HOMER score 11.1 – corresponding to 6294 genes); in hPSC-PP cells there were 21,292 peaks (lowest HOMER score – 18.9), corresponding to 11,451 genes; and for human adult islet there were 17,593 peaks (lowest HOMER score 18.2) – corresponding to 8307 genes. The majority of PDX1 peaks were within 2000 bp of the transcription start sites (TSS) (Fig. S1C). The data regarding the H3K4me3 and H3K27me3 histone marks in adult α, β and exocrine cells was previously published by Bramswig et al. (2013). Gene ontology analysis was performed with Ingenuity Pathway analysis software (Qiagen). BioVenn and DeepVenn were used to generate the Venn diagrams (Hulsen et al., 2008).

Single-cell RNA-seq and data analysis

Single islet cells were dissociated with TrypLe, and loaded on the 10x Genomics platform in the Center for Applied Genomics at The Children's Hospital of Philadelphia. Single-cell transcriptome libraries were sequenced on the HiSeq platform (Illumina). The cellranger (v.2.1,10× genomics) pipeline was used for barcode filtering, alignment (to GRCh38) and UMI counting. Secondary bioinformatic analysis was performed using the Seurat (v.2.3.4) packages in R. Briefly, sequenced 10× libraries were individually evaluated on multiple criteria to determine cells of interest, including low expression of Hemoglobin (average log2 UMI of HBA1, HBA2, HBB<8) and the number of UMI (0-65536). A total of 7421 fetal cells and 20,764 adult pancreatic cells were examined. Two single-cell RNA-seq samples of fetal and adult were integrated by anchor-based method implemented in Seurat. The top 30 principal components were used to perform clustering and visualization using a UMAP (uniform manifold approximation and projection) plot. We assessed cell type-specific gene expression signatures of acinar, ductal and endocrine cells (Fig. 2A,B). Using cell lineage-specific markers, we identified pancreatic cell clusters for α-cells (*GCG*), β-cells (*INS*), δ-cells (*SST*), PP cells (*PPY*), acinar cells (*PRSS1*) and ductal cells (*CFTR*) (Fig. 2A, Fig. S3). To identify differentially expressed genes between cell populations in fetal and adult cells, a non-parametric Wilcoxon rank sum test was performed, and *P*-values were adjusted using a Benjamini-Hochberg correction based on total number of genes in the dataset.

Human adult islet-cell transcriptomic data have been previously published (GSE114297; Xin et al., 2018). Mouse transcriptomic data of the developing pancreas were from Krentz et al. (2018) and Bastidas-Ponce et al. (2019).

Immunofluorescence

A small proportion of the human fetal pancreas samples at 14 weeks gestation were fixed with 4% PFA overnight, embedded in OCT and stored at −80C. Cryosections were then used for immunofluorescence for PDX1 with goat anti-Pdx1 (Santa Cruz Biotechnology, A-17, 1:500), rabbit anti-onecut1 (Santa Cruz Biotechnology, H-100, 1:1000) and guinea-pig anti-insulin (Abcam, ab7842, 1:500).

Acknowledgements

We thank Dr. Mitchell Lazar for his comments on our manuscript, and Dr Chris Wright for PDX1 antiserum and for his comments on our manuscript. We thank the University of Pennsylvania Diabetes Research Center (DRC) for the use of the Functional Genomics Core (P30-DK19525) and thank the Center for Applied Genomics at The Children's Hospital of Philadelphia.

Competing interests

The authors declare no competing or financial interests.

Author contributions

Conceptualization: J.C.R., D.E.S., D.A.S.; Formal analysis: X.Y., J.C.R., J.K., R.Y., G.R., K.-J.W., D.E.S.; Investigation: J.C.R., J.Y., C.L., D.E.S.; Writing - original draft: X.Y., D.E.S., D.A.S.; Writing - review & editing: X.Y., K.-J.W., D.E.S., D.A.S.; Supervision: D.A.S.; Funding acquisition: D.A.S.

Funding

J.C.R. was supported by the National Institutes of Health (F32-DK-089747). D.E.S. was supported by an American Diabetes Association Junior Faculty Development Award (1-16-JDF086). D.A.S. was supported by the National Institutes of Health (R01DK105689 and R01DK121175). Deposited in PMC for release after 12 months.

Data availability

All mouse embryonic and human fetal PDX1 ChIP-seq data generated in this study are available in ArrayExpress (E-MTAB-3354). The human fetal pancreas single-cell RNA-seq data generated in this study have been deposited in GEO under accession number GSE201230.

References

Abou-Kheir, W., Mukherji, D., Hadadeh, O., Saleh, E., Bahmad, H. F., Kanso, M., Khalifeh, M., Shamseddine, A., Tamraz, S., Jaafar, R. et al. (2020). CYR61/CCN1 expression in resected pancreatic ductal adenocarcinoma: A retrospective pilot study of the interaction between the tumors and their surrounding microenvironment. *Heliyon* **6**, e03842. doi:10.1016/j.heliyon.2020.e03842

Ashery-Padan, R., Zhou, X., Marquardt, T., Herrera, P., Toube, L., Berry, A. and Gruss, P. (2004). Conditional inactivation of Pax6 in the pancreas causes early onset of diabetes. *Dev. Biol.* **269**, 479-488. doi:10.1016/j.ydbio.2004.01.040

Bastidas-Ponce, A., Tritschler, S., Dony, L., Scheibner, K., Tarquis-Medina, M., Salinno, C., Schirge, S., Burtscher, I., Bottcher, A., Theis, F. J. et al. (2019). Comprehensive single cell mRNA profiling reveals a detailed roadmap for pancreatic endocrinogenesis. *Development* **146**, dev173849. doi:10.1242/dev.173849

Bramswig, N. C., Everett, L. J., Schug, J., Dorrell, C., Liu, C., Luo, Y., Streeter, P. R., Naji, A., Grompe, M. and Kaestner, K. H. (2013). Epigenomic plasticity enables human pancreatic α to β cell reprogramming. *J. Clin. Invest.* **123**, 1275-1284. doi:10.1172/JCI66514

Du, A., Hunter, C. S., Murray, J., Noble, D., Cai, C. L., Evans, S. M., Stein, R. and May, C. L. (2009). Islet-1 is required for the maturation, proliferation, and survival of the endocrine pancreas. *Diabetes* **58**, 2059-2069. doi:10.2337/db08-0987

Ediger, B. N., Lim, H. W., Juliana, C., Groff, D. N., Williams, L. T., Dominguez, G., Liu, J. H., Taylor, B. L., Walp, E. R., Kameswaran, V. et al. (2017). LIM domain-binding 1 maintains the terminally differentiated state of pancreatic β cells. *J. Clin. Invest.* **127**, 215-229. doi:10.1172/JCI88016

Gao, T., McKenna, B., Li, C., Reichert, M., Nguyen, J., Singh, T., Yang, C., Pannikar, A., Doliba, N., Zhang, T. et al. (2014). Pdx1 maintains β cell identity and function by repressing an α cell program. *Cell Metab.* **19**, 259-271. doi:10.1016/j.cmet.2013.12.002

Guz, Y., Montminy, M. R., Stein, R., Leonard, J., Gamer, L. W., Wright, C. V. and Teitelman, G. (1995). Expression of murine STF-1, a putative insulin gene transcription factor, in beta cells of pancreas, duodenal epithelium and pancreatic exocrine and endocrine progenitors during ontogeny. *Development* **121**, 11-18. doi:10.1242/dev.121.1.11

Hani, E. H., Stoffers, D. A., Chèvre, J. C., Durand, E., Stanojevic, V., Dina, C., Habener, J. F. and Froguel, P. (1999). Defective mutations in the insulin promoter factor-1 (IPF-1) gene in late-onset type 2 diabetes mellitus. *J. Clin. Invest.* **104**, R41-R48. doi:10.1172/JCI7469

Hulsen, T., de Vlieg, J. and Alkema, W. (2008). BioVenn - a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC Genomics* **9**, 488. doi:10.1186/1471-2164-9-488

Jennings, R. E., Berry, A. A., Kirkwood-Wilson, R., Roberts, N. A., Hearn, T., Salisbury, R. J., Blaylock, J., Piper Hanley, K. and Hanley, N. A. (2013). Development of the human pancreas from foregut to endocrine commitment. *Diabetes* **62**, 3514-3522. doi:10.2337/db12-1479

Jennings, R. E., Scharfmann, R. and Staels, W. (2020). Transcription factors that shape the mammalian pancreas. *Diabetologia* **63**, 1974-1980. doi:10.1007/s00125-020-05161-0

Jonsson, J., Carlsson, L., Edlund, T. and Edlund, H. (1994). Insulin-promoter-factor 1 is required for pancreas development in mice. *Nature* **371**, 606-609. doi:10.1038/371606a0

Kaestner, K. H., Hiemisch, H., Luckow, B. and Schütz, G. (1994). The HNF-3 gene family of transcription factors in mice: gene structure, cDNA sequence, and mRNA distribution. *Genomics* **20**, 377-385. doi:10.1006/geno.1994.1191

Kaihara, K. A., Dickson, L. M., Jacobson, D. A., Tamarina, N., Roe, M. W., Philipson, L. H. and Wicksteed, B. (2013). β-Cell-specific protein kinase A activation enhances the efficiency of glucose control by increasing acute-phase insulin secretion. *Diabetes* **62**, 1527-1536. doi:10.2337/db12-1013

Kaneko, K., Ueki, K., Takahashi, N., Hashimoto, S., Okamoto, M., Awazawa, M., Okazaki, Y., Ohsugi, M., Inabe, K., Umehara, T. et al. (2010). Class IA phosphatidylinositol 3-kinase in pancreatic β cells controls insulin secretion by multiple mechanisms. *Cell Metab.* **12**, 619-632. doi:10.1016/j.cmet.2010.11.005

Keller, D. M., McWeeney, S., Arsenlis, A., Drouin, J., Wright, C. V., Wang, H., Wollheim, C. B., White, P., Kaestner, K. H. and Goodman, R. H. (2007). Characterization of pancreatic transcription factor Pdx-1 binding sites using promoter microarray and serial analysis of chromatin occupancy. *J. Biol. Chem.* **282**, 32084-32092. doi:10.1074/jbc.M700899200

Khoo, C., Yang, J., Weinrott, S. A., Kaestner, K. H., Naji, A., Schug, J. and Stoffers, D. A. (2012). Research resource: the pdx1 cistrome of pancreatic islets. *Mol. Endocrinol.* **26**, 521-533. doi:10.1210/me.2011-1231

Krentz, N. A. J., Lee, M. Y. Y., Xu, E. E., Sproul, S. L. J., Maslova, A., Sasaki, S. and Lynn, F. C. (2018). Single-cell transcriptome profiling of mouse and hESC-derived pancreatic progenitors. *Stem Cell Reports* **11**, 1551-1564. doi:10.1016/j.stemcr.2018.11.008

Liu, J. S. and Hebrok, M. (2017). All mixed up: defining roles for β-cell subtypes in mature islets. *Genes Dev.* **31**, 228-240. doi:10.1101/gad.294389.116

Lorberbaum, D. S., Kishore, S., Rosselot, C., Sarbaugh, D., Brooks, E. P., Aragon, E., Xuan, S., Simon, O., Ghosh, D., Mendelsohn, C. et al. (2020). Retinoic acid signaling within pancreatic endocrine progenitors regulates mouse and human β cell specification. *Development* **147**, dev189977. doi:10.1242/dev.189977

Malenczyk, K., Szodorai, E., Schnell, R., Lubec, G., Szabó, G., Hökfelt, T. and Harkany, T. (2018). Secretagogin protects Pdx1 from proteasomal degradation to control a transcriptional program required for β cell specification. *Mol. Metab.* **14**, 108-120. doi:10.1016/j.molmet.2018.05.019

Offield, M. F., Jetton, T. L., Labosky, P. A., Ray, M., Stein, R. W., Magnuson, M. A., Hogan, B. L. and Wright, C. V. (1996). PDX-1 is required for pancreatic outgrowth and differentiation of the rostral duodenum. *Development* **122**, 983-995. doi:10.1242/dev.122.3.983

Pan, F. C. and Wright, C. (2011). Pancreas organogenesis: from bud to plexus to gland. *Dev. Dyn.* **240**, 530-565. doi:10.1002/dvdy.22584

Perelis, M., Marcheva, B., Ramsey, K. M., Schipma, M. J., Hutchison, A. L., Taguchi, A., Peek, C. B., Hong, H., Huang, W., Omura, C. et al. (2015). Pancreatic β cell enhancers regulate rhythmic transcription of genes controlling insulin secretion. *Science* **350**, aac4250. doi:10.1126/science.aac4250

Ramond, C., Beydag-Tasöz, B. S., Azad, A., van de Bunt, M., Petersen, M. B. K., Beer, N. L., Glaser, N., Berthault, C., Gloyn, A. L., Hansson, M. et al. (2018). Understanding human fetal pancreas development using subpopulation sorting, RNA sequencing and single-cell profiling. *Development* **145**, dev165480. doi:10.1242/dev.165480

Raum, J. C., Soleimanpour, S. A., Groff, D. N., Coré, N., Fasano, L., Garratt, A. N., Dai, C., Powers, A. C. and Stoffers, D. A. (2015). Tshz1 Regulates pancreatic β-cell maturation. *Diabetes* **64**, 2905-2914. doi:10.2337/db14-1443

Remedi, M. S. and Emfinger, C. (2016). Pancreatic β-cell identity in diabetes. *Diabetes Obes. Metab.* **18** Suppl. 1, 110-116. doi:10.1111/dom.12727

Sharon, N., Vanderhooft, J., Straubhaar, J., Mueller, J., Chawla, R., Zhou, Q., Engquist, E. N., Trapnell, C., Gifford, D. K. and Melton, D. A. (2019). Wnt Signaling separates the progenitor and endocrine compartments during pancreas development. *Cell Rep.* **27**, 2281-2291. doi:10.1016/j.celrep.2019.04.083

Spaeth, J. M., Walker, E. M. and Stein, R. (2016). Impact of Pdx1-associated chromatin modifiers on islet β-cells. *Diabetes Obes. Metab.* **18** Suppl. 1, 123-127. doi:10.1111/dom.12730

Stoffers, D. A., Ferrer, J., Clarke, W. L. and Habener, J. F. (1997a). Early-onset type-II diabetes mellitus (MODY4) linked to IPF1. *Nat. Genet.* **17**, 138-139. doi:10.1038/ng1097-138

Stoffers, D. A., Zinkin, N. T., Stanojevic, V., Clarke, W. L. and Habener, J. F. (1997b). Pancreatic agenesis attributable to a single nucleotide deletion in the human IPF1 gene coding sequence. *Nat. Genet.* **15**, 106-110. doi:10.1038/ng0197-106

Swift, G. H., Liu, Y., Rose, S. D., Bischof, L. J., Steelman, S., Buchberg, A. M., Wright, C. V. and MacDonald, R. J. (1998). An endocrine-exocrine switch in the activity of the pancreatic homeodomain protein PDX1 through formation of a trimeric complex with PBX1b and MRG1 (MEIS2). *Mol. Cell. Biol.* **18**, 5109-5120. doi:10.1128/MCB.18.9.5109

Teo, A. K., Tsuneyoshi, N., Hoon, S., Tan, E. K., Stanton, L. W., Wright, C. V. and Dunn, N. R. (2015). PDX1 binds and represses hepatic genes to ensure robust pancreatic commitment in differentiating human embryonic stem cells. *Stem Cell Reports* **4**, 578-590. doi:10.1016/j.stemcr.2015.02.015

Wang, A., Yue, F., Li, Y., Xie, R., Harper, T., Patel, N. A., Muth, K., Palmer, J., Qiu, Y., Wang, J. et al. (2015). Epigenetic priming of enhancers predicts developmental competence of hESC-derived endodermal lineage intermediates. *Cell Stem. Cell* **16**, 386-399. doi:10.1016/j.stem.2015.02.013

Wang, X., Sterr, M., Burtscher, I., Chen, S., Hieronimus, A., Machicao, F., Staiger, H., Häring, H. U., Lederer, G., Meitinger, T. et al. (2018). Genome-wide analysis of PDX1 target genes in human pancreatic progenitors. *Mol. Metab.* **9**, 57-68. doi:10.1016/j.molmet.2018.01.011

Watt, A. J., Garrison, W. D. and Duncan, S. A. (2003). HNF4: a central regulator of hepatocyte differentiation and function. *Hepatology* **37**, 1249-1253. doi:10.1053/jhep.2003.50273

Westmoreland, J. J., Drosos, Y., Kelly, J., Ye, J., Means, A. L., Washington, M. K. and Sosa-Pineda, B. (2012). Dynamic distribution of claudin proteins in pancreatic epithelia undergoing morphogenesis or neoplastic transformation. *Dev. Dyn.* **241**, 583-594. doi:10.1002/dvdy.23740

Wu, H. and Sun, Y. E. (2006). Epigenetic regulation of stem cell differentiation. *Pediatr. Res.* **59**, 21R-25R. doi:10.1203/01.pdr.0000203565.76028.2a

Wu, L., Zhou, L., Lu, Y., Zhang, J., Jian, F., Liu, Y., Li, F., Li, W., Wang, X. and Li, G. (2012). Activation of SIRT1 protects pancreatic β-cells against

palmitate-induced dysfunction. *Biochim. Biophys. Acta* **1822**, 1815-1825. doi:10.1016/j.bbadis.2012.08.009

Xin, Y., Dominguez Gutierrez, G., Okamoto, H., Kim, J., Lee, A. H., Adler, C., Ni, M., Yancopoulos, G. D., Murphy, A. J. and Gromada, J. (2018). Pseudotime ordering of single human β-cells reveals states of insulin production and unfolded protein response. *Diabetes* **67**, 1783-1794. doi:10.2337/db18-0365

Yaney, G. C. and Corkey, B. E. (2003). Fatty acid metabolism and insulin secretion in pancreatic beta cells. *Diabetologia* **46**, 1297-1312. doi:10.1007/s00125-003-1207-4

Yu, X. X. and Xu, C. R. (2020). Understanding generation and regeneration of pancreatic β cells from a single-cell perspective. *Development* **147**, dev179051. doi:10.1242/dev.179051

Yu, X. X., Qiu, W. L., Yang, L., Wang, Y. C., He, M. Y., Wang, D., Zhang, Y., Li, L. C., Zhang, J., Wang, Y. et al. (2021). Sequential progenitor states mark the generation of pancreatic endocrine lineages in mice and humans. *Cell Res.* **31**, 886-903. doi:10.1038/s41422-021-00486-w

Zhang, J., McKenna, L. B., Bogue, C. W. and Kaestner, K. H. (2014). The diabetes gene Hhex maintains δ-cell differentiation and islet function. *Genes Dev.* **28**, 829-834. doi:10.1101/gad.235499.113